Deliverable

D1.1 Legal Landscape and Requirements Plan

Project Acronym:	DUET	
Project title:	Digital Urban European Twins	
Grant Agreement No.	870697	
Website:	www.digitalurbantwins.eu	
Version:	1.0	
Date:	2020	
Responsible Partner:	AEG	
Contributing Partners:	GSL	
Reviewers:	Nils Walravens (IMEC) Lieven Raes (AIV) Andrew Stott (external)	
Dissemination Level:	Public	X
	Confidential – only consortium members and European Commission	

This project has received financial support from the European Union's Horizon 2020 Programme under grant agreement no. 870697

Revision History

Revision	Date	Author	Organization	Description
0.1	27 February 2020	Kletia Noti	GSL	Initial structure
0.2	3 April 2020	Kletia Noti, Valeria Comegna	GSL	First version
0.3	28 April 2020	Geert Mareels, Andrew Stott	AIV	Review
0.4	2 May 2020	Kletia Noti, Tomáš Pavelka, Valeria Comegna	GSL	Second version
0.5	24 May 2020	Tomáš Pavelka	GSL	Review
0.6	2 June 2020	Valeria Comegna	GSL	Pre-final version
0.7	10 June 2020	Nils Walravens	IMEC	Review
1.0	12 June 2020	Valeria Comegna	GSL	Final version

Table of Contents

List of Abbreviations

BEREC	Body of European Regulators for Electronic Communications
Charter	Charter of Fundamental Rights of the European Union
Cybersecurity Act	Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013
CFR	Charter of Fundamental Rights of the European Union
CSIRT	Computer Security Incident Response Team
CJEU	Court of Justice of the European Union
DPA	Data Protection Authority
DPIA	Data Protection Impact Assessment
DPO	Data Protection Officer
DUET	Digital Urban European Twins
ECHR	Convention for the Protection of Human Rights and Fundamental Freedoms
ECtHR	European Court of Human Rights
EDPS	European Data Protection Supervisor
EDPB	European Data Protection Board
ENISA	European Union Agency for Cyber-security
GDPR	Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/ (General Data Protection Regulation)
ICO	UK Information Commissioner's Office
NIS Directive	Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union (Directive on security of networks and information systems)
NRA	National Regulatory Authority
Open Data Directive	Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information
PSI Directive	Directive 2013/37/EU of the European Parliament and of the Council of 26 June 13 20131 amending Directive 2003/98/EC on the re-use of public sector information

Executive Summary

This document outlines the legal and ethical requirements/plan for the DUET project, aimed at supporting the activities of the DUET Digital Twins consortium. It addresses the requirements of D.1.1 under the Grant Agreement entered into with the European Commission. As the project continues, these legal requirements should be continuously checked for feasibility, fine-tuned and updated accordingly, in the light of the evolving legal landscape at EU level and across Member States.

Data is the backbone of the smart city of tomorrow. In a complex ecosystem where IoT enables collecting and analysing data on usage patterns to benefit residents and their communities, and provides policy makers with tools to take agile, real-time decisions, unleashing the potential of data as a valuable resource requires creating and furthering trust and confidence in technologies that are vital to the uptake of smart cities. In turn, making use of data to unleash such potential requires the law to act as an enabler. It also requires organisations handling data, including personal data, to put in place adequate ethical mechanisms abiding by the applicable EU and national laws. In addition, when the applicable legal framework does not provide clear-cut guidance or does point at legal gaps, this process also requires having recourse to soft law issued by competent authorities, and best practices, at times in the context of existing smart city practices, effectively to deal with these gaps and grey areas.

We have identified several main cluster areas of the Legal and Ethics landscape and requirements for the **DUET** project. First, we tackle what we will hereinafter call **data governance**, which boils down to the privacy aspects of data sharing that impact DUET's Digital Twin and its component parts (**Chapter 2**). After a brief overview of the purpose of the Chapter, we identify privacy-related concerns and gaps, and also outline risk mitigation aspects linked to each of those identified concerns. Among the main concerns is DUET's compliance with the EU legislation on privacy protection when the data it handles falls under the notion of "personal data". To this end, we have drawn on operational guidance from the European Data Protection Supervisor (EDPS) and the European Data Protection Board (EDPB). We have also, among others, focused on certain types of data of potential relevance to DUET, such as location data (to which several pieces of legislation, such as the GDPR or the ePrivacy Directive, apply concurrently), and video-surveillance data.

The challenges of **data security,** and a sound plan for risk mitigation are discussed under **Chapter 3**. The legal framework encompasses the various pieces of legislation concerning the different layers where security risks may occur, as well as the related soft law instruments. We also look at literature highlighting what the risks may be. Finally, looking at best practices from the European Cloud Security Alliance, we also draw some guidance on what elements a sound risk mitigation plan could contain.

Across the two Chapters, we refer also to the **ethical aspects** of the technologies that underpin the project - spanning from the IoT, Artificial Intelligence (AI), to their intersection, big data analytics, as well as the cloud. In particular, the legal overview will also point at gaps in the regulation of those technologies in relation to the smart city context. Moreover, such ethical aspects are dealt with in **Chapter 4 through a high-level analysis** which purports to illustrate as well as address those ethical issues beyond data protection and data security, i.e. IoT and AI safety, liability, IP, trade secrets, as well as open data aspects. We will also graphically sketch out a number of potential risks and related mitigation measures: for practical reasons, IoT and AI will be analysed separately. Finally, **Chapter 5** draws preliminary conclusions in the light of the findings contained in the preceding chapters. In the **Annex,** we briefly outline certain data governance related issues of the Sidewalk Toronto Project as an example of an existing smart city initiative, now abandoned, with the aim of

learning from this experience, scoping and analysing the thorniest legal challenges that cities powered by technology may present.

1. Introduction

The scope of this deliverable is to shed light on both the legal landscape that underpins the DUET Digital Twins consortium activities and the requirements plan with the aim of providing a framework in compliance with EU and, where applicable, national legislation. As such, this deliverable purports to answer, in this initial phase, the requirements of Task 1.1 under WP 1 in the Grant Agreement.

In particular, the deliverable aims at providing an overview of the legal landscape that underpins the project activities, as well as an analysis of the risks. It also addresses how to operationally mitigate those risks while ensuring ethics by design: we look at both privacy, and security concerns, as well as other potentially relevant ethical aspects of the project. The 'by design' approach requires laying out risk mitigation procedures prior to the DUET architecture being built, rather than as an *ex post* compliance tool: such procedures will then span throughout the whole life of the project.

In the context of smart cities[1], most of the literature comes from technological and urban studies, looking at the issues mostly from an environmental and sociological rather than a legal standpoint. Such literature has placed emphasis on the social, urban, policy-making and environmental benefits of smart cities, rather than their challenges in terms of the ethical and legal conundra they present. The word 'smart city' has become a buzzword referring to a concept often presented as a panacea to solve the difficulties of the twenty-first century accelerated urbanisation. Most of the legal literature analysing the concerns that the phenomenon of smart cities gives rise to comes from outside the EU context. It is only over the past years that a growing number of EU scholars have started to analyse some of the crucial issues from a legal and ethics standpoint. Edwards, a leading academic in the field of Internet Law, for example, highlighted "*the lack of opportunity in an ambient or smart city environment for the giving of meaningful consent to processing of personal data; the degree to which smart cities collect private data from inevitable public interactions, the "privatisation" of ownership of both infrastructure and data, the repurposing of "big data" drawn from IoT in smart cities and the storage of that data in the Cloud*".

Against this background, this deliverable is organised as follows: first, in **Chapter 2**, we provide an overview of the potential data governance aspects that the project may entail. In particular, we look at the legislation applicable when personal data are processed by DUET, and the risks that such personal data handling could entail. We also analyse the EU legal framework (including when laid out in soft law) when data other than personal data is handled by DUET. We also look at the applicable risk mitigation mechanisms with the help of guidance from EU bodies such as the European Data Protection Board (EDPB) or the European Data Protection Supervisor (EDPS), as well as, as the case may be, national data protection authorities. We also analyse, in terms of specific risks and risk mitigations, to certain specific data which is of relevance for the DUET project, such as location data, video-surveillance data and also data which belongs to mixed data sets.

Second, in **Chapter 3** we delve into the specific aspects concerning data and also software and hardware security that the DUET project may be confronted with. First, we look at the legal landscape that underpins cybersecurity in the EU, and analyse the various pieces of legislation and soft law, including soft law issued by

[1] Edwards, L., "Privacy, Security and Data Protection in Smart Cities: a critical EU law perspective", CREATe working paper series, (2015).

the competent agency of the European Union, the European Union Agency for Cyber-security (ENISA). Second, we pinpoint the possible risks that may arise, as described by the main stakeholders in the industry. Third, drawing from sector-specific guidance, we provide some procedural considerations in terms of risk mitigations. We explore the notion of 'security by design' and we also look into what this notion entails for DUET. Finally, in **Chapter 4** we look at the ethical aspects beyond privacy and (cyber)security that may be relevant during the project, with particular regard for technologies such as IoT and AI. In the specific, we look at the potential liability aspects that the DUET consortium could be exposed to and provide risk mitigation plans in this respect.

When scoping potential legal issues, we have sought to learn lessons by looking at other Smart Cities initiatives, including analysing in more detail the now halted Sidewalk Toronto project (**Annex 1**). Sidewalk Toronto is an urban development project operated until May 2020 by Sidewalk Labs, an Alphabet (Google) subsidiary, at Quayside, a waterfront area in Toronto, Ontario, Canada. In charge of steering the project in line with public interest is Waterfront Toronto, a body created by the governments of Canada, Ontario, and the City of Toronto.

After winning a request for proposals in October 2017, Sidewalk Labs committed USD 50 million to test pilot projects. In June 2019, it published the Master Innovation Development Plan, a detailed set of project documentation. Sidewalk Labs has withdrawn from the project as of 7 May 2020. From the outset, data governance and privacy had raised public and privacy experts' concerns, and eventually became threshold issues in deciding whether the project would move forward at all. Even if Sidewalk Labs' official statement quotes the "*unprecedented economic uncertainty*" around Covid-19 as the reason for withdrawal, it is likely that the project was withdrawn due to the privacy concerns it raised. Prior to the withdrawal, one of the key pillars for reaching an agreement on data governance and privacy management was the envisaged adherence by all stakeholders (including Sidewalk Labs) to emerging, but not yet published, "Intelligent Community Guidelines". Such guidelines were a set of rules combining input from government stakeholders, industry and the broader community on digital governance issues and privacy, which was supposed to be enforceable against private parties (including Sidewalk Labs) through contract. Despite the project having been ultimately abandoned, the concerns it gave rise to may be of interest for DUET as a lesson learning experience.

2. Data Governance

2.1 Purpose

The purpose of this Chapter is to highlight the main privacy-related concerns with respect to the data employed in the context of the three technologies that intersect in the DUET project, namely the IoT, big data and the cloud, as well as their combined use. To this end, an overview of the legal landscape under EU law and, where applicable, national law will be provided. Among others, both difficulties of interpretation and legal gaps could be identified under EU law. To the extent necessary, reference to soft law and guidance clarifying the applicable legal instruments shall be made.

Data analytics and IoT are transforming connected cities into smart cities: while the data stream that the smart city unleashes is enormous in terms of potential to improve city services and citizens' lives, these developments have given rise to a debate on data ownership and the use of data. This requires compliance with national and EU-wise data privacy laws, like the EU's Regulation 2016/679 (GDPR). Yet, in the context of the EU legal framework, given the adoption of the GDPR, enhanced rights of data subjects may bring about challenges for the development of smart cities.

One example is the static notion of personal data under the GDPR. Data comes in various shapes and often in complex datasets. While the law puts them in boxes of "personal data", falling under the scope of the GDPR, or "non-personal data" which is subject to a separate legal framework under EU law, it is worth remembering that, in practice, this categorisation does not reflect the dynamic nature of the data that DUET will deal with. Indeed, the big data that DUET will deal with also probably mostly feed in mixed data sets. In this respect, EU data protection laws and, *in primis*, the Regulation on the Free Flow of non-personal data, and its interplay with the GDPR, as interpreted by the European Commission, may provide some guidance. While the distinction appears straightforward, in practice it is not: the Chapter also highlights some of the gaps that may arise in terms of legal framework when dynamic data is taken into account, and how to minimise risks associated with them.

Besides the concerns related to data from a dynamic standpoint, and the difficulties that a static legal framework may entail, the GDPR has laid down, as will be seen below, some principles which may raise practical difficulties in a smart city context: consider, for example, the impracticality of always obtaining 'informed and freely given consent'. There are indeed difficulties to obtain, each time, GDPR-compliant consent by the data subjects to handle their personal data, as real time, granular data are harnessed thanks to the new technologies employed. In this case, alternative grounds for lawful processing under the GDPR, when the data is personal data, must be identified. This comes with its own set of challenges. In addition, under other EU legislation such as the ePrivacy Directive, certain difficulties when it comes to obtaining consent to process electronic communications data arise: these concerns touch upon the DUET project as data in rest and in transit falls under its scope.

Consider also, for instance, another GDPR-related principle such as purpose limitation: the principle under which personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes. Yet, some authors observe, in the context of smart cities, purpose limitation may appear antithetical to the nature of big data[2]. The GDPR does not ban

[2] Kitchin, R., "Big Data, New epistemologies and paradigm shifts", in Big Data & Society, (2014).

the use of information for purposes other than those for which it was collected, provided that certain safeguards are abided by. However, those safeguards also give rise to certain grey areas. While data can be used by municipalities to carry out mobility improvements, they can be relied on by private actors, such as those providing mobility and ridesharing services. This raises the issue of data use and re-use. What are, in this respect, the safeguards and the data subject's access rights? Finally how can it be ensured that data is stored in a GDPR-compliant way, and what are the protections that apply in terms of accountability, integrity, confidentiality, monitoring and auditability throughout the data lifecycle (collection, storage, use and re-use, sharing and disposal)?

At the outset, definitions, such as 'data governance', 'privacy', 'privacy by design', 'data sharing' carried out during the project activities, are worth clarifying.

By 'data governance', this Section refers to the framework to manage the various categories of data (including personal data), taking into account the policies, procedures and systems to employ to ensure that data is properly protected, that risks are adequately managed and that the privacy legislation applicable, when at stake is personal data, is complied with[3]. As such, in accordance with the OECD's interpretation[4], we refer to data governance as a synonym for 'data management framework', including both the ethical and legal concerns of data management, in the context of the IoT platform environment and the ecosystem that characterises DUET Digital Twin's evolution.

The notion of 'privacy' relates to the right of individuals to control the collection, processing, and use (including re-use) of their personal data, as clarified in EU data protection laws (including of a constitutional order). The right to privacy is enshrined under Art. 8 of the Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR)[5], Art. 7 of the Charter of Fundamental Rights of the European Union (CFR)[6], as well as the abovementioned GDPR, which puts the focus on enhanced rights of data subjects. When it comes to the right to privacy, in the civil law tradition, the legal protection of one's own image is linked to it, and such right to one's image is also enshrined under the scope of Article 8 of the ECHR.

The term 'privacy by design'[7] has been coined by Lessig in relation to both hardware and software and refers to building privacy into the code[8]. This principle is now enshrined under Art. 25 of the GDPR.

By 'sharing' we intend, even if this notion is not spelled out in the GDPR[9], sharing of personal data by organisations within the EU, including with third party organisations (other administrations but also commercial actors). For example, in the UK, Chapter 5 of Part 5 of the Digital Economy Act facilitates the linking and sharing of de-identified data by public authorities.

In the light of the above, it becomes crucial to take steps to create, review and continuously enhance, as the project unfolds, data governance. Such governance would ensure accountability through clear roles for data

[3] OECD, "OECD Privacy Principles", (2013), available at: http://oecdprivacy.org/.
[4] OECD, "Report: "Enhancing Access to and sharing of data: Reconciling risks and benefits for data re-use across societies", (2019).
[5] Council of Europe, Convention for the Protection of Human Rights and Fundamental Freedoms (European Convention on Human Rights, as Amended) ECHR, (1950).
[6] Charter of Fundamental Rights of the European Union (2016) Official Journal C202 (hereinafter, Charter).
[7] Lessig, L., "Code 2.0", in Basic Books, (2006).
[8] Edwards, cited.
[9] Christophi, A., "Sharing personal data in smart cities", Citip conferences (2019).

controllers and processors, data ethics frameworks and periodic audits concerning the controls, together with periodic privacy impact assessments.

Good data stewardship is therefore needed. We first provide, under **Section 2.2**, a brief overview of the applicable legal framework at EU level. Then under **Section 2.3**, we will move onto some privacy-related concerns we have identified as well as the risk mitigation measures to be followed in this respect. We herewith provide a plan for a sound and ethical governance of data: this concept, in turn, relates to the notion of data protection by design, which refers to the need for any action involving processing of personal data being done with data protection in mind. This notion is in more details explored in r **Section 2.3**. **Section 2.4** will draw some preliminary conclusions.

2.2 Legal landscape

This Section first focuses on EU legislation as it now stands (**Section 2.2.1**), but also the extent to which national legislation complements such EU law (**Section 2.2.2**). In addition, it is also important to mention EU-wide legislation proposals concerning future pieces of legislation, such as the Data Act announced by the European Commission in 2020 or the upcoming ePrivacy Regulation (**Section 2.2.3**). We also refer to, where relevant, the main soft-law instruments which may be of use for the present analysis (**Section 2.2.4**).

2.2.1 Current EU data protection legal landscape

The EU data governance legal landscape encompasses several pieces of legislation, the most important among which are the General Data Protection Regulation, the Regulation on the Free Flow of non-personal data, the Open Data Directive, the Data Protection Law Enforcement Directive, the ePrivacy Directive and the Data Protection Regulation for EU institutions and bodies[10]. This framework is to be complemented by the ePrivacy Regulation which is currently pending in the legislative process.

The **General Data Protection Regulation**[11] represents the milestone of EU data protection, which lays down the general principles, rules and procedures to be followed by any actors involved in personal data processing *lato sensu* (thereby including collection, storage, use, re-use, or sharing of personal data) related to individuals-natural persons residing in the EU. Notably, the **principle of free movement of personal data**, under which the free flow of personal data within the EU shall neither be restricted nor prohibited is laid down under Art. 1(3) thereof. In terms of material scope, Art. 2(1) provides that the Regulation applies to the processing of personal data wholly or partly by automated means. It is thus relevant in terms of DUET activities, where big data is at stake.

The GDPR clarifies what constitutes personal data, identifies the main actors and sets out rights and obligations upon them. It lays down rules and procedures for processing personal data. In particular, the

[10] The specific piece of legislation will not be illustrated here, as - for the time being - the team presumes it falls outside of the scope of this project.
[11] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC [2016] OJ L 119/1 (hereinafter, GDPR).

GDPR outlines six data protection principles that organisations shall abide by when collecting, processing and storing individuals' personal data:

- personal data must be processed in a lawful and transparent manner, ensuring fairness towards the individuals whose personal data is being processed ('lawfulness, fairness and transparency');
- there must be specific purposes for processing the data and the company/organisation must indicate those purposes to individuals when collecting their personal data. A company/organisation cannot simply collect personal data for undefined purposes and further use the personal data for other purposes that are not compatible with the original purpose, except for under certain narrow circumstances ('purpose limitation');
- the company/organisation may collect and process only personal data to the extent that is necessary to fulfil that purpose ('data minimisation');
- the company/organisation must ensure personal data is accurate and up-to-date, having regard to the purposes for which it is processed, and rectify the data where it is not ('accuracy');
- the company/organisation must ensure that personal data is stored for no longer than necessary for the purposes for which it was collected ('storage limitation');
- the company/organisation must install appropriate technical and organisational safeguards that ensure the security of personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technology ('integrity and confidentiality').

Further to facilitate cross-border exchange of data and boost the data economy, in November 2018 the European Parliament and the Council adopted the **'Free Flow of Non-Personal Data Regulation'**[12]. This instrument is based on the **principle of free flow of non-personal data**. Under this principle, Member States are prohibited from imposing requirements on where data should be localised in all cases, unless such requirements are proportionate and duly justified by public security. On top of this, Member States have to communicate any data localisation restrictions to the European Commission on a single online information point readily available for users and service providers. In addition, the Regulation also lays down the **principle of data availability**, rendering competent authorities able to access data for supervisory control wherever it is stored or processed in the EU. Finally, as will be better seen under Chapter 3, the Regulation enshrines the principle of data portability, which allows for users to port data between cloud service providers. To this encourages **cloud service providers and cloud users jointly to develop codes of conduct** based on the principles of transparency and openness, that will make it easier to switch cloud service providers. This should make the European cloud market more competitive and lead to lower prices. The codes must be developed and implemented by mid-2020. The Regulation also clarifies that any security requirements that already apply to businesses storing and processing data will also apply when they store or process data across EU borders or in the cloud.

As a part of the **EU Open Data policy**, rules have been adopted to encourage Member States to facilitate the re-use of data from the public sector with minimal or no legal, technical or financial constraints. In full compliance with the GDPR, the new **Open Data Directive**[13] updates the framework setting out conditions under which public sector data should be made available for re-use, with a particular focus on the increasing amounts of high-value data that is now available. The Directive replaces the Public Sector Information

[12] Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union, [2018], OJ L 303.
[13] Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information [2019], OJ L 172.

Directive, also known as the 'PSI Directive'[14], which dated from 2003 and it was subsequently amended in 2013[15]. The Open Data Directive entered into force on 16 July 2019 and the Member States will have to transpose it by 16 July 2021.

Recital 16 of the Open Data Directive defines open data as "*data in an open format that can be freely used, re-used and shared by anyone for any purpose*", "*for private or commercial purposes, with minimal or no legal, technical or financial constraints, and which promote the circulation of information not only for economic operators but primarily for the public*". Such data can span anything from anonymised personal data on household energy use to more general information about national education or literacy levels. The same recital provides that Member States are encouraged to promote the creation of data based on the principle of '**open by design and by default**' with regard to all documents falling within its scope. This should be done while guaranteeing a consistent level of protection of public interest objectives, such as public security, including where sensitive critical infrastructure protection-related information is concerned. Member States are required to ensure protection of personal data, "*including where information in an individual data set does not present a risk of identifying or singling out a natural person, but when that information is combined with other available information, it could entail such a risk*". The Directive lays down an actual obligation upon Member States to make all existing documents held by 'public sector bodies' and public undertakings re-usable, unless access is restricted or excluded under national rules on access to documents or subject to the other exceptions. Public authorities can limit the making available of public data by imposing conditions in the standard licenses as regards the re-use by the licensee dealing with issues such as liability, the protection of personal data, the proper use of documents, guaranteeing non-alteration and the acknowledgement of source. Art. 8 of the Open Data Directive specifies that such conditions should be "*objective, proportionate, non-discriminatory and justified on grounds of a public interest objective*".

While the new Open Data Directive envisions the use of open licenses for data sharing and re-use to become common practice across the EU, such licensing arrangement is currently meant to apply only to public sector bodies. Therefore ,DUET partners may not be able to make full use of these regulated licenses (see also Chapter 4).

Furthermore, of relevance to DUET, the new rules will stimulate the publishing of dynamic data and the uptake of Application Programme Interfaces (APIs). The new rules also limit the exceptions which currently allow public bodies to charge more than the marginal costs of dissemination for the re-use of their data. The new Directive provisions enlarge the scope of the previous PSI Directive to:

- ➢ data held by public undertakings, which the undertakings make available for re-use. Charges for the re-use of such data can be above marginal costs for dissemination;

[14] Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information, [2003], OJ L 345. In the words of the European Commission, "Public sector information, sometimes also referred to as **government data**, refers to all the information that public bodies produce, collect or pay for. Examples are geographical information, statistics, weather data, data from publicly funded research projects and digitised books from libraries". (Digital Agenda for Europe, 2013). This data is referred to as PSI. The policies rest on the premise that the re-use of this type of data generates value on the economy and society. Open public data are PSI that can be readily and widely accessible and re-used. See:
https://ec.europa.eu/digital-single-market/en/european-legislation-reuse-public-sector-information

[15] Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information Text with EEA relevance, [2013], OJ L 17.

- ➢ research data resulting from public funding – Member States will be asked to develop policies for open access to publicly funded research data. New rules will also facilitate the re-usability of research data that is already contained in open repositories.
- ➢ strengthen the transparency requirements for public–private agreements involving public sector information, avoiding exclusive arrangements.

A relevant EU piece of legislation that applies as *lex specialis* is **Law Enforcement Directive,**[16] which aims at protecting the fundamental right to data protection whenever personal data is used by criminal law enforcement authorities. It is intended to ensure, in particular, that the personal data of victims, witnesses, and suspects of crime are duly protected and to facilitate cross-border cooperation in the fight against crime and terrorism. The Directive was initially proposed in 2012 as part of the data protection reform package launched by the EU Commission. The final text was adopted in April 2016 and published in the Official Journal of the EU on 4th May 2016, together with the General Data Protection Regulation (GDPR).

Finally, the **ePrivacy Directive**[17] also applies to to complement the EU data protection regime. It provides *inter alia* a privacy protection framework for transmission and processing of data in connection with use of public communication networks (such as Internet, mobile, or telephone networks) and in the provision of electronic communication services. The directive aims to provide an enhanced protection to personal data in, and ensure confidentiality of, electronic communications, and is in a relationship of speciality to the GDPR (provisions of the ePrivacy directive serve to particularise and complement the GDPR). The GDPR, in turn, provides with respect to the ePrivacy Directive that it (GDPR) does not impose additional obligations in relation to processing of personal data in connection with provision of publicly available electronic communication services in relation to matters for which they are subject to specific obligations with the same objective set out in the ePrivacy Directive.

2.2.2 Current national legal framework: Belgium, Czech Republic, Greece and France

Belgium

Similarly to all other Member States, the GDPR has direct effect and is directly enforceable before Belgian courts. The **Law of 30 July 2018 on the protection of natural persons with regard to the processing of personal data**[18] was published in the Belgian Gazette and entered into force on 5 September 2018.

Insofar as the private sector is concerned, the law only complements the GDPR and deviates from it to a limited extent. Insofar as the territorial scope is concerned, the Law will apply to companies and

[16] Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA, [2016], OJ L 119.

[17] Directive 2002/58 of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector, [2002] Official Journal L 201.

[18] SERVICE PUBLIC FÉDÉRAL JUSTICE, SERVICE PUBLIC FEDERAL INTERIEUR ET MINISTÈRE DE LA DÉFENSE [C−2018/40581], 30 JUILLET 2018 Loi relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel.

organisations that process personal data: 1) in relation to activities of an establishment which is situated in the Belgian territory, irrespective of the place where the processing takes place, 2) in relation to data subjects residing in the Belgian territory, when the company is not established in Belgium, 3) where it offers services to these subjects or monitors the behaviour of data subjects for as far as this behaviour takes place in the Belgian territory. The Law does not apply to a processor established in Belgium if the controller is established in another Member State, when the processing takes place in this Member State. Another important aspect is the provision where the processing of special categories of personal data is allowed (such as, for example, political beliefs), which the Member State can do when it deems such processing necessary for reasons of substantial public interest. Insofar as sensitive data such as genetic, biometric data or data concerning health are concerned, the data controller or processor must: (a) indicate which categories of persons have access to the data and explain their relation to the processing of this personal data; (b) maintain a list of these categories for the Belgian DPA; (c) to ensure that the designated persons are subject to a legal, statutory or equal contractual obligation to ensure the confidential nature of the data[19].

Czech Republic

In the Czech Republic, the GDPR has direct effect and is directly enforceable given the nature of a regulation under EU law. **Law no. 110/2019** Coll., on processing of personal data, complements GDPR as regards its internal application, and particularizes certain legal grounds for data processing. The same piece of legislation further transposes into Czech national law the Law Enforcement Directive (see above **Section 2.2.1**). This latter part concerns the processing of data by competent authorities, such as law enforcement authorities, insofar as this data falls under the definition of personal data according to the GDPR. Finally, Law no. 89/2019 Coll., the Czech Civil Code, contains provisions protecting several aspects of natural persons' privacy, including protection of one's personality, image, reputation, etc. The Code provides for standalone means of enforcement, such as cease and desist court orders or actions for damages. The protection conferred by the Civil Code may be concurrent with that provided by GDPR and the law on processing of personal data.

Greece

On 29 August 2019, **Law 4624/2019**[20] came into force. This piece of legislation enacts into national level the provisions of the GDPR as well as the Law Enforcement Directive. Since Greece failed to transpose the Law Enforcement Directive within the two years period required, the law has been said to transpose some of the Law Enforcement Directive's and the GDPR' principles in a hasty fashion, and a complaint was filed with the European Commission on the compatibility of this law's provision with the EU acquis, as well as with the Greek data protection authority.

On certain points, Greece has a higher threshold of protection for sensitive data than what is prescribed under the GDPR. For instance, according to Art. 23 of the Greek statute the processing of genetic data for health and life insurances is prohibited to avoid discrimination. In addition, Greece has also introduced exceptions from the purpose limitation principles under Art. 23 GDPR. In addition, some rights of the data subject such as those of erasure, access, rectification may be limited under the limitations set out under Art. 23 of the GDPR, but are also subject to a narrow discipline. The Greek legislator "has also made extensive use of the limitations permitted by Article 23 of the GDPR to restrict the data subjects' right to information, the right to access and the right to rectification and erasure, without fully complying with the safeguards provided in Article 23, para 2 GDPR"[21]. Indeed, authors have questioned the compatibility of those provisions

[19] KPMG, Belgian Data Protection Legislation, 2018.
[20] Greek Law regarding the protection of personal data has been published in the Government Gazette (137/A/29-08-2019).
[21] EDRI, Greece: The new data protection law raises concerns, 2019.

with the GDPR. For instance, Greek law introduces provisions that allow the data controller not to erase data upon request of the data subject, in case the controller has reason to believe that erasure would adversely affect legitimate interests of the data subject[22]. Thus, the data controller is in effect allowed by the Greek legislator to substitute the will of the data subject in certain cases.

France

In 2016, France passed the **Digital Republic Act**[23], a landmark piece of legislation where the GDPR provisions and EU open data legislation were enacted into national level. Such legislation made France the first Member State to mandate local and central government automatically to publish documents and public data.

The Act contains several provisions to help achieve the two following goals: to introduce concrete strategies and the digital transformation of the economy and "*forge a resolutely contemporary digital policy underpinned by citizens, users, entrepreneurs, civil servants, consumers, 'makers' and by a whole host of these people to empower them and bolster their rights in the digital universe*"[24]. Notably, it grants data subjects a right to self determination entailing a right to control and decide over the use of his/her personal data. It furthermore crystalizes the right to be forgotten and foresees a procedure for individuals to access, erase and rectify data by electronic means.

In order to achieve these goals, the Act focuses - *inter alia* - on bolstering and broadening the open data policy (only anonymised data can be open) and building a data-oriented public service. It also introduces the new concept of data of general interest, including data coming from both public and private entities, public service concession holders or entities whose activities are subsidised by the public authorities, in order to to make the best possible use of data in the public interest, foster an open environment, uphold the principles of network neutrality and interoperability and transferability of data and set out privacy-compliant principles for data access.

2.2.3 Potential future legislation: The Data Act and the ePrivacy Regulation

In its February 2020 **Data Strategy**[25] (adopted alongside the **White Paper on Artificial Intelligence**[26]), the European Commission announced the first pillars of the new digital strategy of the Commission, including its intention to:

- dopt legislative measures on data governance, access and reuse, for example for business-to-government data sharing for the public interest;

[22] *Id*.
[23] JORF n°0235 du 8 octobre 2016 texte n° 1 LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique.
[24]Please find the Explanatory Memorandum of the Bill at:
https://www.republique-numerique.fr/pages/digital-republic-bill-rationale .
[25]European Commission, Communication from the Commission to the European Parliament, the Council, the European and Social Committee and the Committee of the Regions, A European strategy for data, Com(2020) 66 Final. This follows the European Commission's Communication from the Commission to the European Parliament, the Council, the European economic and social Committee and the Committee of the Regions, Towards a common European data space" COM(2018) 232 final. Also see European Commission, "Guidance on sharing private sector data in the European data economy", accompanying the document, "Communication from the Commission to the European Parliament, the Council, the European economic and social Committee and the Committee of the Regions, Towards a common European data space", SWD(2018) 125 final.
[26] *Infra*.

- ➢ Make data more widely available by opening up high-value publicly held datasets across the EU and allowing their reuse for free;
- ➢ Invest €2 billion in a European High Impact Project to develop data processing infrastructures, data sharing tools, architectures and governance mechanisms for thriving data sharing and to federate energy-efficient and trustworthy cloud infrastructures and related services;
- ➢ Enable access to secure, fair and competitive cloud services by facilitating the set-up of a procurement marketplace for data processing services and creating clarity about the applicable regulatory framework on cloud framework of rules on cloud;
- ➢ Empower users to stay in control of their data and investing in capacity building for small and medium-sized enterprises and digital skills;
- ➢ Foster the roll out of common European data spaces in crucial sectors such as industrial manufacturing, green deal, mobility or health.

It is expected, but it is to be monitored, that the above-mentioned act could clarify how data could be re-used in the context of smart cities.

The **proposed ePrivacy Regulation** is expected to replace the ePrivacy Directive with added relevance for Smart Cities. The proposal is considered politically sensitive, and, after the European Commission submitted a draft in early 2017, it has only very slowly progressed through the legislative process. The European Parliament and the Council have revised the draft in several iterations (the latest publicly available being the Council Presidency compromise proposal from March 2020), but there is no clear timeline for the measure's adoption in place.[27] While for the purposes of the report we will refer to, where applicable, the ePrivacy Directive, this legislative initiative must be closely monitored and the following paragraphs provide a brief overview of the ePrivacy Regulation novelties potentially relevant for DUET.

The ePrivacy Regulation aims - *inter alia* - to close certain application gaps of the ePrivacy Directive as regards the concept of electronic communication services and its inclusion of machine-to-machine communication services (M2M) and (provisionally called) "IoT services", and to clarify the rules on use and storage of information on users' terminal equipment (such as mobile phones or connected vehicles), as well as regulate further types of data processing activities, such as use of helpful metadata for wider benefit. These elements are essential features of the IoT and large-scale data analysis and thus likely of high relevance for Smart Cities initiatives. The proposal is part of the EU's Digital Single Market Strategy, the overarching objective of which is to increase trust in and the security of digital services in the internal market. Three main aspects of the ePrivacy Regulation proposal appear highly relevant to Smart Cities:

- ➢ **Confidentiality of communications and processing of data in M2M/IoT services - scope of ePrivacy rules' application**.

  One of the pillars of both extant and proposed ePrivacy rules is the achievement of confidentiality and security of data in the specific context of their transmission through electronic communication networks. As such, they build on the framework protecting personal data (but are not limited to personal data) by imposing special obligations on providers of electronic communication services, typically mobile network or internet connection providers, and also to persons who make use of

[27] The process has also been impacted by the COVID-19 pandemic; Council Presidency Progress report on the ePrivacy regulation proposal dated 29 May 2020.

information related to end user's terminal equipment (such as website operators' placement of cookies on users' mobiles and computers).

One of the features typical of IoT is that it relies on machine-to-machine communication (M2M, also called in the latest ePrivacy Regulation proposal as "IoT services"), which is an automated transfer of data and information between devices or software-based applications with limited or no human interaction[28] (for example, transfer of data between sensors or from a sensor to an automated processing unit). The ePrivacy Directive and related legislation is not entirely clear whether such devices and their data transmission carried out within closed networks but then transmitted over the Internet or other public network fall within the scope of application of that Directive, which creates a potential application gap.[29] The ePrivacy Regulation proposal makes clear that M2M/IoT services are intended to be covered. Even though the rules should not apply to networks of closed groups of end users, such as home or corporate networks, the proposal ePrivacy Regulation clarifies that as soon as electronic communication data is transferred from such closed group network to a public communications network (such as the Internet), the ePrivacy rules will apply to such data including when it is M2M/IoT data.

To the extent any DUET's activities will involve any use of M2M/IoT services that fall under the above definitions, DUET may be considered an electronic communication service provider in the sense of ePrivacy rules. This means, in essence, that electronic communication data in/from such activities may be processed only where it is necessary to provide an electronic communication service (that is, transmit the information from source to an end-user or end-equipment), and once the data are no longer needed for transmission, they should be deleted or at least anonymized. However, such data (and importantly, metadata) may be further processed or stored, but subject to specific conditions in more detail described elsewhere in this chapter where relevant. Note that the requirements on confidentiality and electronic communication data processing apply (already under the existing ePrivacy Directive) not only with respect to personal data, but also other traffic data and metadata, and data related to legal persons which are not personal data (and which thus fall out of the scope of GDPR altogether). Anonymization/de-identification does not remove that data from the scope of ePrivacy rules.

- **Use of end users' terminal equipment**. Under the existing ePrivacy Directive, the use of the processing and storage capabilities of users' terminal equipment (such as mobile phones, smart watches, laptops, or connected vehicles or other devices in the IoT) or access to information stored in such equipment is possible principally only with the user's consent or where the user requests a provision of a service. The latest ePrivacy Regulation proposal (March 2020) introduces a new, flexible legal ground for processing that is based on legitimate interest pursued by the service provider, except when such interest is overridden by the interests or fundamental rights and freedoms of the end-user. End-user's interests are deemed to be overriding the provider's in certain cases (provider cannot process data for user profiling purposes or information that contains special categories of personal data). In any event, such legal ground may be used only after a careful assessment and subject to further safeguards, and the data thus processed cannot be shared with third parties unless anonymised. Other types of information processing related to terminal equipment will remain subject to end-user's consent; these may include IoT-related services that need information emitted by the terminal equipment to enable it to connect to another device. The

[28] As per recital 12 of March 2020 ePrivacy Regulation proposal.
[29] Edwards, L., cited, page 19.

ePrivacy Regulation proposal does, however, allow a non-consent based processing of equipment-emitted information such as in provision of physical movements' tracking services (e.g., services enabling statistical people counting in a specific area), subject to further safeguards to minimise impact on individuals' privacy.

- **Processing of metadata**. The ePrivacy Regulation proposal recognises that metadata such as location data can be useful for businesses, consumers and the society as a whole and aims at broadenign the possibilities for providers to process such data vis-à-vis the fairly restrictive regime of the ePrivacy Directive (see **subsection 2.4.5** discussing processing of location data under the currently applicable law). In addition to processing based on user's consent, or where it is necessary for the provision of a service based on contract with the end-user or for billing, metadata can be also processed based on legitimate interest pursued by the service provider, except when such interest is overridden by the interests or fundamental rights and freedoms of the end-user. The ePrivacy Regulation proposal recitals give several examples of possible legitimate interests, such as processing for scientific research or statistical counting purposes for example for creation of heat maps (a geographical representation of data using colours to indicate the presence of individuals); the proposal acknowledges that such usage of electronic metadata may benefit public authorities and transport operators to define where to develop new infrastructure. Metadata thus processed may not be used to build end-user profiles (the legislator wishes to prevent illegitimate practices such as use of data for segregation of people), the data should not contain special categories of personal data and in any event cannot be shared with third parties unless made anonymous. Further safeguards not mentioned in this summary will apply to such use of metadata.

The wording and scope of the ePrivacy Regulation proposal provisions are still subject to debate and amendments in the legislative process, and a close eye must be kept on these developments. Note, further, that the definition of "electronic communication service" will have changed to include M2M transmission services with the effectiveness of the newly adopted European Electronic Communications Code (Directive 2018/1972)[30] as of 21 December 2020, so such communications/services will fall into the scope of the current ePrivacy Directive irrespective of when or whether the ePrivacy Regulation proposal gets adopted at the end of the day.

The ePrivacy Regulation will be, similarly to the existing ePrivacy Directive, a *lex specialis* to the GDPR as regards personal data, which means that all matters concerning the processing of personal data not specifically addressed by ePrivacy rules are governed by the GDPR. Currently, there are some overlaps between GDPR and ePrivacy rules that aim to achieve similar objectives, and the proposed ePrivacy Regulation aims to reduce such overlaps and help decrease the administrative burden placed on data controllers by repealing certain ePrivacy Directive provisions (for example, Article 4 mandating data processing security obligations, which is similar to Article 32 of the GDPR[31]).

Finally, the ePrivacy Regulation proposal contains a newly designed set of penalties that follow the stricter trend brought about by the GDPR: infringements of ePrivacy rules may become subject to administrative fines up to EUR 10M/20M or up to 2%/4% of the total worldwide annual turnover of the persons liable.

[30] Directive (EU) 2018/1972 of the European Parliament and of the Council of 11 December 2018 establishing the European Electronic Communications Code [2018], *OJ L 321.*
[31] *Infra*, Section 3.

2.2.4 Soft law and other relevant policy initiatives

EDPS and EDPB Guidance

The EDPS, the EU body overseeing data protection compliance at the EU institutions, provides the European institutions and bodies with policy advice on all matters relating to the processing of personal data. Among other things, its mandate includes the provision of practical recommendations and practical solutions through adoption of guidelines, serving as a useful source of inspiration for other organisations outside the EU institutions as well as an additional guidance to that offered by national data protection authorities.
In addition, the EDPB has adopted guidelines clarifying data processing through the use of some technologies in the context of smart cities. As of 2018, the EDPB has succeeded the Art. 29 Working Party and, in light of the GDPR, is in charge of ensuring the consistency of its application.

Relevant to the present report are the **EDPS Guidelines for assessing the proportionality of measures that limit the fundamental rights to privacy and to the protection of personal data**[32], which complement the **EDPS Necessity Toolkit**[33], the **EDPB Guidelines on processing of personal data through video devices**[34], as well as **ePrivacy Directive Guidelines 1/2020 on the processing of personal data in the context of connected vehicles and mobility related applications**[35], **EDPB Opinion 5/2019 on the interplay between the GDPR and the ePrivacy Directive**[36]. Of specific relevance are also:

EDPS Guidelines on assessing the proportionality of measures that limit the fundamental rights to privacy and to the protection of personal data[37] of 19 December 2019 are intended to help with the assessment of compliance of proposed measures with EU law on data protection, having regard to the fundamental right to the protection of personal data enshrined under Art. 8 CFR. They have been developed to better equip EU policymakers and legislators responsible for preparing or scrutinising measures that involve the processing of personal data and limit the rights to protection of personal data and to privacy.
Once they have identified the measures which have an impact on data protection and the priorities and objectives behind these measures, policy makers and legislators are assisted in finding solutions which minimise conflict between these priorities, while being proportionate.

[32] EDPS, "Guidelines on assessing the proportionality of measures that limit the fundamental rights to privacy and to the protection of personal data", 25 February 2019, available at: https://edps.europa.eu/data-protection/our-work/publications/guidelines/edps-guidelines-assessing-proportionality-measures_en .
[33] EDPS, "Assessing the necessity of measures that limit the fundamental right to the protection of personal data: A Toolkit", 11 April 2017, available at: https://edps.europa.eu/sites/edp/files/publication/17-06-01_necessity_toolkit_final_en_0.pdf.
[34] EDPB, "Guidelines 3/2019 on processing of personal data through video devices", 10 July 2019, available at: https://edpb.europa.eu/sites/edpb/files/consultation/edpb_guidelines_201903_videosurveillance.pdf .
[35] EDPB, "Guidelines 1/2020 on the processing of personal data in the context of connected vehicles and mobility related applications", 7 February 2020, available at: https://edpb.europa.eu/our-work-tools/public-consultations-art-704/2020/guidelines-12020-processing-personal-data-context_en.
[36] EDPB, Opinion 5/2019 on the interplay between the GDPR and the ePrivacy Directive, 12 March 2019, available at: https://edpb.europa.eu/our-work-tools/our-documents/stanovisko-vyboru-cl-64/opinion-52019-interplay-between-ePrivacy_en .
[37] EDPS, "Guidelines on assessing the proportionality of measures that limit the fundamental rights to privacy and to the protection of personal data", 19 December 2019, available at: https://edps.europa.eu/sites/edp/files/publication/19-12-19_edps_proportionality_guidelines2_en.pdf .

EDPB Guidelines on processing of personal data through video devices (3/2019)[38] focus on how and when the use of video devices interplays with the application of the GDPR, i.e. video-surveillance implicates personal data processing and may pose risks of unauthorized uses, lawfulness of the processing, purpose limitation, processing of special categories of data, transparency of the processing, storage periods, security measures.

EDPB Guidelines on connected vehicles (**Guidelines 1/2020 on the processing of personal data in the context of connected vehicles and mobility related applications**)[39] are important insofar they further clarify certain concepts contained in the GDPR, such as data processor or controller, as well as highlight some privacy and data protection concerns of IoT, which can be relevant in the context of smart cities. In addition, they also explain the interplay between the GDPR and the ePrivacy Directive when it comes to some of this Directive's provisions.

The **EDPB Opinion 5/2019 on the interplay between the GDPR and the ePrivacy Directive**[40] clarifies the interplay between the two legal instruments. For example, it clarifies aspects concerning the obligations the controller must abide by when processing personal data, what notion of consent shall prevail, etc.

Finally, on 21 April 2020, the EDPB adopted **Guidelines 04/2020 on the use of location data and contact tracing tools**[41] and **Guidelines 03/2020 on the processing of data concerning health for the purpose of scientific research in the context of the COVID-19 outbreak**[42]. In order to support the response to the pandemic, the first set of guidelines explains the use of location data to support the response to the pandemic so as to assess the overall effectiveness of confinement measures and contact tracing, with the aim of notifying individuals when they have been in close proximity with a confirmed carrier of the virus. The second set of guidelines addresses the latest legal questions concerning the use of health data pursuant to Art. 4(15) GDPR for research purposes connected to the fight against the COVID-19, namely the legal basis, the implementation of adequate safeguards for the processing of health data and the exercise of the data subjects rights. These initiatives should be taken into account since they clarify how the processing of certain sensitive data under the GDPR can occur in ways that are compatible with the purpose limitations required by Art. 23 of the GDPR.

Other relevant policy initiatives

A further policy initiative worth mentioning is the 2020 **European Commission's White Paper on AI**[43], touching in particular, on aspects that should be taken into account when addressing some of the privacy-related risks AI technology may generate iis-à-vis certain fundamental rights, such as personal data

[38] Cited.
[39] *Id.* 28.
[40] *Id*. 29.
[41] EDPB, "Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the COVID-19 outbreak", 21 April 2020, available at: https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_20200420_contact_tracing_covid_with_annex_en.pdf .
[42] EDPB, "Guidelines 03/2020 on the processing of data concerning health for the purpose of scientific research in the context of the COVID-19 outbreak", 21 April 2020, available at: https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-032020-processing-data-concerning-health-purpose_en .
[43] European Commission, On Artificial Intelligence - A European approach to excellence and trust, COM(2020) 65 Final, 19 February 2020.

and privacy protection, but also the risks that AI may foster bias and thus threaten the principle of non-discrimination, or risks related to its opacity problem. The White Paper posits that "*AI is a strategic technology that offers many benefits for citizens, companies and society as a whole, provided it is human-centric, ethical, sustainable and respects fundamental rights and values*". With the White Paper and the accompanying Report on the safety and liability framework, the Commission launches a broad consultation of Member States civil society, industry and academics, of concrete proposals for a European approach to AI, including both policy means to boost investments in research and innovation, enhance the development of skills and support the uptake of AI by SMEs, and proposals for key elements of a future regulatory framework.

Furthermore, the above-mentioned **Commission's Communication on a European strategy for data**[44] advocates, among others, the use of public sector information by business (government-to-business – G2B – data sharing) as well as the use of privately-held data by government authorities (business-to-government – B2G –data sharing). This latter is particularly relevant in the context of smart cities. The Commission's Communication is accompanied by a report on business-to-Government (B2G) data sharing. In the report, experts advise to make data sharing in the EU easier by "*taking policy, legal and investment measures in three main areas:*

(a) ***Governance of B2G data sharing across the EU****: such as putting in place national governance structures, setting up a recognised function ('data stewards') in public and private organisations, and exploring the creation of a cross-EU regulatory framework.*
(b) ***Transparency, citizen engagement and ethics****: such as making B2G data sharing more citizen-centric, developing ethical guidelines, and investing in training and education.*
(c) ***Operational models, structures and technical tools****: such as creating incentives for companies to share data, carrying out studies on the benefits of B2G data sharing, and providing support to develop the technical infrastructure through the Horizon Europe and Digital Europe programmes*"[45].

What is of relevance for our purposes is also that the report revises existing principles of private sector data sharing in B2G contexts[46] and includes as new principles both accountability and fair and ethical data use, which should guide B2G data sharing for the public interest.

ITU, Artificial Intelligence for Development Series, 2018

In 2018, the telecommunication development bureau of the International Telecommunications Union (ITU) started promoting an initiative to deepen the understanding of, and promote further discussion and collaboration among policy makers and regulators of the significance of Artificial Intelligence (AI). The AI series is part of this initiative. Among others, this initiative deals with AI and data, as well as privacy. Speaking about access to data, the initiative highlights that open data and open standards for public data are likely to be an important enabler of AI. It recommends that governments promote open standards to build a robust data ecosystem, making systems and data interoperable. These also include common standards for metadata, which will allow the provenance of data to be traced as data is used and reused for different

[44] COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS A European strategy for data, COM(2020) 66 final.
[45] High Level Expert Group on Business to Government Data Sharing, Report, Towards a European Strategy on Business-to-Government data sharing in the public interest, 19 February 2020.
[46] See principles in the Commission's Communication 'Towards a common European data space' and its Staff Working Document.

purposes. Among others, this report also mentions the OECD privacy framework, containing recommendations where transborder flows of personal data are at stake. It highlights the tension between, on the one hand, data for public good, and, on the other hand, the need to respect privacy of personal data, in relation to which de-identification techniques are highlighted. It also mentions some hurdles with commercial or proprietary data, which have strategic value, but also present costs and risks. Among the costs, the report mentions the costs of anonymization or de-identification, and risks such as re-identification but also third party confidentiality rights. A separate section deals with personal data, and it highlights that the tension between data protection and the realization of IoT will create new grey areas with space to circumvent legislative hurdles. It also highlights whether traditional data protection laws and paradigms are fit for an IoT era, and calls for a more-context specific approach. It also advocates not always relying on consent given by an individual, when the risks to privacy are minimal, and suggests that other grounds for processing could *inter alia* be the notion of 'legitimate interest'.

2.3 Privacy risks in the context of smart cities

Trust building

A smart city is successful when there is trust in the relationship its stakeholders have with the city, its services and its service providers.

Prior to delving into the description of the specific risks that data processing by the DUET partners may encounter throughout the roll-out of the pilots, and the overall project, a few preliminary remarks regarding the very first general risk that the DUET partners may well come across must be made: this risk concerns a general public suspicion/lack of trust in respect to how the data collected is handled. **Lack of trust in data privacy and system integrity could represent a major barrier to a smooth unfolding and functioning of the DUET projects.**

Kitchin discusses several types of privacy concerns with the increased datafication that smart city technologies may unleash[47]. The data collected is often indiscriminate and exhaustive, distributed across multiple devices, services and places, platform independent, i.e. flowing easily across platforms and devices as well as continuous, i.e. generated on a routine and automated basis. As such, the production of detailed datasets that can easily be examined through data analytics, stored in digital databases, conjoined with other datasets and shared impacts privacy in ways that could negatively impact trust of citizens: according to Kitchin, the first of the privacy related concerns is that people potentially become subject to increased **surveillance and dataveillance** as ever before. For instance, location and movement tracking can occur through gathering data through CCTV cameras installed in cities, smartphones and smartphone apps, sensor networks deployed across street infrastructure, wifi mesh, GPS data devices in cars and vehicles, increased digital footprint and voluntary sharing of data by individuals. Organising, storing and sharing of big data also changes the use to which such data can be put, with some uses being unpredictable and unexpected, i.e. which concern **data being repurposed in ways that have little to do with the original purpose for which they were generated and without giving prior notice to the individuals concerned by the data**. This risks

[47] Kitchin, R., "Getting smarter about smart cities: improving data privacy and data security", 2016, Data Protection Unit, Department of the Taoiseach, Dublin, Ireland, (2016).

bring about potential erosion of the purpose limitation or the data minimisation principles that underpin the GDPR.

At the one end of the spectrum, in a smart city, the data with the potential of re-use in terms of improved services is often also data coming from personal data being processed, which is difficult to anonymise without diminishing the potential to use them: this is for example, the case for data on mobility[48]. On the other hand, some examples of existing smart cities initiatives, such as now halted Google's Sidewalks Labs[49] project in Waterfront Toronto, show that questions about data use arise, as a complaint was filed by an NGO before a competent court and a scrutiny was launched in 2019 by a panel, composed - *inter alia* - of privacy experts, advising decision makers. According to privacy experts, the project would have entailed no-surveillance free zones and the privatisation of so-called 'urban-data'. Risk mitigations put in place such as de-identification by default and and independent so-called 'civic data trust' which would mediate access to and use of this urban data were deemed insufficient to address the concern that this project would result in 'control creep'[50]. This example shows that good, adequate data governance may be crucial to ensure citizen support in the potential of smart cities to improve citizen's lives, and that handling data requires a need to be privacy-cautious throughout the life cycle of a project. Annex I provides an overview of the legal issues that had arisen in the context of the Sidewalk Toronto project.

Against the above-mentioned concerns, building trust in the DUET Digital Twins may require communication and educational initiatives, i.e. awareness raising campaigns, to explain what the DUET project is about. It also requires devising a sound data governance policy, compatible with the data protection laws of the European Union, as they currently stand, something this Section will focus on.

DUET may also need to focus on promoting transparency and fostering citizen participation. Indeed, **DUET is created for the aim of realising 'the full potential of city data to drive an era of informed, smart and co-created policy making' (see D8.1). Hence, it is fundamental to allow the citizens to partake in Data Policy Making in line with DUET's Policy-Ready-Data-as-a-Service (PRDaaS).** Literature provides examples on how to actively involve citizens into smart city design and policies, including through citizen participation[51], crowd-sourcing[52], citizen-centered approaches[53], or co-creation and living labs[54] .

Therefore, some of the questions that arise are: How to persuade users of the benefits of smart city projects? How to do so by ensuring that citizens willingly give up on part of their privacy for the convenience of living

[48]De Montjoie, A., Hidalgo, C.A., Verleysen, M., Blondel, V.D., "Unique in the Crowd: the privacy bounds of human mobility", Nature Scientific Reports (2013), and CNIL, Cahier IP5-La Plateforme d'Une Ville, 2017.
[49] Tusikov, N., "Sidewalks Toronto Master Plan raises urgent questions about data and privacy", 2019. About the lessons drawn, Goodman, E., Powles, J., "Urbanism under Google: lessons from Sidewalks Toronto", 2019.
[50] Kitchin, R., "The data revolution: big data, open data, data infrastructures and their consequences, 2014.
[51] Berntzen, L., Johansson, M.R., "The role of citizen participation in municipal Smart City projects: Lessons learned from Norway", Smarter as the new urban agenda, Springer International Publishing (2016).
[52] Schuurman, D., Baccarne, B., De Marez, L., Mechant, P., "Smart ideas for smart cities: Investigating crowdsourcing for generating and selecting ideas for ICT innovation in a city context", Journal of theoretical and applied electronic commerce research, 7 (3) (2012).
[53] Gaved, M., Jones, A., Kukulska-Hulme, A., Scanlon, E., "A citizen-centred approach to education in the smart city: Incidental language learning for supporting the inclusion of recent migrants", International Journal of Digital Literacy and Digital Competence, 3 (4) (2012).
[54] Schaffers, H., Sällström, A., Pallot, M., Hernández-Muñoz, J.M., Santoro, R., Trousse, B.,
"Integrating living labs with future internet experimental platforms for co-creating services within smart cities", Concurrent enterprising (ICE), 2011 17th international conference on, IEEE (2011).

in a city that makes their lives more beneficial and easier[55]? How to exploit the full potential of participatory democracy (civil tech) when running a smart city project? Furthermore, how can open data help strengthen trust in smart cities by the public, but also how to "*prevent a smart city project from becoming just a juxtaposition of separate initiatives, without synergies*"[56]? Finally, what processes need to be put in place to ensure the development of a citizen-centric resilient smart city, where sound data governance and security policies are created throughout the life cycle of the project and not just *a posteriori.*

The following subsections dive deeper into some specific privacy-related risks that could arise in the context of the DUET project. For convenience purposes, we will follow the macro-areas that correspond to the abovementioned principles enshrined in the GDPR. Prior to this, we will also look into what typologies of data may give rise to privacy-concerns.

2.3.1 Categories of data: personal and non-personal data, mixed data-sets

We anticipate that in the context of the DUET Digital Twins development, including the pilots, a substantial amount of data will be processed. At times, such processing may involve personal data, which needs to be handled in compliance with the strict regime of the GDPR, directly applicable across all EU Member States. 'Processing', is a broad concept defined by Art. 4 of the GDPR, and in accordance with the EDPB clarifications, it includes any operation or set of operations which is performed on personal data, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction, etc[57]. This data processing should be made in compliance with the EU data protection legal framework as well as applicable national legal frameworks, such as that of the Member States where the pilots take place, when this national framework complements the relevant EU legislation.

For the purposes of our analysis, the data which will be processed during DUET Digital Twin's development can be grouped into three main categories:

- personal data;
- non-personal data;
- mixed data-sets combining the two.

While in theory the distinction is easy to grasp, in practice it is not. Consider, for example, the notion of 'location data': while this data is mentioned under Art. 4(1) of the GDPR, clarifying what personal data is ("*any data that may identify an individual*"), when this data is made anonymous, then it is no longer personal and the GDPR ceases to apply, even though ePrivacy rules may still be applicable. Yet, for legal purposes, the distinction is necessary in order to identify the applicable legislation to each set of data, the risks related to the processing for each type of data and, therefore, which risk mitigation measures to be taken by the organisations handling this data at each stage of the processing.

[55] Bertels, N., "Smart City innovation: should you be willing to trade your privacy for utility?", 4 July 2017, KU Leuven CiTiP, available at: https://law.kuleuven.be/citip/blog/smart-city-innovation-should-you-be-willing-to-trade-your-privacy-for-utility/ .
[56] Baudoin, P., Trust is key to the success of smart cities, (2016).
[57] EDPB, Guidelines 1/2020.

This Section will briefly define the main actors, the three categories of data, pinpoint the legal rules governing their processing, as well as identify the risks related to such data.

Definitions and applicable law

➢ *Who's who: Data subject, data processor, data controller and third parties*

Under Art. 4(1) of the GDPR, a 'data subject' is any natural person, i.e. a living individual.

According to Art. 4 of the GDPR, the main actors involved in data processing are the following:

- Controller "*means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data*";
- Processor "*means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller*";
- Third party "*means a natural or legal person, public authority, agency or body other than the data subject, controller, processor and persons who, under the direct authority of the controller or processor, are authorised to process personal data*".

➢ *Personal data*

Art. 4(1) of the GDPR defines personal data as "*any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person*".

While most of these categories are straightforward, online identifiers present a more difficult concept. The GDPR provides several examples of these in Recital 30 that include: Internet protocol (IP) addresses, cookie identifiers and other identifiers such as radio frequency identification (RFID) tags. These identifiers refer to information that is related to an individual's tools, applications, or devices, like their computer or smartphone. Any information that could identify a specific device, like its digital fingerprint, is an identifier. The broad definition of personal data has remained essentially unchanged in the GDPR as compared to the previous legislation. Various aspects of the definition of personal data, such as 'any information', 'relating to', 'identified or identifiable', were already clarified by Art. 29 Working Party 11 in its Opinion 4/2007 on the concept of personal data.[58] The GDPR only applies to personal data processed in one of two ways:

- Personal data processed wholly or partly by automated means (or, information in electronic form); and
- Personal data processed in a non-automated manner which forms part of, or is intended to form part of, a 'filing system' (or, written records in a manual filing system).

When the data processor is a competent authority under Art. 3 of the Law Enforcement Directive, then such piece of legislation regulates the processing of such data. The Law Enforcement Directive underpinning principles mirror those enshrined in the GDPR, the main piece of legislation to which we will refer.

[58] Of 20 June 2007, WP 136.

➢ *Non-personal data*

In accordance with the Commission's Communication **Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union**[59], where the data is not 'personal data' as defined in the GDPR, these are non-personal data and the GDPR does not apply.

Based on the description of the DUET Digital Twins made in Deliverable D8.1, it is likely that certain applications will make use of large quantities of non-personal data (for example, data on environmental footprint which cannot be traced to an individual or depreciation of road infrastructure). The **Free Flow of Non-Personal Data Regulation** provides a legal framework for the free flow of non-personal data in the EU. Examples are machine-generated data or commercial data, which are either non-personal in nature or refer to personal data that has been made anonymous.

The abovementioned Communication clarifies the notion of non-personal data, which can be categorised by origin either as:

- ❖ data which originally did not relate to an identified or identifiable natural person, such as data on weather conditions and air pollution generated by sensors installed on wind turbines or data on maintenance needs for industrial machines.
- ❖ data which were initially personal data, but were **later made anonymous**. The 'anonymisation' of personal data is different from pseudonymisation, **as properly anonymised data cannot be attributed to an identifiable person, not even by use of additional data** and are therefore non-personal data.

At the same time, it is worth noting that privacy concerns with respect to the above-mentioned data may still arise, although the scope of the Free Flow of Non-Personal Data Regulation is different from that of the GDPR. This is the case for example, when non-personal data goes through machine learning or deep learning processing, insofar as there is still a slight chance it can be linked to an individual thanks to the computational capabilities of itself.

To conclude, personal data are those which can be traced back to an individual, causing them to be either directly or indirectly identifiable. For the latter, the reasonableness test is used. Under Recital 26 of the GDPR, to determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly.

As the definition of personal data refers to 'natural persons', datasets containing the names and contact details of legal persons are in principle non-personal data. However, they may be regarded as personal data, if the name of the legal person is the same as that of a natural person who owns it or if the information relates to an identified or identifiable natural person. While the distinction appears straightforward in theory in practice it is not.

➢ *Mixed data-sets*

[59] Communication from the Commission to the European Parliament and the Council - Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union - COM(2019)250.

A mixed dataset consists of both personal and non-personal data. Mixed datasets represent the majority of datasets used in the data economy and are commonplace in a smart city environment, whereIoT systems, AI and technologies enabling big data analytics are deployed.

Art. 2(2) of the Free Flow of Non-personal Data Regulation provides that "*where personal and non-personal data in a data set are inextricably linked, the Regulation on the free flow of non-personal data shall not prejudice the application of GDPR*". In accordance with the Commission's recommendation in its Guidance [60], in a case of a dataset composed of both personal and non-personal data, several pieces of legislation will apply: (a) the Free Flow of Non-Personal Data Regulation applies to the non-personal data part of the dataset; (b)·the General Data Protection Regulation' free flow provision (Art. 1(3) GDPR) applies to the personal data part of the dataset; and, importantly, (c)·if the non-personal data part and the personal data parts are 'inextricably linked', the data protection rights and obligations stemming from the GDPR fully apply to the whole mixed dataset, even when personal data might represent only a small part of the dataset.

However, the Guidance acknowledges that the concept of 'inextricably linked' is not defined by either of the two Regulations. It goes on to specify that, for practical purposes, "*it can refer to a situation whereby a dataset contains personal data as well as non-personal data and separating the two would either be impossible or considered by the controller to be economically inefficient or not technically feasible*". In addition, the Guidance also specifies that "*separating the dataset is also likely to decrease the value of the dataset significantly*" and that "*the changing nature of data, makes it more difficult to clearly differentiate and thus separate between different categories of data*".

It is possible that mixed datasets will often be handled in the context of the DUET Digital Twins' project. Indeed, the Guidance expressly says that "*data related to the Internet of Things, where some of the data allow assumptions to be made about identifiable individuals (e.g. presence at a particular address and usage patterns); will contain mixed data sets*". Given the above clarifications, therefore, the GDPR will be the main legislation governing their processing.

Risks identified with regard to categories of data

Risk of re-identification: De-identification of data at source provides a "weak form of privacy protection" because it is always possible to reverse engineer the process by combining datasets

Scholars dealing with privacy and big data are skeptical about the notion of perfect anonymisation[61]. They opine that the law has been slow in adopting a holistic approach when it comes to the protection of data subjects when data sets are released to others. The law takes instead a snapshot approach focusing on whether an individual can be identified within a given data set. However, as seen above, when combining data sets, such re-identification is indeed possible. Namely, re-identification risks relate to the mechanisms that can be used which make it possible to identify data subjects within datasets. While some anonymisation techniques are more difficult to reverse than others, they may also still be vulnerable. According to recital 26 of the GDPR, the principles of data protection should apply to any information concerning an identified or identifiable person. Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information, should be considered to be information on an

[60] European Commission, Communication to the European Parliament and the Council, "Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union", COM/2019/250 final.
[61] De Montjoie, A., *et al*, (2013), cited. Rubinstein, I., Hartzog, W., "Anonymization and Risk" 91 Washington Law Review 703, NYU School of Law, Public Law Research Paper No. 15-36, (2016).

identifiable natural person. To determine whether a natural person is identifiable, according to this recital, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller, or by another person to identify the natural person either directly or indirectly.

Article 29 Working Party's 2014 Opinion on Anonymisation Techniques[62] further clarifies what those means of re-identification are and thus complements the GDPR in concretely understanding what the above-mentioned risks consist in. This Opinion describes three common risks as follows:

- 'Singling out' → the "*possibility to isolate some or all records which identify an individual in the dataset*."
- 'Linkability' → the "*ability to link at least two records concerning the same data subject or a group of data subjects (either in the same database or in two different databases)"*.
- 'Inference' → the "*possibility to deduce, with significant probability, the value of an attribute from the values of other attributes.*"

The distinction between personal and non-personal data is not always straightforward

Data should be thought of as a continuum, with the dividing line between what is personal data and what is not being subtle. In this vein, the abovementioned Open Data Directive has introduced the concept of dynamic data. According to Art. 2(1), 'dynamic data' means documents in a digital form, subject to frequent or real-time updates, in particular because of their volatility or rapid obsolescence; data generated by sensors are typically considered to be dynamic data. 'Citizen data' can be defined as personal and non-personal data, directly or indirectly generated in the digital public sphere, using digital technologies and collected through different infrastructures such as IoT, etc[63]. DUET public administrators and partners are required to identify if and how data handled by the project could fall under the scope of the EU data protection legal framework. In other words, in order to identify the obligations that the data controller(s) and the data processor(s) may be subject to, it will be necessary to distinguish between personal and non-personal data. Furthermore, within the category of personal data, it will also be necessary to distinguish between non-sensitive and sensitive data[64], this latter being subject to a higher level of protection under the GDPR.

Risks associated with location data

Several-fold risks associated with location data have been identified: first, location data can provide information on an identifiable or identified individual, and some of this information can reveal sensitive data of the data subjects. Suppose, for example, an individual who goes in a given LGBT bar: when location data about this frequentation are collected, processed and shared, they touch upon the sexual life of the individual, a specific category of sensitive personal data according to the GDPR. In this respect, a higher level of protection of this type of data as per the GDPR must be ensured. Second, location data can be re-used for purposes other than those for which they have been collected. In the risk mitigation part we will look further at risk mitigations in this respect. Third, when location data is anonymized, they can also be subject to re-identification under certain circumstances.

[62] Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques, Adopted on 10 April 2014, available at: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf .
[63] Eurocities Principles on Citizen Data, available at: http://nws.eurocities.eu/MediaShell/media/Citizen_data_principles_final_draft.pdf.
[64] See *infra*.

Data ownership and obligations for the various stakeholders

There may be a separate question of who owns the data collected in the context of smart cities. A spart city may involve multiple interacting data flows or multiple data owners/controllers[65]. In a data environment, the notion of ownership does not have a legal connotation. It refers to other concepts such as assurance of data quality and security. First, the actors in the value data chain who could claim ownership are several. Second there is no specific data-ownership legislation at EU level. Chapter 4 below explains the protection for certain types of data or dataset. The obligations for the various actors are clearer when a specific ownership right in data is spelled out. Because the current legal framework is not yet fully developed, there is a gap. That gap could be filled in with use of appropriate contractual arrangements. It will also be necessary to monitor the upcoming Data Act for novelty in this respect.

2.3.2 Legal grounds for processing

Recital 40 of the GDPR states that in order for processing to be lawful, personal data should be processed on the basis of the consent of the data subject concerned, or some other legitimate basis laid down by EU or national law. Having an appropriate legal basis for the processing of personal data is thus essential to ensure the DUET project's legality[66].

In terms of legal framework, the basic requirements for consent to be valid under the GDPR are laid down under Art. 7 and are explained further by recital 32 of the GDPR. Pursuant to Art. 7 of the GDPR, consent must be freely given, specific, informed and unambiguous, as explained by Recital 32 (i.e. given on the basis of a statement or a clear affirmative action). In order to be considered as freely given consent, consent must be given on a voluntary basis, i.e. the consent must imply a real choice by the data subject.

Consent becomes especially important when data cannot be de-identified at source since it is then that the GDPR applies. Against this background, this section will focus on the notion of a GDPR-compliant consent conundrum, its hurdles in the context of this project and what other alternative grounds for processing may be chosen.

<u>Consent under the GDPR and hurdles to obtain it in a smart city context</u>

A key issue in an ambient or smart city environment is whether or not obtaining meaningful consent to processing of personal data is at all possible[67]. The difficulty of obtaining freely given and informed consent in this context has long been recognized. The ubiquitous computing which a smart city depends on creates a hurdle both with initial consent giving, but also in the context of automated resharing of data. The European Union Agency for Cyber-security (ENISA) opines that "*the continuous repurposing and making use of already processed or inherent data sets, has made the traditional consent models insufficient and obsolete in big data. This has led to many arguments against the very concept of consent*". However, it acknowledges also

[65] Edwards, L., cited.
[66] See: https://smit.vub.ac.be/wp-content/uploads/2019/09/Report-roundtable-data-protection-in-smart-cities_def.pdf
[67] Id.

that "*consent is a fundamental data protection element and, like many other, it has to adapt to the new technological landscape with new usable and practical techniques*"[68].

Edwards, for example, points at the issue of how to obtain a meaningful prior consent in Internet of Things systems, especially where data is collected in public, as e.g. by smart road or smart transport systems. She highlights that smart cities further dilute the level of consent in the IoT: "*While consumers may at least have theoretically had a chance to read the privacy policy of their Nest thermostat before signing the contract, they will have no such opportunity in any real sense when their data is collected by the smart road or smart tram they go to work on, or as they pass the smart dustbin*"[69].

The technology the smart city relies on makes this hurdle seem insurmountable. Obtaining a freely given, specific, informed and unambiguous consent from each and every participant/citizen may be expensive or excessively burdensome in relation to the obligations incumbent on the data controllers. At times, obtaining and managing GDPR-compliant consent by providing timely information may prove operationally ineffective or impossible, such as when data must be used at the time of capture to meet the final objective, for example, react to an emergency[70].

According to Recital 42 of the GDPR, "*Consent should not be regarded as freely given if the data subject has no genuine or free choice or is unable to refuse or withdraw consent without detriment*." In the context of the Sidewalk Toronto project for some time run by Alphabet's subsidiary Sidewalk Labs, the lack of tools to provide express consent, including consent obtainable by box ticking, gave rise to privacy related concerns. Unlike an app, streets and parks cannot require their users to check a dialog box consenting to how their personal information will be used before granting access. In public spaces where personal information is collected—for example, a video footage that records people's faces in a crowd—there is no easy way for people to opt out of giving their consent. For more information on the Toronto Sidewalks project, see the Annex.

<u>Other alternative grounds for lawfully processing personal data</u>

If a consent cannot be obtained, then the GDPR foresees other lawful grounds for processing of personal data. There has to be, in other words, a need to identify a legal basis for any data processing without consent.

Data controllers may well opt for other grounds for processing personal data foreseen under the GDPR, namely the ground of 'public interest' (Art. 6 (1) (e) GDPR) and the 'legitimate interest' ground' (Art. 6(1)(f) GDPR), given the flexibility these notions present. For the sake of completeness, all the GDPR grounds for processing alternatives to consent will be outlined below. Special attention will be paid to the two aforementioned grounds.

- ➢ *Public interest*

[68] ENISA, Privacy by design in big data. An overview of privacy enhancing technologies in the era of big data analytics, 2015.
[69] Edwards, L., cited.
[70] Tarin, D., "Privacy and Big Data in Smart cities", The Smart City Journal, available at: https://www.thesmartcityjournal.com/en/technology/341-privacy-and-big-data-in-smart-cities .

Public interest is a basis for lawful processing alternative to consent. Art. 6(1)(e) GDPR gives an entity a ground for lawful processing of personal data when: "*processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller*".
Yet, as stated by Recital 40, as well as Art. 6(3) GDPR, the relevant task or authority must be laid down by EU law or national law. Therefore, this ground can apply only either when the entity is: (a) carrying out a specific task in the public interest which is laid down in law; or (b) is exercising official authority which is laid down in law[71]. In both cases, processing must be necessary, i.e. a targeted and proportionate way of achieving a set purpose. If another reasonable and less intrusive way to achieve the result exists, then the necessity test fails.

This is straightforward when it comes to the collection and use of personal data by a certain authority, e.g.. tax authority, whose legal powers are provided by national law along with data processing rules on income and tax returns and creation of a database. Yet, Recital 41 clarifies that this does not have to be an explicit statutory provision, as long as the application of the law is clear and foreseeable. Any organisation who is carrying out a specific task in the public interest may benefit from this ground for processing.

GDPR Recital 45 contemplates the possibility for private entities to process personal data for public interest purposes, provided it be determined by Union or Member State law.[72] Therefore, no particular issue as regards the participation of private parties in data processing operations should come into consideration, provided that all the other requirements are met.

Art. 6(2) GDPR enables Member States to maintain or introduce more specific provisions to adapt the application of the public task legal basis. For example, the 2017 UK Digital Economy Act[73] gives public authorities powers to share personal information across organisational boundaries to improve public services.

➢ *Contractual necessity*

The second legal ground for lawful processing as provided by the GDPR is the necessity of personal data processing in the context of a contract.

GDPR Recital 40 mentions '*the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract*' as a legitimate basis forlawful processing and GDPR Recital 44 simply states that processing should be lawful where it is necessary in the context of a contract or the intention to enter into a contract. Indeed one cannot enter in any contractual relationship without providing personal data and identifiers. At the very least this concerns contact information. In specific types of contracts, far more is required. The required data to enter into a contract or perform a contract really need to be provided in the scope of the contract and services offered.

➢ *Legal obligations*

[71] ICO Guidance on the GDPR.
[72] "*It should also be for Union or Member State law to determine whether the controller performing a task carried out in the public interest or in the exercise of official authority should be a public authority or another natural or legal person governed by public law, or, where it is in the public interest to do so, including for health purposes such as public health and social protection and the management of health care services, by private law, such as a professional association*".
[73] UK Digital Economy Act 2017, available at : http://www.legislation.gov.uk/ukpga/2017/30/contents/enacted .

The third legal basis for lawful processing is compliance with legal obligations.

The controller has a legal duty for which particular personal data needs to be processed, then such processing is permitted. However, particular rules apply. Recital 45 states that "*where processing is carried out in accordance with a legal obligation to which the controller is subject or where processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority, the processing should have a basis in Union or Member State law*". The use of this legal basis is thus limited to what is provided by EU law or EU Member State laws.

➢ *Vital interests*

The protection of the 'vital interests' of a natural person is a fourth ground for lawful processing.
In this case the natural person can be a natural person other than the data subject. It is not up to the controller to define what a vital interest is. There should be a life threatening circumstance where there is no other legal ground for processing but where not processing personal data would essentially mean that someone's life would be in actual danger (in case of a serious accident information about the victim's medical history such as allergies towards specific medication).

Recital 46 of the GDPR states that processing of personal data based on the vital interest of another natural person should in principle take place only where the processing cannot be manifestly based on another legal basis. Some types of processing may serve both important grounds of public interest and the vital interests of the data subject as for instance when processing is necessary for humanitarian purposes, including for monitoring epidemics and their spread or in situations of humanitarian emergencies, in particular in situations of natural and man-made disasters.

➢ *Legitimate interests*

Art. 6 (1)(f) GDPR states that processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data. The GDPR explicitly provides that the legal ground of legitimate interest does not apply to personal data processing by public authorities in the performance of their tasks.
GDPR Recitals give some examples of legitimate interest:
Recital 47: "Such legitimate interest could exist for example where there is a relevant and appropriate relationship between the data subject and the controller in situations such as where the data subject is a client or in the service of the controller".
Recital 47: processing for direct marketing purposes or preventing fraud;
Recital 48: transmission of personal data within a group of undertakings for internal administrative purposes, including client and employee data;
Recital 49: processing for the purposes of ensuring network and information security, including preventing unauthorised access to electronic communications networks and stopping damage to computer and electronic communication systems;
Recital 50: reporting possible criminal acts or threats to public security to a competent authority.

Legitimate interests must be weighed against data subject's rights and risks. They must be proportionate, clearly explained, more than economic in nature and of course make processing necessary. However, the use of this legal basis is often difficult in practice.

As the ICO explains in his online Guide on the General Data Protection Regulation[74]: "*legitimate interests is the most flexible of the six lawful bases. It is not focused on a particular purpose and therefore gives you more scope to potentially rely on it in many different circumstances*".

➢ *Special rules for specific processing*

There are special rules regarding data concerning criminal convictions and offences and Member States can determine more precisely the requirements for processing and also can determine other measures for lawful processing, among others in the scope of provisions regarding specific processing situations which are foreseen in Chapter IX of the GDPR.

Risks specific to consent in the context of the ePrivacy Directive

The widespread use of electronic communications is likely to reveal special categories of personal data, either explicitly or because of the combination of content and metadata. In turn, this may jeopardise the data subject's privacy.

Confidentiality of communications is a fundamental right protected under Article 7 of the Charter of Fundamental rights of the European Union. It is recognized by the ePrivacy Directive which lays down the protection of electronic communications, including the confidentiality of the users' communications. In line with what the EDPB says in its statement on the revision of the ePrivacy Directive and its impact on the protection of individuals with regard to the privacy and confidentiality of communications, this "confidentiality must be applied to "every electronic communication", regardless of the means by which such communication is sent, at rest or in transit, from the sender to the receiver, and must also protect the integrity of every user's terminal equipment".

In a nutshell, the ePrivacy Directive has established a general prohibition upon carriers of electronic communications (as DUET likely could be) to process electronic communications and metadata, except for:

(a) With the prior consent of the user subscriber (natural or legal person). Under Article 5(3) of the Directive, consent of users is required to store any information on an individual's terminal equipment or to gain access to the information stored;

(b) Or if they meet one of the exceptions thereof, namely transmission of an electronic communication (Article 5(3), unless the storage or access is made with the sole purpose of facilitating the transmission of a communication), or billing (strictly necessary to provide an information society service explicitly requested by the user).

[74] ICO, When can we rely on legitimate interests?, available at: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/legitimate-interests/when-can-we-rely-on-legitimate-interests/.

The ePrivacy Directive provisions also ensure the Integrity of end-users' terminal equipment. This means that not only cookies, but every tracking technology is either subject to consent of the user or must fall under one of the exceptions under the ePrivacy Directive. As clarified in the case C-673/17 Planet 49,[75] this applies regardless of whether the information constitutes personal data. According to the court, user's consent must comply with the requirements of the GDPR. E.g., for an online identifier, both the GDPR and the ePrivacy Directive applies. When metadata is genuinely anonymized, then it can be processed further without consent (see Article 29 Data Protection Working Party, Opinion 05/2014 on anonymization techniques.

The ePrivacy Regulation establishes also such general prohibition for processing of communications (regardless of whether they do include personal data or not), unless when permitted under the legal basis which vary from content and metadata. On the one hand, the EDPB cautions that the ePrivacy Regulation's approach is welcome, since it is based on broad prohibitions, narrow exceptions and the use of consent. It also cautions that there should be no possibility to process electronic communications metadata based on open grounds such as "legitimate interests", which goes beyond necessary for the provision of the service. Yet, on the other hand, the ePrivacy Regulation would limit companies from processing metadata and content unless there is either consent or there are the narrowly construed exceptions. This could, according to stakeholders, clash with the GDPR, which allows legitimate interest as a lawful, flexible ground for processing. This approach could in turn impact negatively security and privacy protection for the user. Allowing the legal basis of legitimate interest would entail that processing of metadata and content data for purposes of legitimate interest (e.g. a cyber-threat) would also be lawful under the ePrivacy Regulation. This concern appears to have been integrated in the latest Council Presidency version of the text which introduces legitimate interest as ground for processing the users' data. However, strict conditions apply to it, and such legal ground for processing may not be used when the user is a child, for profiling, and when sensitive data is involved.

Another concern is the issue of scope, which touches upon the machine to machine communications (M2M). This is relevant in the context of IoT. It is first necessary to recall that the ePR would extend confidentiality rules for traditional telecommunication players such as phone companies, to internet-based services, such as Whatsapp, which would fall under the notion of electronic communication service (ESC).

The latest ePrivacy Regulation proposal suggests to expand the ECS category (or rather, make it more specific, as against an unclear text of the ePrivacy Directive) to also include M2M communications. The text (Recital 12) does not distinguish between M2M communications which entail human interaction and those that do not. As long as the transmission occurs via a public network, M2M services fall under the notion of ECS. Services offering the technical transmission of M2M should abide by obligations under the ePrivacy Regulation, which ensures that such communication should not be tampered with. M2M communications that are part of an interpersonal communication service would fall under the scope of the ePrivacy Regulation, which, again, does not distinguish between M2M communications which entail human interaction and those that do not. The business community was concerned that such overly broad scope would hinder innovation in the data economy. This concern was not fully addressed in the latest version of the proposal. Second, under a certain interpretation of the law, the end user's consent might be required before sending data via sensors. According to some stakeholders such as Amcham, obtaining such consent is impractical and not giving it would negatively impact security of IoT. Requiring consent for processing the latter, as would a driver do e.g. using the car entering the range of a new sensor network when the exchanges data with road sensors would be unfeasible (while the processing under the GDPR would be the

[75] Judgment of the CJEU, Grand Chamber, of 1 October 2019.

ground of pre-existing contract with the driver, and consent each time would not be needed). Therefore, literature suggests that M2M communication be brought outside the scope of the ePrivacy Regulation.[76]

2.3.3 Data minimisation, storage limitation and purpose limitation principles

Data minimisation

The principle of 'data minimisation' means that data controllers should limit the collection of personal information to what is directly relevant and necessary to accomplish a specific purpose. They should also retain the data only for as long as is necessary to fulfil that purpose. The data minimisation principle is expressed in Art. 5(1)(c) of the GDPR, which provides that personal data must be "*adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed*". In other words, data controllers should collect only the personal data they really need, and should keep it only for as long as they need it.

Risks identified: ensuring data minimization in an IoT context

First, storing large amounts of data – among them personal data - increases the likelihood of potential data breach compared to the scenario where a lower amount of data has been collected. In the context of smart cities, it is not always possible to minimize the collection of data, since the smart city often relies on big data.

Second, collecting and storing large amounts of data also increases the risk of "*using the data in a way which departs from the consumers' reasonable expectations*"[77]. In an IoT context, the processing is more diversified and multi-purpose, cross-organization and cross device. Several (public and private) actors can be involved in processing operation(s) as controllers, joint controllers, processors and sub-processors. Smart cities are nurtured by big data sharing. The sole concept of data sharing appears to contrast the notion of data minimisation, as a smart city is characterized by multiple interacting data flows, multiple and varying data owners/controllers and different jurisdictions for storage and processing. The city mayor or municipal government may well feel they have the power and duty to control the final design – but actual control may rest with private vendors or investors and their sub and sub-sub-providers in the Cloud (see South Korean example of the city of Songdo or the Sidewalk Toronto example, Annex to the Deliverable). Future cities may even have 'adaptive architectures' which begin to decide themselves what data to collect and how to process it. Algorithms may by their nature tend to opacity and change as they learn in ways such that even data controllers may have low visibility on what exactly is happening in their data silos and conduits.

Storage limitation principle

The above considerations bring us to the principle expressed in the requirement that personal data are kept in a form that permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed (Art. 5(1)(e)GDPR), so-called storage limitation principle). For archiving purposes in the public interest, scientific, or historical research, or statistical purposes personal

[76] S. Storms, Quo Vadis ePrivacy: Confidentiality of Machine to machine communications, 2018.
[77] Podnar Zarko, I. et al, Interoperability and Open-Source Solutions for the Internet of Things, 2016, page 113.

data may be stored for longer periods, provided the safeguards pursuant to Art. 89(1) GDPR are adopted, which will be discussed under Section 2.4.

Risks identified: storage limitation and large scale routing in information networks.

In the context of real-world applications at a large scale, abiding by the principle of storage limitation enshrined under the GDPR may prove a challenge. In particular, large scale routing (the coordination of routing between multiple routing domains), requires data to be stored: compliance with the abovementioned principle would require storage resting upon personal data to be done for no longer than necessary. It could be then helpful to explore whether for public interest purposes such storage can be longer. However, to this end, it will be necessary, as seen above, to abide by Art. 89(1) of the GDPR, which requires the data controller to put in place adequate technical safeguards and procedures. They are discussed under Section 2.4.

Purpose limitation

The principle of purpose limitation essentially requires that personal data may only be processed for the original purpose of collection of the data, or in the words of the OECD Privacy Guidelines[78], at least, so long as it is not incompatible with the original purpose. This principle is enshrined under Art. 5(1)(b) of the GDPR, and requires personal data to be collected for specified, explicit and legitimate purposes, and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific, or historical research, or statistical purposes may be allowed subject to compliance with Art. 89(1) GDPR (not-incompatible purposes clauses).

The principle compels the controller to be clear from the outset why is personal data collected and what is the intended use of it, and if data is used or disclosed for purposes other than the initial ones, the new use is fair, lawful (i.e. on the basis of the above article 89(1) GDPR) as well as transparent (see ICO Guidance on GDPR[79], Principle b) purpose limitation). The controller must thus specify the purpose of the data processing, as a precautionary protection instrument obliging the data controller to prevent and minimize specific risks caused by the processing of such personal data against the individual's fundamental rights to privacy, freedom and non-discrimination.

Once data is collected for a specific purpose, the GDPR does not ban using it for other purposes. Yet, the new purpose must be compatible with the original purpose (as specified under Art. 5(1)(b), i.e. archiving purposes in the public interest, scientific, or historical research, or statistical purposes), or the individual's specific consent is obtained for this new purpose or a clear legal provision is identified (including under national law) requiring or allowing the new processing in the public interest.

To decide whether a new purpose is compatible, as per the GDPR, the following must be taken into account:

- any link between the original purpose and the new purpose;
- the context in which the personal data is originally collected;

[78] OECD, Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, (2013), available at: https://www.oecd.org/internet/ieconomy/privacy-guidelines.htm.
[79] ICO, Guide to the General Data Protection Regulation (GDPR), available at: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/.

- ➢ the nature of the personal data;
- ➢ the possible consequences for individuals of the new processing;
- ➢ whether there are appropriate safeguards.

Risks associated with the purpose limitation principle: function creep

The risk is basically purpose or function creep, which has been described as "*the use of technology to perform a function it was not originally intended for"*.
The risk of re-purposing without consent or knowledge from the data subject, or lacking another lawful legal basis, is obvious in a smart city context. The data generated in the course of using smart city services can be used to personalize the service, which can be beneficial, but also to profile the customer, for example. In particular, "*The use of data for a different goal than it was collected for results in purpose creep. The purpose of the IoT to realize a smooth functioning information society may (also) turn into the perfect tool to realize a surveillance society*"[80].

Relatedly, when re-purposing occurs, the fundamental rights of the individual concerned could be jeopardized. This could occur at different stages of the data processing: while the classic rights to privacy, such as at home or of communications, are typically concerned the moment that personal data is collected, a risk against the fundamental rights to freedom rather arises through the later use of data. In a data-driven innovative smart city context, data controllers are hardly able to predict, when the data is first collected, all possible future purposes of data processing because the outcome of innovation processes is hardly predictable. Yet, the principle of purpose limitation does not require data controllers to predict all possible purposes in advance, as long as, as seen above, all latter uses are fair and transparent.
The principle of purpose limitation requires the controller to limit the data processing to the initial agreed or declared purpose (in relation to which a lawful ground for processing exists), and, in doing so, aims to limit the risk caused by the subsequent data processing for purposes other than the original ones.

2.3.4 Integrity and confidentiality of data

Creating reliable wireless connectivity among devices is one of the challenges in IoT. In accordance with Art. 5(1)(f) of the GDPR, the personal data must be processed in a manner that ensures the appropriate security of such data, including the protection against unauthorized or unlawful processing, and against accidental loss, destruction or damage.

Under the GDPR, controllers are responsible for ensuring personal data is kept secure. The data must be protected both against external threats (e.g. malicious hackers) and against internal threats (e.g. poorly trained employees).

For the purposes of this chapter it is worth recalling that under the GDPR controllers and processors must ensure that appropriate security measures are in place to prevent data – in this case personal data – from being accidentally or deliberately compromised. In addition, Art. 32(1) of the GDPR further specifies what security of processing of personal data entails, stating that: "*taking into account the state of the art, the costs*

[80] Wisman, T.H.A., "Purpose and function creep by design: Transforming the face of surveillance through the Internet of Things", European Journal of Law and Technology, Vol. 4, No. 2, 2013.

of the implementation and the nature, scope, context and purposes of processing, as well as the risk of varying likelihood and severity for the rights and freedoms of the natural persons, the controller and the processor shall implement appropriate technical and organizational measures to ensure a level of security appropriate to the risk". The ICO specifies a checklist, which will be better spelled out under Section 2.4. A sound policy entails an assessment of the risks.

When data is being transmitted via electronic communication networks, the abovementioned ePrivacy Directive rules apply on electronic communication services providers and to those who use information related to end-user's terminal devices (see, in particular, Art. 4(2) thereof). The security aspects of such legislation and what it means for DUET will be dealt with under Chapter 3.

Therefore, with respect to data integrity, the processing of personal data triggers the application of both the GDPR and the ePrivacy Directive.

Risks

The 2015 FTC paper discusses at length the security risks of smart cities. Such vulnerabilities are twofold: on the one hand, they relate to the devices (e.g,) the sensors themselves, on the other hand, they relate to communications, i.e. potential to spread vulnerabilities across networks. Examples are vulnerabilities with connected cars, smart meters which may allow burglars to spot houses which are empty.
Such risks will be tackled more in depth under Chapter 3.

2.3.5 Accuracy of data

Art. 5(1)(d) of the GDPR provides that personal data shall be accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data which are inaccurate, having regard for the purposes for which they are processed, are erased or rectified without delay ('accuracy'). The GDPR does not spell out what accurate data is. Sometimes, what the data is used for may affect whether the data is accurate.
Cities need accurate data in order to thrive. In this respect, it is worth noting that the accuracy of the data will often depend on the quality of the instruments (e.g. sensors) they rely on.
The risks concerning inaccuracy of data is relevant in the context of DUET, since reliance on policy making (for example, on traffic, smart agriculture, waste collection, etc.) depends on the accuracy of the data gathered by the sensors. The same goes for inaccurate machine behavior. Decisions and actions based upon inaccurate data are problematic. Accuracy is linked to integrity, since a security vulnerability can entail incorrect data. In turn, poor data quality obstructs high quality decision making.

2.3.6 Fairness and transparency of data

Aside from the abovementioned lawful basis for processing, Art. 5(1)(a) of the GDPR, requires the use of personal data to be fair and transparent. The fairness principle entails considering how the processing may affect the data subjects concerned and also justifying any adverse impact on the individual: namely, data

should be handled in ways that the individual can reasonably expect. This requires, for example, not misleading the data subject when the individual's personal data is collected. Transparency is linked inextricably to fairness. Transparency has been central to the EU data protection regime and is also laid down under Art. 5(1)(a) of the GDPR. The transparency principle is further laid down under Art. 13 and 14, which -*inter alia* -entails the right of the individuals to be informed, as well as Art. 17(1)(d), which entails the right for the individual to have the information erased (so called right to be forgotten). Of equal importance are also the right to access by the data subject (Art. 15 of the GDPR), as well as the right to rectification (Art. 16). All these rights correspond to obligations for the data controller. Insofar as transparency is concerned, that requires informing individuals how and why their personal data will be used.

As the Article 29 Working Party guidelines[81] highlight, transparency, fairness and accountability (which will be discussed under 2.4) are inextricably linked. Transparency is an obligation for the data controllers and processors, and relates to the following:

- the provision of information to data subjects related to fair processing. It is important that individuals be informed about: (a) purposes for processing personal data; (b) retention periods; (c) who will this data be shared with.
- How data controllers communicate with data subjects with respect to the data subjects' rights.
- How data controllers facilitate the exercise by data subjects of their rights.

This information should explain in clear, and easily accessible language the rights of the data subject. In particular, the details of automated decision-making, including profiling, should be provided to the data subject.

Risks associated with fairness and transparency

First, concerning fairness, the risk is treating personal data so as to create bias. This is more so the case when automated data processing is at stake since these modalities could entail profiling and risk putting in place discrimination practices or repressive measures to the detriment of individuals' freedom of expression. Another risk related to fairness is for the data controller to obtain consent by the data subject by misrepresentation.
The right of the individual not to be subject to automated decisions, including profiling, is laid down under Art. 22(1) GDPR. As a result, a data subject has the right not to be evaluated on the basis of automated processing only, unless, as foreseen under Art. 22(2), the decision is either necessary for entering to or the performance of a contract between the data subject and the data controller, or such decision is authorized by EU or Member State law, or is otherwise based on the data subject's explicit consent. This is subject to safeguards foreseen under Art. 22(3) GDPR. Yet, under Art. 22(4) of the GDPR, no such automated decision making can be done with respect to special categories of data (sensitive data) pursuant to Art. 9 GDPR, unless specific measures are taken to safeguard the individual's fundamental rights. Other important rights enshrined in the GDPR, include the right of access, the right to rectification, the right to erasure, the right to be forgotten, mentioned above, as well as the right to restrict processing (Art. 18), the right to data portability (Art. 20), the right to object (Art. 21).

[81] Art. 29 Working Party Guidelines on transparency under Regulation 2016/679 Adopted on 29 November 2017, available at: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=622227 .

An identified risk, related to the abovementioned principles, is that data collected to serve citizens and visitors could be misused by the data controllers and processors. One such misuse is selling the personal data to third party advertisers without obtaining the data subject's consent. Processing of personal data for secondary uses (e.g. research) is also a very sensitive area (Art. 6(4) GDPR, since in this respect Art. 89(1) GDPR provides an exemption from certain users' rights, by providing that controllers that process personal data for research purposes must implement "appropriate safeguards". Indeed, under Art. 6(1)(f) GDPR personal data may be processed by an organization for research purposes without obtaining consent under the abovementioned legitimate interest ground for processing. In this respect, it is worth distinguishing, on the one hand, research purposes, from commercial purposes (Article 29 Working Party had also found "marketing research" to be considered as a legitimate interest). The GDPR adopts a broad definition of research under Recital 159, including activities of public and private entities. In addition, under Recital 47, processing for direct marketing purposes can be regarded as carried out for a legitimate interest. This requires a balancing with the data subject's rights. When the processing is done on the basis of research, a focus on the principles of transparency must be had. A risk is until how far does research extend to. Even if consent is not needed, when the data is processed for research, Art. 12(1) GDPR requires controllers to take appropriate measures to inform data subjects of the nature of the processing activities and the rights available to them, in a "concise, transparent, intelligible and easily accessible form, using clear and plain language". Such notice must be given at the time the data is collected, but also later on when a controller intends to further process data for a different purpose, including for research. Providing up front notice poses a challenge since it may be difficult to *ex ante* identify the purposes, especially in the context of big data or data mining.

2.4 Risk mitigation plan: Privacy by design

Privacy by design (PdB) is an approach to protecting privacy by embedding it into the design specifications of technologies, business/organisational practices, and physical infrastructures. Privacy by design solutions that could be particularly relevant to the DUET project are:

- restricting the amount of data applications collect to the minimum.
- encrypting data flows as default.
- anonymisation and pseudonymisation of personal data.
- embedding privacy notices systems in user-friendly ways at appropriate times.
- restricting the retention periods of data ('data expiry').
- providing easy-to-understand menus of privacy settings in clear language.
- using flash cards to make system designers think about privacy issues as they build their systems.
- Engineers and coders training and awareness on privacy requirements and legislation. They should be able to incorporate privacy concerns into the applications they develop and shape. New engineering approaches able to implement functionalities and features responding to privacy and data protection concerns.[82]

The most radical solution via PbD to the problems around the IoT is to have data collected by devices locally (and as far as possible processed locally), in this way maintaining them under the control of the user. While

[82] Edwards, L., cited.

this solution, known in the computer science world as 'personal data containers', is receiving a great deal of attention from researchers, it does not apply to a Cloud infrastructure.[83]
Recital 78 of the GDPR[84] provides guidance, by stating that data controllers should be able to demonstrate compliance with the Regulation by adopting internal policies and measures in accordance with the principles of data protection by design and default.

- Privacy Impact Assessments (PIAs) are one approach to making PbD more viable and effective. They are also mandated by the GDPR in certain cases. It may be recommended to make smart-city Data Protection Impact Assessments **participatory and collaborative**, so as to enhance data protection and societal acceptance (trust) of the proposed smart-city innovations.

- The GDPR anticipates that "*an approved certification mechanism*" may be used to demonstrate compliance with the data protection by design and default mandate (GDPR, art 25(3)). Art. 42 and Art. 43 allows for the certification of data protection compliance by certification bodies endorsed by the relevant supervisory authority or other appropriately empowered authority (GDPR, Artt. 43-34).

2.4.1 Anonymisation/pseudonymisation techniques

- *Anonymisation*

Recital 26 GDPR clarifies that data protection rules do not apply to anonymous information. The rationale of this exclusion is based on the reasoning that there is no privacy harm – and as a result no privacy interest – implicated in the processing of non-personally identifiable data.

Anonymisation, orde-identification refers to the process of collecting or changing a dataset in a way that individuals are no longer identifiable. This can be done either at the point of data collection (de-identification at source), or at a later stage. The benefits of anonymisation are: issues of consent do not exist, the data can

[83] *Id.*

[84] Recital 78 GDPR: Appropriate Technical and Organisational Measures:

1. The protection of the rights and freedoms of natural persons with regard to the processing of personal data require that appropriate technical and organisational measures be taken to ensure that the requirements of this Regulation are met.
2. In order to be able to demonstrate compliance with this Regulation, the controller should adopt internal policies and implement measures which meet in particular the principles of data protection by design and data protection by default.
3. Such measures could consist - *inter alia* - of minimising the processing of personal data, pseudonymising personal data as soon as possible, transparency with regard to the functions and processing of personal data, enabling the data subject to monitor the data processing, enabling the controller to create and improve security features.
4. When developing, designing, selecting and using applications, services and products that are based on the processing of personal data or process personal data to fulfil their task, producers of the products, services and applications should be encouraged to take into account the right to data protection when developing and designing such products, services and applications and, with due regard to the state of the art, to make sure that controllers and processors are able to fulfil their data protection obligations.
5. The principles of data protection by design and by default should also be taken into consideration in the context of public tenders.

be exported internationally, the data can be kept for however long the controller wants to. However, where de-identification happens only at a later stage and not at the source, all the earlier stages of processing handle personal data and are thus fully subject to GDPR provisions.
Where personal information is needed to provide the service, anonymisation is not an option and consent will have to be obtained or another legal basis for processing needs to be envisaged.

However, data points can be correlated across different databases and "*it is now rare for data generated by user activity to be completely and irrevocably anonymised*"[85]. "*In an age of big data the strategy of deidentification provides only a weak form of privacy because it is possible to reverse engineer the process by combing and combining data sets*"[86] . In fact, evidence suggests that just four "*spatio-temporal points*" are required in order to uniquely identify 95% of individuals[87]. The risk of re-identification may thus be inherent in many datasets.

De-identifying data at source eases many privacy concerns, but it also may strip the data of its most valuable details. It is "*a bit like building a powerful telescope to see far into space, only to put frosted glass over the lens*"[88]. University of Toronto professors David Lie and Lisa Austin have proposed a solution in what they call safe sharing sites, a type of technical and legal interface for privacy-protective sharing of personal information.

Data sanitisation techniques

Article 29 Working Party Opinion on Anonymisation Techniques[89] examines the robustness of data sanitisation techniques against those risks.

Data sanitisation techniques process data in a form that aims to prevent re-identification of data subjects. Randomisation and generalisation are considered as two main families of sanitisation techniques.

Article 29 Working Party distinguishes data sanitisation techniques into 'randomisation', 'generalisation', 'masking direct identifiers' and 'pseudonymisation'.

- ➢ Randomisation (noise addition, permutation and differential privacy) and generalisation (k-anonymity, l-diversity) are methods of anonymisation.
- ➢ Masking direct identifiers as a security measure.
- ➢ Pseudonymisation.

Techniques under the randomisation group aims at altering the veracity of data. Examples are 'noise addition', 'permutation' and 'differential privacy'. More specifically, noise addition and permutation can reduce linkability and inference risks, but fail to prevent the singling out risk. Differential privacy is able to

[85]Article 29 Working party, Opinion 05/2014 on anonymisation techniques, 10 April 2014, available at: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/legitimate-interests/when-can-we-rely-on-legitimate-interests/ . Ohm, P., Broken Promises of Privacy: Responding to the surprising failure of anonymisation, 2010. Edwards, cited.
[86] Kitchin, cited; Narayanan, A., and Shmatikov, V., Robust de-anonymisation of large sparse datasets, 2010.
[87] De Montjoye, Y-A, Hidalgo,C.A, Verleysen, M. and Blonde, V., Unique in the crowd: the privacy bounds of human mobility, 2013; see also Ohm, cited.
[88] Ryan, A., Can smart cities help their residents without hurting their privacy?, Quartz, Yahoo Finance, 27th November 2019.
[89] Article 29 Working party, Opinion 05/2014 on anonymisation techniques, 10 April 2014, available at: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/legitimate-interests/when-can-we-rely-on-legitimate-interests/ .

prevent all the risks up to a maximum number of queries or until the predefined privacy budget is exhausted but queries must be monitored and tracked when multiple queries are allowed on a single dataset.

As regards the generalisation category, 'K-anonymity' is considered robust against singling out, but linkability and inference risks are still present. 'L-diversity' is stronger than K-anonymity provided it first meets the minimum criterion of k-anonymity, as it prevents both the singling out and inference risks.

The GDPR definition of pseudonymisation is more restrictive than merely masking direct identifiers. Masking direct identifiers is conceived as a security measure by the Article 29 Working Party because it does not mitigate the three risks. It rather simply removes/masks the direct identifiers of data subjects.

➢ *Pseudonymisation and encryption*

The processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person" (Art. 4(5) GDPR). This additional information is kept separately and is secured through organisational or technical measures (e.g. encryption). For instance, a research study would qualify as pseudonymisation, if the personal data of study participants would be replaced by unique attributes (e.g. number or code) in the research documentation and their personal data would be kept separately with the assigned unique attributes in a secured document (e.g. in a password protected database)[90]. Nonetheless, data which have been pseudonymised are still considered information about an identifiable person if they can be attributed to this person by using additional information and, as such, that data would constitute personal data in the meaning of the GDPR.
The pseudonymisation of data can be used as evidence of data protection by design and the implementation of appropriate security measures within an organisation (GDPR, arts 25(1) and 32(1)(a)).
Pseudonymisation is the only technical or organisational measure explicitly mentioned in Art. 25 GDPR, signaling it be considered good practice.

➢ *Contextual controls*

Recital 78 GDPR elaborates on additional measures aside from prompt pseudonymisation that can minimise the processing of personal data and suggests "***transparency*** *with regard to the functions and processing of personal data*", enabling "*the data subject to monitor the data processing*", and *"enabling the controller to create and improve security features*" .

On this note, contextual controls of a legal, organisational and technical character are necessary to help tackle the risks of re-identification.

These comprise three sets of controls:

- Legal and organisational controls such as obligations between parties and/or internal policies adopted within one single entity aimed at directly reducing re-identification risks, e.g. obligation not to re-identify or not to link.

[90] EUROPEAN COMMISSION, COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL, Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union, Brussels, 29.5.2019, COM(2019) 250 final.

- Security measures such as data access monitoring and restriction measures, auditing requirements, monitoring of queries, aimed at ensuring the de facto enforcement of the first set of controls.
- Legal, organisational and technical controls relating to the sharing of datasets aimed at ensuring that the first set of legal controls are transferred to recipients of datasets. They include obligations to share the datasets with the same set of obligations or an obligation not to share the datasets, as well as technical measures such as encryption to make sure confidentiality of the data is maintained during the transfer of the datasets. These measures are used to balance the strength of data sanitisation techniques with the degree of data utility. [91]

In practice, the selection of contextual controls depends on a specific data sharing scenario.

2.4.2 Lawful grounds for processing of personal data: consent and safeguards when consent cannot be handily obtained

First, the question of how to obtain consent in a smart city scenario arises. The US Federal Trade Commission has come up with a number of existing good practices to obtaining such consent designed to be as unobtrusive as possible, including:

- directing customers to video tutorials to guide them through privacy settings pages (drawn from Facebook) or alternately providing set up wizards to get data collection choices right.
- homes or other locations might have detailed control dashboards or management portals where consumers could review with some clarity what data they had chosen to share from time to time across different applications or via different devices.
- putting QR codes on IoT devices, which could be scanned by customers using their smartphones, to give them easy access to privacy policies or other advice.
- providing icons to convey privacy-related information, such as a flashing light that appears when an IoT device connects to the Internet; different icons might flash up to show different levels of risk, and/or different types of data collection.
- Customers might ask ‘just in time’ for privacy and security settings to be sent to them via emails or texts .

An alternative approach is to reconsider how consent might be given in the IoT world, conceiving it as an ongoing process, rather than a one-time choice at the point of data collection.

Another suggestion consists in decoupling the time of giving consent from the time of collection of data, which is that of ‘sticky privacy preferences’. The idea here is that the privacy choices you made earlier are remembered by smart systems, and applied the next time a choice needs to be made. The FTC suggests that a single device in a smart home–a home appliance that acts as a hub–could learn a consumer’s preferences based on prior behaviour and apply them to new appliances and new uses.

[91] Runshan, H., and Stalla-Bourdillon, S., and Yang, M., and Schiavo, V. and Sassone, V., “Bridging Policy, Regulation, and Practice? A Techno-Legal Analysis of Three Types of Data in the GDPR (September 1, 2017)” in 'Data Protection and Privacy: The Age of Intelligent Machines' Edited by Ronald Leenes Rosamunde van Brakel, Serge Gutwirth and Paul De Hert, Hart Publishing, (2017).

Note, however, that GDPR is fairly strict in its requirements on the quality of data subjects' consent with personal data processing (see above), and any innovative approach to obtaining consent should be designed with these in mind.

During the VUB-SMIT Roundtable Personal data protection in Smart Cities that took place in September 2019 in Brussels[92], some participants suggested considering the 'public interest' legal ground for processing as a solution for data sharing in public spaces and for the public benefit. In this regard, it was emphasised that - albeit the right to personal data ranks higher in the legal sources hierarchy and the control by the data subject over his personal data must be warranted through a solid control architecture - the level of control may dilute in a data sharing scenario. As seen above, the public interest legal basis applies when processing is necessary for the performance of a public task and should be "*laid down by Union law or Member State law*", thus meaning that express statutory powers to share might be needed, unless one would argue in favor of implied legal powers bestowed upon public municipalities. The first solution appears to be the most loyal to the principles of legal certainty and the rule of law. National laws should provide for express legal gateways for data sharing in public, as Art. 6 (2) GDPR allows for. For instance, the UK Digital Economy Act of 2017 has granted public authorities powers to share personal information across organisational boundaries to improve public services, e.g. e-government, enable better public services[93].

2.4.3 Risk mitigation to abide by the purpose limitation and the data minimisation principle: the case of video-surveillance

The use in public spaces of video devices that process a massive amount of personal data is part and parcel of smart city appliances, e.g. smart cameras and video analysis softwares. The risk of violations of privacy rights and discriminatory outcomes is recognised by the GDPR, where it requires a **data protection impact assessment** in the case of systematic monitoring on a large scale of a public accessible area (Art. 35 (3) (c)) and **the designation of a DPO** in case of regular and systematic monitoring of data subjects on a large scale (Art. 37 (1)(b)). The **EDPB Guidelines 3/2019 on processing of personal data through video devices[94]** give guidance on the legal requirements and exceptions that personal data processing through video devices shall meet in order to avoid secondary use or "*misuse for totally different and unexpected purposes*"[95], e.g. marketing, employee performance monitoring.

The EDPB Guidelines 3/2019 can be summarized as follows:

- Video surveillance systems employ techniques entailing different degrees of intrusiveness: certain techniques can be more intrusive, e.g. complex biometric technologies, other techniques can be more privacy-friendly (simple counting algorithms). The data protection issues vary from case to case.
- Issues lie also in the state-of-the art technology that can still be inaccurate and induce biases. It is reported that software used for facial identification, recognition or analysis performs differently

[92] VUB Chair Data Protection on the ground, "Personal data protection in smart cities", Roundtable report, September 2019. See also, Christofi, A., "Sharing personal data to build the smart city: legal barriers and enablers", (2019).
[93] *Id*.
[94] EDPB, Guidelines 3/2019 on processing of personal data through video devices, (2019).
[95] *Id*., p. 4.

based on the age, gender, and ethnicity of the identified person. In this regard, the EDPB recommends data controllers to subject data processing through the surveillance devices to regular assessment.

- The GDPR does not apply to the processing of personal data by a natural person in the course of a purely personal or household activity. According to the EDPB restrictive interpretation, this exemption shall cover only the private and family activities of the person.

- The purposes of processing have to be documented in writing and need to be specified for every camera (also collectively if more cameras are used for the same purpose). Data subjects must be informed of the purpose(s) (transparency and information obligations under Art. 13 GDPR). The mere purpose of safety is not sufficiently specific (Art. 5(1)(b)).

- Legal bases can be grounded on Art. 6(1). The provisions most likely to be used are Art. 6(1)(f) (legitimate interest) and article 6(1)(e) necessity to perform a task carried out in the public interest or in exercise of official authority). In exceptional cases Art. 6(1)(a) of the GDPR (consent) might be used as a legal basis by the controller.

- In order to assess the existence of a valid legitimate interest pursued by a controller or a third party it is necessary to carry out and document a balancing test taking into account the interests, rights and fundamental freedoms of the data subjects.

- "*Personal data should be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation'), see Art. 5(1)(c) GDPR. Before installing a video-surveillance system the controller should always critically examine if this measure is firstly suitable to attain the desired goal, and secondly adequate and necessary for its purposes*"[96].

- Any disclosure of videos to third parties needs to have an autonomous legal basis.

- Video surveillance may entail the processing of special categories of personal data. In order to assess whether a specific legal basis pursuant to Art. 9 GPPR is necessary, it shall be addressed if processing of particular categories of data is the objective of the data controller.

- In order to ensure compliance with transparency and information obligations, data controllers should inform data subjects by first displaying a warning sign that a video camera is installed in a certain place and, second, providing mandatory details and information with the privacy information notice.

- Video surveillance images must be kept only for the time strictly necessary for the pursued purposes (ideally a few days).

[96] *Id.*, p. 8.

2.4.2 Processing of sensitive categories of data: safeguards in place under the GDPR

Prospectively, future DUET implementations will demand a careful consideration on whether and how to handle sensitive data and optimise/channel their use for the public good. Sensitive data are special categories of data that, by virtue of being rooted in a paramount value, i.e. a fundamental human right or fundamental freedom, are subject to a stricter legal regime/protection.

Under the GDPR, sensitive data are those revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, genetic data, biometric data (where used for identification purposes), data concerning a person's sex life and orientation and, notably health data (Art. 9 GDPR).

Among all the above mentioned kinds of sensitive data, health data are more likely to be of good use for DUET's future applications, not just for the benefit of private-to-private patient-physician interactions, but also to help the public tackle a health crisis/pandemic.

At present, health data sharing for the pursuit of public interest is at the centre of the public debate and it is being implemented by a number of States across the globe. In China, South Korea, Singapore and Israel, just to mention a few, mobile phone data tracking, public mapping of infected individuals and mass surveillance techniques are being used to monitor and enforce lockdown, quarantine and social distancing policies[97]. The focus lies in the trade-off between the interest of public health and the principles of democracy, good governance, due process and the fundamental human rights and freedoms, among which the right to protection of personal data.

In a statement published on 6th April 2020, the European Data Protection Supervisor (EDPS), has clarified that measures that weaken the protection of the right to privacy should comply with both a necessity and a proportionality test. "*The GDPR clearly states that the processing of personal data should be designed to serve mankind [...]. The GDPR states also that the right to* ***the protection of personal data is not an absolute right; it must be considered in relation to its function in society and be balanced against other fundamental rights, in accordance with the principle of proportionality.*** *Legality of processing the personal data – even so called sensitive data like data about health – can be achieved when processing is necessary for reasons of substantial public interest, on the basis of Union or Member State law which shall be proportionate to the aim pursued. [...] The GDPR also permits processing of sensitive data when it is* ***necessary for reasons of public interest in the area of public health, such as protecting against serious cross- border threats to health.***"

The EDPS further asked Member States to urgently adopt a harmonised approach for the protection of personal data when tackling the COVID-19 crisis. The EDPS also called for a pan-European model COVID-19 mobile application.

Moreover, the EDPS confirms that, given the urgency and nature of the crisis, exceptional measures can be justified, but these should observe safeguards designed to prevent a lasting impact on fundamental rights and freedoms. As a result, exceptional measures should (i) be temporary; (ii) be limited to the specific

[97] Renda, A., (2020), Will privacy be one of the victims of covid-19?, CEPS, available at: https://www.ceps.eu/will-privacy-be-one-of-the-victims-of-covid-19/ .

purpose of fighting the COVID-19 crisis; (iii) restrict access to the data; and (iv) contain rules governing the fate of the data after the crisis.

The above instructions are fully in line with the general approach taken by the same body in its guidelines on necessity and proportionality[98].
The Guidelines start with the premise that the fundamental rights to privacy and the protection of personal data are enshrined in Art.s 7 (right to respect for private life) and 8 (right to the protection of personal data) ECHR and that those rights are not absolute and may be limited, provided that the limitations comply with the requirements laid down in Art. 52(1) CFR, i.e. measures must be necessary and proportionate to the aim to be achieved.
Art. 52 CFR requires a measure to be compliant with the following criteria:

- it must be provided for by law,
- it must respect the essence of the rights,
- it must genuinely meet objectives of general interest recognised by the Union or the need to protect the rights and freedoms of others,
- it must be necessary, and
- it must be proportional.

Necessity is a fundamental principle when assessing the restriction of fundamental rights, such as the right to the protection of personal data. According to the case-law of the Court of Justice of the European Union (CJEU)[99], because of the role the processing of personal data entails for a series of fundamental rights, the limiting of the fundamental right to the protection of personal data must be strictly necessary.
Necessity shall be justified on the basis of objective evidence and is the first step before assessing the proportionality of the limitation. Necessity is also fundamental when assessing the lawfulness of the processing of personal data. The processing operations, the categories of data processed and the duration the data are kept shall be necessary for the purpose of the processing.

Proportionality is a general principle of EU law. It restricts authorities in the exercise of their powers by requiring them to strike a balance between the means used and the intended aim. In the context of fundamental rights, such as the right to the protection of personal data, proportionality is key for any limitation on these rights.
More specifically, proportionality requires that advantages due to limiting the right are not outweighed by the disadvantages to exercise the right. In other words, the limitation on the right must be justified. Safeguards accompanying a measure can support the justification of a measure. A precondition is that the measure is adequate to achieve the envisaged objective. In addition, when assessing the processing of

[98] See supra n. 25. See also EDPS, "The EDPS quick guide to necessity and proportionality", 20 January 2020, available at: https://edps.europa.eu/data-protection/our-work/publications/factsheets/edps-quick-guide-necessity-and-proportionality_en .

[99] CJEU, joined cases C-293/12 and C-594/12, Digital Rights Ireland, paragraphs 34 - 36; see also joined cases C-92/09 and C-93/09 Volker und Markus Schecke, paragraph 58. 33 See for instance, joined cases C-92/09 and C-93/09 Volker und Markus Schecke, paragraph 55 and joined cases C-468/10 and C-469/10, Asociación Nacional de Establecimientos Financieros de Crédito (ASNEF) and Federación de Comercio Electrónico y Marketing Directo (FECEMD), v Administración del Estado, paragraph 41. The CJEU held only in one case that there was no limitation on the right to private life when the personal data related to salaries were processed by the employers for their original purpose, see CJEU, Joined Cases C-465/00, C-138/01 and C-139/01, Rechnungshof et al v. Österreichischer Rundfunk, paragraph 74. 34 CJEU, Joined Cases C-465/00, C-138/01 and C-139/01, Rechnungshof et al v. Österreichischer Rundfunk, paragraph 75 and joined cases C-293/12 and C-594/12 Digital Rights Ireland, paragraph 33.

personal data, proportionality requires that only that personal data which is adequate and relevant for the purposes of the processing is collected and processed.

2.4.3 Risk mitigations on location data under the GDPR and the ePrivacy Directive

Guidance 4/2020 of the EDPB[100] provides useful information on how location data can be used and shared in full respect with EU law. While it has been issued in the context of the COVID-19 emergency, it can provide operational guidance for DUET in the context of its activities. The following text complements **subsection 2.2.1** above providing a focused overview of certain ePrivacy rules development.

The EDPB clarifies that:

(a) location data can be collected by electronic communication service providers in the course of the provision of the service.
(b) location data collected by information society services providers' applications whose functionality requires the use of such data (e.g. navigation, transportation services and so on).

Under (a), the data collected by electronic communication service providers – which may, but need not to contain, personal data - can only be processed under the one of the grounds of Art. 6(1) of the ePrivacy Directive (allowing the processing of location data included in traffic data for the purpose of transmitting a communication) and in accordance with Art. 9(1) of the ePrivacy Directive. Under Art. 9, location data may be processed when they are made anonymous or with the consent of the user or subscribers. When data is anonymized (and the EDPB provides guidance to this end), preference should be given to the processing of anonymized data, rather than personal data. On the other hand, when the data is personal data, then also the GDPR applies. In this respect, the EDPB clarifies that a dataset can be made anonymous as a whole or not. When a single data pattern is anonymized but not the whole data set, then the data set is considered pseudonymised only and thus the GDPR fully applies. In particular, when data is anonymized, a risk mitigation is to ensure that they are not subject to re-identification. Indeed, scientific research has shown that location data thought to be anonymized may not be and they may be vulnerable to re-identification under certain circumstances.[101] The EDPB recommends that location data must be carefully processed to meet a reasonability test, which requires taking into account objective aspects and contextual elements which may vary from case to case. In addition, "*transparency concerning anonymization techniques is highly encouraged*".

A former study carried out on the interplay between the GDPR and the ePrivacy Directive[102] clarifies this interplay. First, in this latter case, a lawful ground for processing under the GDPR must exist. Second, the purpose limitation principle must be abided by: location data containing personal data may be collected only for specified, explicit and legitimate purposes and cannot be further processed in a manner incompatible

[100] See supra n. 34.

[101] De Montjoie, et al, "Unique in the crowd: the privacy bounds of human mobility", cited. De Montjoie *et a*l, On the privacy-conscientious use of mobile phone data, 2018.

[102] EPDB Opinion 5/2019 on the interplay between the ePrivacy Directive and the GDPR, in particular regarding the competence, tasks and powers of data protection authorities

with those purposes. When location data containing personal data is processed for other purposes than that for which it has been collected, then under Art. 6(4) of the GDPR a compatibility assessment must be conducted. This means that appropriate safeguards are taken in this respect. In addition, other GDPR principles, such as data minimization, and storage limitation must be observed.

In particular, as the EDPB observes, the data can be transmitted to authorities or other third parties if they have either been anonymized, as required under Art. 9 of the ePrivacy Directive, or for data indicating the geographic position of the terminal equipment of a user, which are not traffic data, with the prior user consent (where the notion of consent is the one adopted by the GDPR, which then must comply with the requirements under Art. 4(11) and 7 of the GDPR).

Under (b), for information, including location data, collected from a user's terminal equipment, Art. 5(3) of the ePrivacy Directive applies: the storing of information on the user's device is allowed only if (i) the user has given consent; (ii) the storage and or access is strictly necessary for the information society service explicitly requested by the user. When the re-use of such location data is at stake, then additional safeguards must apply. In particular, when the data has been collected as per Art. 5(3) of the ePrivacy Directive, they can be further processed with the additional consent of the data subject or on the basis of EU or Member State law which constitutes a *necessary* and *proportionate* measure in a democratic society to safeguard the objectives referred to under Art. 23(1) of the GDPR (the EDPB then recalls Section 1.5.3 of the Guidelines 1/2020 on processing personal data in the context of connected vehicles).

Finally, the EDPB recalls that derogations to the ePrivacy Directive are possible under Art. 15 thereof, when "*they constitute a necessary, appropriate and proportionate measure within a democratic society*". We recall here the considerations already made by the EDPB on what is necessary and proportionate explained elsewhere in this chapter.

As mentioned in subsection 2.2.1 above, the ePrivacy Regulation may, if adopted, relax the current strict requirements on processing of location metadata and also with regard to use of information related to user's terminal equipment by introducing a legal ground that will enable lawful processing based on data controller's legitimate interests.

2.4.4 Function creep and risk mitigation

As seen, the purpose limitation principle can be thwarted as a result of making use of big data in a smart city context.
The data controller can reduce the risk against the data subjects' fundamental rights, by implementing further protection instruments, such as further rights of information or participation of the individual in the data processing. This information may be necessary in order to find a legitimate balance between the risks to the individual's fundamental rights specifically concerned and the controller's fundamental rights and, thus, in order to legitimize the data processing, overall.

By means of self-regulation mechanisms, data controllers can set up private standards for specific cases and certain purposes of data processing in a way that guarantees that the individuals' decision-making process is so designed that they can effectively and efficiently manage the risks caused by the data processing (i.e. determined by the corresponding purposes). Such standards, be it in the form of a certificate,

a code of conduct or binding corporate rules, specify the conditions of the data processing and can thus signal to the individual concerned, as well as business customers of the data controller, the level of data protection. Data controllers can hence create themselves legal certainty and use this as a competitive advantage on the market. Finally, such standards simultaneously provide the basis for two additional advantages. First, they provide the basis for further privacy-enhancing technologies. If machines shall, one day, manage the risks on behalf of the individual concerned, the purpose of the data processing and, thus, all further requirements must be formalized, in order to enable machines to communicate the requirements to each other. In particular, formalizing purposes makes it possible that a third party (potentially, a machine), which receives personal data from another party (or machine), can obtain all purposes previously specified in an automated way.

The big data assault on purpose limitation can be dealt with by a number of legal strategies:

- asking consent for plausible re-uses at the start, this is to say identify all the several possible specific and limited purposes that are likely to justify the data processing.
- obtaining new consent to re-uses of data as they arise.
- using a non-consent based ground, such as legitimate interests to make repurposing lawful.

However, some risks still remain. A one-off blanket consent to any prospective reuse risks to be too vague and fail the specific and limited purposes test. Seeking to obtain consent anew could be costly for data controllers.[103] As well the use of the legitimate interests legal basis has been criticized for it entrusts the controller with the delicate task of balancing user fundamental rights any other interests pursued with the data processing, without there being a system of oversight in place.

2.4.5 Accuracy: data sanitization techniques

DUET should ensure that it has taken all reasonable steps that the personal data processed and stored is not incorrect or misleading as to any matter of fact. In addition, the personal data, depending on the use, must be kept up to date (although for certain uses historical data may have value). When it is discovered that personal data is incorrect or misleading, reasonable steps must be taken to ensure that it is corrected or erased as soon as possible. Finally, any challenges to the accuracy of the personal data by data subjects or 3rd parties must be carefully considered and, as a matter of good practice, a note on them should be kept. Data cleansing, cleaning or scrubbing is the process which fixes or removes incorrect or inaccurate data.

2.4.6 Accountability

Accountability is closely linked with the abovementioned principles of data accuracy, integrity and confidentiality and fairness and transparency. The GDPR integrates accountability as a principle which requires that organisations (controllers and processors) put in place appropriate technical and organizational

[103] Edwards, L., cited.

measures apt at demonstrating compliance with such a piece of legislation. The principle is enshrined under Art. 5(2) of the GDPR, which provides that the controller shall be responsible for, and be able to demonstrate compliance with, paragraph 1 (which enshrines the other data protection principles). Two are then the key elements for DUET: first, DUET is responsible for complying with the GDPR. Second, it must be able to demonstrate compliance with the GDPR.

What are some of the measures in order to demonstrate compliance? There are a number of procedures that an organization can take in order to abide by this principle. The UK Information Commissioner Officer (ICO), in its **Guide to the GDPR, Accountability and Governance**, lists them:

- Adopting and implementing data protection policies. Indeed, the GDPR says at Art. 24(2), that, where proportionate, implementing data protection policies is one of the measures to take in order to demonstrate compliance with the GDPR. Such policies, in particular, should ensure various levels of protection depending on the categories of data at stake. As seen, indeed, when DUET handles large volumes of personal data, then the policies must be robust and comprehensive. The same goes for special categories of data, such as particularly sensitive information, in relation to which specific data protection policies must be drafted. In accordance with Art. 24(1) GDPR, "*taking into account the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for the rights and freedoms of natural persons, the controller shall implement appropriate technical and organizational measures to ensure and to be able to demonstrate*" that processing is carried out in compliance with the GDPR. Those measures shall be reviewed and shall be updated where necessary;

- Adopting a data protection **by design** and **by default** approach. This is enshrined under Art. 25(1) of the GDPR. This means that a data protection approach must be embedded throughout the life cycle of the project: this means, according to the provision, ensuring that, both at the time of the determination of the means of processing and of processing, appropriate technical and organizational measures must be implemented, such as the abovementioned pseudonymisation, abiding by the GDPR principles such as the data minimization principle, in an effective manner and integrating the necessary safeguards to abide by the GDPR and protect the rights of the data subject. Under Art. 25(2) GDPR, the no more than necessary processing principle – for each specific purpose - must be ensured by default. This in turn requires DUET setting out a data protection policy which ensures, for each category of data, identifying specific purpose in order to list the necessary data to be processed. This relates to the amount of personal data, as well as extent of processing, the period of their storage and accountability. As the data is usually big data, and dynamic, this static approach may prove complex to abide by. This complexity must be taken into account in the context of risk mitigation. To this end, an appropriate certification mechanism pursuant to Art. 42 of the GDPR may be used as an element to identify compliance;

- Putting written contracts in place with processors that the controller interacts with that process personal data. Whenever a controller uses a processor to handle personal data on their behalf, it needs to put in place a written contract that sets out each party's responsibilities and liabilities. The contract shall require the processor to take appropriate measures to ensure security of processing and to assist the controller in allowing individuals to exercise their rights under the GDPR. Art. 28 of the GDPR lays down what a processor must abide by. *Inter alia*, under Art. 28(3)(c), the processor assists the controller in putting forth appropriate technical and organizational measures to fulfill the

controller's obligation to respond to requests for exercising the data subject's rights, as well as demonstrative, under 28(3)(d), compliance with Artt. 32 to 36 GDPR;

- Maintaining documentation of processing activities;

- Implementing appropriate security measures. Such measures are enshrined under Art. 32(1) of the GDPR, which provides that the controller and the processor shall implement appropriate technical and organizational measures to ensure a level of security appropriate to the risk, including the pseudonymisation and encryption (see infra), the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services, the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident, as well as a process for regularly testing, assessing and evaluating the effectiveness of technical and organizational measures for ensuring security of processing. In this respect, under Art. 32(2), account shall be taken in particular, of the risks that are presented by processing, in particular from accidental or unlawful destruction, loss, alteration, unauthorized disclosure of or access to personal data transmitted, stored or otherwise processed;

- Recording and, where necessary, reporting personal data breaches. Art. 33 of the GDPR enshrines the procedures to notify a personal data breach to a competent supervisory authority by the controller without delay, and where feasible, no later than 72 hours after having become aware of it. Under Art. 33(2) such notification shall at least: describe the nature of the data breach, including where possible, the categories and approximate number of data subjects concerned, and the categories and approximate number of personal data records, communicate the name and contact details of the data protection officer or other contact point where more information can be obtained, describe the likely consequences of the personal data breach, and describe the measure taken, or proposed to be taken, to address it, including where appropriate, measures to mitigate the possible adverse effects. Art. 33(5) of the GDPR provides that the controller shall document any personal data breaches, comprising the facts relating to the personal data breach, its effects and the remedial action taken. Under Art. 34 of the GDPR such breach should be communicated, as per the procedures and under the circumstances laid down thereof, to the affected data subject;

- Carrying out data protection impact assessments for uses of personal data that are likely to result in high risk to individuals' interests. Section 3 of the GDPR (Art.s 35 and following) speaks about the data protection impact assessment. This is particularly important where a type of processing involving using new technologies is at stake as may be the case here. Under Art. 35(1), when this is the case, and taking into account the nature, the scope, the context and the purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data. Art. 35(7) contains the minimum information that the assessment must contain and that encompass a systematic description of the envisaged processing operations and the purposes of processing, including, where applicable, the legitimate interest pursued by the controller, an assessment of the necessity and proportionality of the processing in relation to the purposes, and assessment of the risks to the rights and freedoms of the data subjects as well as the measures envisaged to address those risks including safeguards, security measures and mechanisms to ensure protection of personal data and demonstrate compliance with the GDPR;

- Appointing a **data protection officer.** The appointment of the data protection officer by the controller and the processor is enshrined under Section 4 of the GDPR, and in particular, Art. 37 and following. It will be necessary to establish whether, as foreseen under Art. 37(1)(c), the core activities of the controller and the processor consist of processing on a large scale of special categories of data pursuant to Art. 9 GDPR (sensitive data) or under Art. 37(1)(b), the processing operations require regular and systematic monitoring of data subjects on a large scale; and

- Adhering to relevant codes of conduct and signing up to certification schemes. Indeed, adherence to approved codes of conduct as referred to under Art. 40 of the GDPR or approved certification mechanisms as referred to in Art. 42 may be used as an element by which to demonstrate compliance with the obligations of the controller.

It is important to consider these as elements of comprehensive holistic privacy-compliant data governance, rather than a ticking the box exercise.

3. DUET Digital Twins Security: Security by design

3.1. Purpose

With the rapid growth of technology involved in smart cities, it is becoming vital to identify and implement security controls for their secure operation. Smart city security is essential for the technologies to be incorporated in the smart city infrastructure, and establish citizen's trust in such projects.[104]

Recent smart city developments have raised the importance of cyber security concerns. For example, as free-wifi develops so do threats to security of networks arise. Also, the use of the smart grid infrastructure exposes city users to threats linked to the smart city cyber-infrastructure. For example, not only patterns of users (for example, when users are at home, etc) can be known, but also in the context of weather prediction or floods, e.g., a problem with such infrastructure would raise security issues for the smart city inhabitants. In addition, the smart city also publishes open data to the citizens about their city. Some of this open data (for example, which areas of cities have less criminality, which are the patterns of open consumption) are beneficial to citizens to make informed decisions. Yet, this open data can also be exploited by malevolent parties to create smart city security vulnerabilities. Risks are thought as a formula of vulnerability times threat, times consequence. Vulnerabilities are weaknesses in a system which give rise to specific risks when combined with a threat[105].

Ensuring the security of personal data is essential to ensuring the protection of privacy. Concurrently, companies will also be required to adopt reasonable technical, physical and administrative measures in order

[104] Ralko, Sh., Kumar, S, "Smart City Security", KSU Conference on Cybersecurity Education, Research and Practice, 2016.
[105] Cloud Security Alliance, CSA Security Guidance for critical areas of focus on Cloud Computing Security Guidance v4.0, (2017), available at: https://cloudsecurityalliance.org/research/guidance.

to protect a wide range of data, including personal data, from loss, misuse or alteration[106]. In the context of smart cities, some of the risks faced are[107]:

- unauthorized disclosure of personally identifiable information which results in security breaches. These may involve data and identity theft, for instance, in the context of parking lots, where cyber attackers risk accessing an ample amount of targeted personal information that can be potentially exploited for fraudulent transactions at the detriment of data subjects;
- device hijacking: the attacker hijacks and effectively assumes the control of a device (for example, an autonomous car, or a smart meter system).
- man in the middle (MitM): when an attacker interrupts or redirects communications between two systems.
- a distributed denial of service attack attempt which renders a machine or a network unavailable to its intended users by disrupting services of a host connected to the Internet, or a permanent denial of service attack (which requires re-installation of hardware).
- convergence of legacy and new technologies: most legacy systems do not allow for live updates or data encryption. Merging disparate technology platforms can create, for instance, holes in the security perimeter.
- integration of the digital and the physical environmental: a dense web of interconnected sensors, a diverse range of resource-constrained devices and constant flow of data among them increases the peril of having countless points of entry for attackers who seek to compromise systems[108]. Thus every endpoint represents a potential gate for attackers.

Against this background, DUET must ensure that security is established throughout the life cycle of the project and not simply *ex post*. The security risks may touch upon the various layers of the technology infrastructure: edge (the front end of the smart city, which gathers the data from IoT devices, then sends it through the communication layer to the core), core (a cloud of IoT data platform that processes data and generates output that makes sense of the data streaming from the edge), communication (connects the core and the edge by a network system such as WiFi, Bluetooth).

This Chapter will discuss the security issues that may arise in this context, and will lay down some guidelines, also drawing from best practices adopted by the industry, as to the security solutions that need to be implemented to keep the infrastructure secure. Indeed, there are a number of measures that DUET can take to minimize cybersecurity risks. The notion of “security by design” shall be explored to this end.

3.2 Legal landscape

This Section contains an overview of the EU-level legal instruments dealing with the security aspects of the technologies that the DUET Digital Twin will consist of.

At the outset, a clarification must be made. Throughout the EU, as well as at Member State level, legal frameworks have been adopted requiring public and private organisations to safeguard the security of information systems. This requirements plan will presuppose that EU laws – or Member State laws

[106] *Id*.
[107] Koren, A., “The Biggest Smart City Security Challenges in 2019”, September 16, 2019.
[108] ENISA, Cyber Security for Smart Cities,Guidelines, (2015).

transposing them - apply. However, the caveat is that when cloud computing is at stake, various laws may apply concurrently, in accordance with the following[109]:

- the location of the cloud provider;
- the location of the cloud user;
- the location of the data subject(s);
- the location of the servers;
- the legal jurisdiction of the contract between the parties;
- any legal frameworks between those various locations.

Therefore, applicable legal requirements will vary a lot depending on the various jurisdictions, as well as legal entities and frameworks involved. This said, our caveat here is that EU law applies (and when relevant, EU Member States' specific laws).

The following Section describes these legal instruments, which encompass both hard law legislation and soft law documents. In addition, the Section covers the question of how some of these instruments have been transposed into national law, to the extent relevant.

Legislation at EU level

The Directive on security of network and information systems (NIS Directive): The first piece of EU-wide legislation dealing with cyber security applicable to digital services[110] providers is Directive (EU) 2016/1148 of the European Parliament and the Council of 6 July 2016 (NIS Directive). The Directive aims at ensuring a level playing field across Member States which guarantees a high common level of security of network and information systems across the EU in the context of the Digital Single market. To this end, Member States must put in place the following measures to increase cybersecurity:

- **National Capabilities**: Member States preparedness against incidents and security breaches must be ensured by requiring them to be appropriately equipped, e.g. via the creation of a Computer Security Incident Response Team (CSIRT) in order to contribute to the development of trust and confidence between Member States and to promote swift and effective operational cooperation. The Directive also foresees the obligation for Member States to designate a competent national authority. In addition, cyber exercises should be carried out.
- **Cooperation among Member States:** To this end, a cooperation group (composed of representatives of the Member States, the Commission and ENISA, the European Union Agency for Cybersecurity[111]) must be set up in order to support and facilitate strategic cooperation and the exchange of information among Member States. This is the avenue where Member States cooperate, exchange information and agree on the consistent implementation of the Directive. The Directive also provides that a CSIRT Network (composed of representatives of the Member States' CSIRTs and CERT-EU) must be set up in order to promote swift and effective operational cooperation on specific cybersecurity incidents and sharing information about risks;

[109] *Id*.

[110] A 'digital service' is defined by the Directive (EU) 2015/1535 as "*any service normally provided for remuneration, at a distance, by electronic means and at the individual request of a recipient of services*" . For the purposes of this Study, cloud services fall within digital services under Directive 2016/1148.

[111] ENISA provides recommendations on cybersecurity and supports policy making. Its mandate, first foreseen in this piece of legislation, has been further strengthened in 2019 through the Cybersecurity Act.

- **National requirements in terms of cybersecurity of various sectors in EU Member States**: A culture of security across sectors which are vital for our economy and society and moreover rely heavily on ICTs, such as energy, transport, water, banking, financial market infrastructures, healthcare and digital infrastructure is fostered by the Directive. Businesses in these sectors which are identified as operators of essential services (OES) by the Member States will have to take appropriate security measures, as well as notify serious incidents to the relevant national authority. Digital service providers (OES) will have to take measures to prevent and minimise the impact of incidents affecting the security of their network and information systems.

Regulation 2019/881 on ENISA and ICT Cybersecurity Certification (Cybersecurity Act)[112]: Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act) represents the regulatory framework laying down ENISA's tasks and prerogatives. It also contains provisions on information and communications technology (ICT) cyber security certification. In particular, the Cybersecurity Act foresees a permanent mandate for ENISA, reinforcing its role in cybersecurity with the conferral of new tasks and resources to carry them out, and the creation of an EU certification framework for ICT products and services. The Regulation acknowledges that the digital single market, and in particular the data economy and the IoT, can thrive only if there is general public trust that such products, services and processes provide a certain level of cybersecurity. Therefore, the Regulation lays down a framework for the establishment of European cybersecurity certification schemes for the purpose of ensuring an adequate level of cybersecurity for ICT products, ICT services and ICT processes in the Union, as well as for the purpose of avoiding the fragmentation of the internal market with regard to cybersecurity certification schemes in the Union. This framework shall provide for a mechanism to establish European cybersecurity certification schemes and to attest that the ICT products, ICT services and ICT processes that have been evaluated in accordance with such schemes comply with specified security requirements for the purpose of protecting the availability, authenticity, integrity or confidentiality of stored or transmitted or processed data or the functions or services offered by, or accessible via, those products, services and processes throughout their life cycle. In accordance with the Act, ENISA shall promote the use of European cybersecurity certification: the mission of ENISA in this respect is to engage with public services and with industry and standardisation organisations to draw up candidate cybersecurity certification schemes. The European cybersecurity certification has the objective to assure a high level of protection of stored, transmitted or otherwise processed data against accidental or unauthorised storage, processing, access, destruction, loss, alteration, lack of availability or disclosure during the entire life cycle of the ICT product, ICT service or ICT process. It is also designed to achieve a good knowledge of their dependencies and vulnerabilities.

The **Regulation on the Free flow of Non-Personal Data**, already referred to in Chapter 2, plays an important role also in the cybersecurity domain: certification schemes and codes of conduct pursuant to the Regulation on the Free Flow of Non-Personal data are worth mentioning.

Preamble 33 of the Regulation on the free flow of non-personal data provides that enhancing trust in the security of cross-border data processing should reduce the propensity of market players and the public sector to consider data localization as a proxy for data security. When it comes to national law, Preamble 35 provides that security requirements at national level should be **necessary** and **proportionate** to the risks

[112] Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013, [2019], OJ L 151.

posed to the security of data processing. Preamble 36 also makes explicit reference to the abovementioned NIS Directive. While it is left to Member States to ensure that digital services providers whom the Regulation applies to identify and take appropriate and proportionate technical and organizational measures to manage the risks posed to the security of network and information systems, a minimum level of security pursuant to the Regulation should encompass the following: incident handling, business continuity management, monitoring, auditing and testing, and compliance with international standards. This is further specified under Art. 6 of such regulation, enshrining upholding data portability. To this end, under Art. 6(1), the Commission shall facilitate the development of self-regulatory codes at EU level, which - *inter alia* - taking into account the principles of transparency and interoperability and due account of open standards. It shall also ensure approaches to certification schemes that facilitate the comparison of data processing products and services, including information security management.

The **European Electronic Communications Code (EECC)**[113] is a comprehensive set of updated rules for the telecoms sector and part of a package of telecom laws, which includes Regulation (EU) 2018/1971 establishing the Body of European Regulators for Electronic Communications (BEREC) and the Agency for Support for BEREC (BEREC Office)[114]. The EEC updates and consolidates the framework for the regulation of electronic communication services across the EU and EEA. It replaces and repeals Directives 2002/19/EC[115], 2002/20/EC[116] and 2002/21/EC[117], as well as Article 5 of Decision 243/2012/EU[118]. It entered into force on 20 December 2018 and must be transposed by the Member States by 21 December 2020. Regardless of this transposition deadline, certain provisions of the EE may have direct effect across Member States.

Its primary aim is to stimulate competition and increase investment in 5G and very high capacity networks in order to give access to high quality connectivity to every citizen and business in the EU. It aims as well at ensuring competition by extending rules to providers that were not regulated by the previous framework, such as over-the-top (OTT) players which offer telecoms services such as interpersonal communication, content and cloud services.

On the one hand, it sets out regulatory tasks for the NRAs and other competent authorities intended to achieve the objectives laid down in Art. 3(2) and (4)[119] and establishes a set of procedures to ensure that the

[113] Directive (EU) 2018/1972 of the European Parliament and of the Council of 11 December 2018 establishing the European Electronic Communications Code, OJ L 321.

[114] Regulation (EU) 2018/1971 of the European Parliament and of the Council of 11 December 2018 establishing the Body of European Regulators for Electronic Communications (BEREC) and the Agency for Support for BEREC (BEREC Office), amending Regulation (EU) 2015/2120 and repealing Regulation (EC) No 1211/2009 [2018] OJ L 321/1.

[115] Directive 2002/19/EC of the European Parliament and of the Council of 7 March 2002 on access to, and interconnection of, electronic communications networks and associated facilities (Access Directive) [2002] OJ L 108/7.

[116] Directive 2002/20/EC of the European Parliament and of the Council of 7 March 2002 on the authorisation of electronic communications networks and services (Authorisation Directive) [2002] OJ L 108/33.

[117] Directive 2002/21/EC of the European Parliament and of the Council of 7 March 2002 on a common regulatory framework for electronic communications networks and services (Framework Directive) [2002] OJ L 108/51.

[118] Decision No 243/2012/EU of the European Parliament and of the Council of 14 March 2012 establishing a multiannual radio spectrum policy programme [2012] OJ L 81/7.

[119] Art. 3 (2) reads: "In the context of this Directive, the national regulatory and other competent authorities as well as BEREC, the Commission and the Member States shall pursue each of the following general objectives, which are not listed in order of priority:

(a) promote connectivity and access to, and take-up of, very high capacity networks, including fixed, mobile and wireless networks, by all citizens and businesses of the Union;

(b) promote competition in the provision of electronic communications networks and associated facilities, including efficient infrastructure-based competition, and in the provision of electronic communications services and associated services;

regulatory framework is harmonised throughout the EU. On the other hand, it places obligations upon 'electronic communications services providers' that will have access to 5G radio spectrum licencing and will have to comply with specific rules, among which, competition and consumer protection provisions and well as end-user protection rules and security requirements.

This Directive redefines and expands the scope of the 'electronic communications service' definition. Pursuant to Art. 2(4) of the EEEC there are three types of electronic communications services: (a) Internet access service; (b) interpersonal communications services (c) services consisting mainly or wholly in the conveyance of signal, such as transmission services used for the provision of machine-to-machine services and for broadcasting. DUET's activities may consist in the provision of machine-to-machine services by way of deploying IoT systems, such as people/car counting sensors, which may put DUET in scope of this legislation and, thus, be supervised by national authorities.

With specific regard to security issues, the EECC introduces significant changes in the security requirements and supervision of the electronic communications services, compared to its predecessors. As ENISA reports, the "*new rules provide an EU-wide definition of security requirements and security incidents for the telecom sector*"[120]. Art. 2(21) defines the umbrella term 'security of networks and services' as "*the ability of electronic communications networks and services to resist, at a given level of confidence, any action that compromises the availability, authenticity, integrity or confidentiality of those networks and services, of stored or transmitted or processed data, or of the related services offered by, or accessible via, those electronic communications networks or services".* Hence, "breaches of confidentiality of communications, or issues with the authentication of users, for example, are in scope. Previously this was left up to interpretation. At present, NRAs have been given a clear mandate to address them"[121].

The new rules require the electronic communication service providers to implement state-of-the-art measures in order to protect the security of networks, services and users' communication. Art. 40(1) requires Member States to ensure that *"providers of public electronic communications networks or of publicly available electronic communications services take appropriate and proportionate technical and organisational measures to appropriately manage the risks posed to the security of networks and services*". Recital 94 further clarifies that "*Having regard to the state of the art, those measures shall ensure a level of security appropriate to the risk presented".*[122]

(c) contribute to the development of the internal market by removing remaining obstacles to, and facilitating convergent conditions for, investment in, and the provision of, electronic communications networks, electronic communications services, associated facilities and associated services, throughout the Union, by developing common rules and predictable regulatory approaches, by favouring the effective, efficient and coordinated use of radio spectrum, open innovation, the establishment and development of trans-European networks, the provision, availability and interoperability of pan-European services, and end-to-end connectivity;

(d) promote the interests of the citizens of the Union, by ensuring connectivity and the widespread availability and take-up of very high capacity networks, including fixed, mobile and wireless networks, and of electronic communications services, by enabling maximum benefits in terms of choice, price and quality on the basis of effective competition, by maintaining the security of networks and services, by ensuring a high and common level of protection for end-users through the necessary sector-specific rules and by addressing the needs, such as affordable prices, of specific social groups, in particular end-users with disabilities, elderly end-users and end-users with special social needs, and choice and equivalent access for end-users with disabilities."

[120] ENISA, Press Release, "Security supervision changes in the new EU telecoms legislation", January 2020.
[121] ENISA, id., cited.
[122] According to this recital, security measures should take into account as a minimum, the following: "as regards security of networks and facilities: physical and environmental security, security of supply, access control to networks

According to Recital 94, *"measures, including encryption where appropriate, shall be taken to prevent and minimise the impact of security incidents on users and on other networks and services"* . Recital 97 makes express reference to encryption, as follows: "*(end-to-end) encryption and, where necessary, encryption should be mandatory in accordance with the principles of security and privacy by default and by design".*

Moreover, services providers are required to inform - free of charge - their customers about possible security threats and measures they can take to protect the security of their communications, for instance by using specific types of software or encryption technologies (Article 40(3)[123]). However, such "*requirement to inform users should not discharge a service provider from the obligation to take, at its own expense, appropriate and immediate measures to remedy any security threats and restore the normal security level of the service*" (Recital 96).

Additionally, the national telecom authorities can ask telecom providers to mitigate specific cyber threats, even prior to there being actual incidents (Art. 41(1)). Providers of public electronic communications networks or services must notify immediately the national authority of a significantly impactful security incident. Art. 40(2) clarifies the parameters to be considered in the security breach reporting in order to assess the significance of breaches. Such parameters are: the number of users affected by the security incident, the duration thereof, the geographical scale, the impact on the functioning of the network or service and the extent of impact on economic and societal activities.

Soft law at EU and international level

Several soft instruments adopted by ENISA (the European Union Agency for Cybersecurity) are relevant in this respect. This Section highlights some of them and the recommendations that can help best practices in the context of security related aspects of the DUET Digital Twins various phases.

This Section first tackles a background analysis of ENISA's mandate. It then focuses on some guidance ENISA has provided throughout its mandate, so as to sketch the legal landscape on cybersecurity soft law at EU level.

ENISA's mandate, tasks and prerogatives are regulated under the abovementioned Cybersecurity Act (Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013. ENISA was originally conferred a mandate under the Regulation (EC) No 460/2004 of the European Parliament and of the Council of 10 March 2004 establishing the European Network and Information Security Agency. It supports the European Union institutions, the Member States and the business community in addressing, responding and especially in preventing network and information security problems. It does so through a series of activities across five areas identified in its strategy: Expertise; Policy; Capacity; Community; Enabling.

Mandate: ENISA's mandate became permanent. The new mandate further clarifies the role of ENISA as the EU agency for cybersecurity and as the reference point in the EU cybersecurity ecosystem, acting in close

and integrity of networks; as regards handling of security incidents: handling procedures, security incident detection capability, security incident reporting and communication; as regards business continuity management: service continuity strategy and contingency plans, disaster recovery capabilities; as regards monitoring, auditing and testing: monitoring and logging policies, exercise contingency plans, network and service testing, security assessments and compliance monitoring; and compliance with international standards". Furthermore,

[123] See also Recital 96 and 97.

cooperation with all the other relevant bodies of such an ecosystem, including private and public actors. The scope of the mandate is better delineated, strengthening those areas where the agency has shown clear added value and adding those new areas where support is needed in view of the new policy priorities and instruments, in particular the NIS Directive, the review of the EU Cybersecurity Strategy, the EU Cybersecurity Blueprint for cyber crisis cooperation and ICT security certification.

Governance: The organisation and the governance of the Agency were moderately reviewed, in particular to make sure that the needs of the wider stakeholders' community are better reflected in its work.

Tasks: EU policy development and implementation tasks: ENISA is tasked with proactively contributing to the development of policy in the area of network information security, as well as to other policy initiatives with cybersecurity elements in different sectors (e.g. energy, transport, finance). To this end, it has a strong advisory role: in particular, it provides independent opinions and preparatory work for the development and the update of policy and law. ENISA also supports EU policy and law in the areas of electronic communications, electronic identity and trust services, with a view to promoting an enhanced level of cybersecurity. This task is especially relevant in the context of the current Study. ENISA assists EU Member States in achieving a consistent approach on the implementation of the NIS Directive across borders and sectors, as well as in other relevant policies and laws. It does so through its input in the context of the abovementioned NIS Cooperation Group. In order to support the regular review of policies and laws in the area of cybersecurity, ENISA also provides regular reporting on the state of implementation of the EU legal framework.

Capacity building tasks: ENISA contributes to the improvement of EU and national public authorities' capabilities and expertise, including on incident response and on the supervision of cybersecurity related regulatory measures. The Agency is also required to contribute to the establishment of Information Sharing and Analysis Centres in various sectors through providing best practices and guidance on available tools and procedures, as well as by appropriately addressing regulatory issues related to information sharing.

Knowledge and information, awareness raising tasks: ENISA became the information hub of the EU in matters relating to cybersecurity, including by means of the promotion and sharing of best practices and initiatives across the EU. To do so, ENISA pools information on cybersecurity deriving from the EU and national institutions, agencies and bodies. The Agency also makes available advice, guidance and best practices on the security of critical infrastructures. In the aftermath of significant cross-border cybersecurity incidents, ENISA furthermore compiles reports with a view of providing guidance to businesses and citizens across the EU. This stream of work also involves the regular organisation of awareness raising activities in coordination with Member States' authorities.

Market related tasks (standardisation, cybersecurity certification): ENISA performs a number of functions specifically supporting the internal market, including through a cybersecurity 'market observatory', by analysing relevant trends in the cybersecurity market to better match demand and supply, and by supporting the EU policy development in the ICT standardisation and ICT cybersecurity certification areas. With regard to standardisation in particular, it facilitates the establishment and the uptake of cybersecurity standards. ENISA also executes the tasks foreseen in the context of the future framework for certification (see infra).
Research and innovation: ENISA contributes its expertise by advising EU and national authorities on priority-setting in research and development, including in the context of the contractual public-private partnership on cybersecurity (cPPP). ENISA's advice on research feeds into the new European Cybersecurity Research and Competence Centre under the next multiannual financial framework. ENISA is also involved,

when asked to do so by the Commission, in the implementation of research and innovation EU funding programmes.

Operational cooperation and crisis management: this stream of work strengthens the existing preventive operational capabilities, in particular upgrading the pan-European cybersecurity exercises (Cyber Europe) by having them on a yearly basis, and on a supporting role in operational cooperation as secretariat of the CSIRTs Network (as per NIS Directive provisions) by ensuring, among others, the well-functioning of the CSIRTs Network IT infrastructure and communication channels. In this context, a structured cooperation with CERT-EU, European Cybercrime Centre (EC3) and other relevant EU bodies is required. Furthermore, a structured cooperation with CERT-EU, in close physical proximity, results in a function to provide technical assistance in case of significant incidents and to support incident analysis. Member States that would request it would receive assistance to handle incidents and support for the analysis of vulnerabilities, artefacts and incidents in order to strengthen their own preventive and response capability.

Tasks on incident management: ENISA also plays a role in the EU cybersecurity blueprint presented as part of this package and setting the Commission's recommendation to Member States for a coordinated response to large-scale cross-border cybersecurity incidents and crises at the EU level. ENISA facilitates the cooperation between individual Member States in dealing with emergency response by analysing and aggregating national situational reports based on information made available to the Agency on a voluntary basis by Member States and other entities.

The Study team did not identify any specific ENISA guidance tackling the security aspects of smart cities. Yet, among the many pieces of soft law that ENISA has produced throughout its mandate, three of them become particularly relevant: its Guidance on IoT, mapping of risks and risk management with respect to operators of essential facilities (which can further help us sketch the risk management and risk mitigation measures below), as well as how digital services providers should handle incident notifications. Below we provide a brief overview of these three pieces of soft law:

Good practices for security of IoT in the context of smart manufacturing: this guidance lays down good practices for security of IoT related to Industry 4.0. Given the smart city and Industry 4.0 intersection, such guidance also becomes relevant in the context of the current project. Indeed, in the concept of Industry 4.0, the IoT shall be used for the development of so-called smart products[124]. As Siemens highlights, Industry 4.0 is a smart city enabler, insofar as Industry 4.0 solutions are enabling smart cities insights determine infrastructure requirements for logistics and interconnected network needs and demands.

Among others the guidance helps in outlining, besides the privacy aspects, the security challenges related to the evolution of industrial systems and services precipitated by the introduction of IoT innovations, with the objective to collect "*good practices to ensure security of IoT in the context of Industry 4.0/Smart Manufacturing, while mapping the relevant security and privacy challenges, threats, risks and attack scenarios*". First, the study identifies the security challenges and vulnerabilities of industry 4.0 and its components, and it later on carries out a taxonomic analysis of risks and threats, by providing concrete examples. On the basis of this risk assessment, it provides a list of security measures containing the policies and procedures to mitigate risks and address the threats identified. In particular, it focuses on the notion of **security by design**, and it clarifies that such notion is linked to the security measures to be applied from the very beginning of product development.

[124] Lom, M., "Industry 4.0 as a part of smart cities", IEEE Conference, Smart City Symposium Prague (2016).

Baseline Security Recommendations for Internet of Things in the context of critical information infrastructures:

Cloud computing soft legal framework at EU law: Aside from the abovementioned NIS Directive, as well as the Regulation on the free flow of non-personal data, in May 2019, the European Commission adopted a new **Cloud Strategy**. This follows and bolsters an initiative adopted in 2014 concerning piloting of a potential of a cloud of public services for the delivery of more flexible public services and providing for service sharing between public and private providers.

The European Commission's cloud strategy - *inter alia* - elaborates on how the cloud is affecting digital solutions, the data ecosystem, IT infrastructure and security services. This strategy aims at a cloud service offering that is secure: such objective must be achieved by identifying and managing security risks (besides from being multi-cloud (so as to avoid dependence on a single provider), energy-efficient, hybrid, and privacy compliant). In the strategy, the Commission observes that it is facilitating self-regulatory work from industry to develop recommendations for the purposes of a European Cloud Certification Scheme. The ENISA may be asked to draw up a candidate scheme in accordance with the abovementioned Cybersecurity Act, which will address both personal and non-personal data. It also observes that at the moment, several codes of conduct have been established by the industry, such as the Cloud Select Industry Group, the Cloud Infrastructure Service Providers in Europe, as well as the Cloud Security Alliance. In particular, the NGO Cloud Security Alliance has adopted guidelines tackling Cyber Security for Smart Cities technologies. Such private sector initiative will be referred to below.

Moreover, it is also worth specifying that ENISA has also adopted some guidance on the cloud. A 2009 Cloud Security Risk Assessment, often referred to as best practice, was followed by several other initiatives including- inter alia - an assurance framework for governing IT risks in the cloud. In addition, throughout its mandate ENISA has adopted several documents providing best practices, including guidelines concerning incident reporting for cloud computing (2013), cloud security guidance for SMEs (2015), as well as a short 2016 paper on cloud forensics. Whenever relevant, they will be mentioned *infra.*

Finally, security standards at international level constitute an important tool to guarantee compliance with legal requirements on security by big data services providers. According to the General Vocabulary of the ISO/IEC Guide 2:2004[125]: "*A Standard is a document, established by consensus and approved by a recognized body, which provides for common and repeated use, rules, guidelines or characteristics for activities or their results, aimed at the achievement of the optimum degree of order in a given context.*" The key organization in the world developing international cybersecurity standards is the International Organization for Standardization and International Electrotechnical Commission Joint Technical Committee 1, Information Technology (ISO/IEC JTC1), in particular Subcommittee 27: "IT Security Techniques". The latter is an internationally recognized centre of information and IT security standards expertise serving the needs of both the business sectors and the governments. Its work covers the development of standards for the protection of information and ICT, spanning over requirements, methods, techniques and guidelines to address aspects of both security and privacy in regard to information security management systems (Member States), cryptographic and security mechanisms, security evaluation, testing and specification, security controls and services, identity management and privacy technologies. Reliance on such standards

[125]ISO/IEC GUIDE 2:2004 Standardization and related activities — General vocabulary, available at: https://www.iso.org/standard/39976.html .

ensures that big data services providers implement the state of the art security-related measures and processes.

3.3 Risk considerations

As Lilian Edwards cautions, cities and their infrastructure, when interweaving with smart cities solutions, rely on sensor networks and integrated communications systems, making them vulnerable to a series of security issues, for example security incidents, cyber-attacks, power failure and software errors[126]. Risks at various levels need to be identified and mapped. In particular, the IoT technology which the DUET Digital Twin would - *inter alia* - rely on, is prone to various vulnerabilities.

Beyond the risks associated with physical components such as sensors, and hardware components, other risks concern the integrity of data itself. In this respect, both the GDPR, as well the ePrivacy Directive become relevant.

Finally, security risks could encompass vulnerability of software applications and components, as well as the cloud infrastructure (infrastructure as a service). To complicate matters further, digital technologies such as IoT display complexity levels due to the interdependency between the different components and layers, encompassing the tangible devices, but also the software interfaces and components, the data level, both at rest, and at transmission and also the features of connectivity[127].

Two levels of risks are particularly worth mentioning: (a) IoT components. A 2015 FTC report shows, for example, that a company making baby monitors attached to the Internet, thus allowing parents to view live feeds of their infants from a distance, was hacked in hundreds of cases[128]. Such vulnerabilities encompass unauthorized access as well as misuse of personal information. Another risk is an attack on device functionality. (b) Cloud: various levels of cloud and associated risks.

This riks mapping needs to be done during the development phase of the DUET project. For the time being, we provide below a brief overview of the main concerns emerging in the literature as well as in best practices from the business sector, and in particular the Cloud Security Alliance[129].

Vulnerabilities arise in the various stages of the process, from the design and planning stage until the implementation phase and finally the operational and maintenance phase.

In a nutshell, in the design and planning stage, the vulnerabilities are several: first concerns with data, such as insufficient cryptographic protection (e.g. encryption), on rest and on transit. Loss of encryption keys by the data controller can enable a malitentioned party to tamper with data. Other security concerns relate to who can access the data (authentication) and what permissions does he or she have (authorisation). Vulnerability of authentication capabilities is related to weak passwords, etc. When it comes to lack of authorisation, this relates to what the user can do: when permissions are more than needed, this increases vulnerability risks. For example, a concern is the vulnerability of secure configuration by default. Moreover, common security problems are issues with updates of software as well as tampering by unauthorized

[126] Edwards, cited.
[127] idem
[128] FTC Staff Report, Privacy and Security in an interconnected world, 2015.
[129] Cloud Security Alliance, cited.

sources. Other vulnerabilities relate to security incidents which concern instances of threats that cause the system or data to be compromised. They occur due to events, such as phishing. When non-basic functionalities are not effectively secured, this can also give rise to vulnerabilities. System malfunctions and crashes can concern the various layers, such as software and hardware, being insecure premises part of some hardware vulnerabilities. There are also risks when it comes to an insufficient baseline for auditing capabilities. Finally, there are concerns with vendor security which also impact the DUET cloud security.

Further vulnerabilities arise at the levels of technology implementation phase and the operation and maintenance phase. They concern, for example, the following risks:

- Vulnerabilities with security tests
- Lack of strong encryption
- Issues with administration interfaces and functionality
- Vulnerabilities of system administration
- Insufficient audit of security risks
- Unauthorised physical access
- Vulnerability of passwords for access to systems
- Vulnerability of user functionality and services

At the operationation and maintenance level, the following concerns can be identified:

- Vulnerabilities with monitoring
- Vulnerabilities with patching
- Insufficient assessment and auditing
- Concerns with the logging environment
- Risks with unsecure access control
- Absence of or insufficient threats intelligence
- Absence of recovery plans in case of incidents

3.4 Requirements plan for risk mitigation

First, as ENISA opines, it is important to "*treat IoT cybersecurity as a cycle, not as an end-to-end process, adopting a security by design approach from the viewpoint of the devices and infrastructure at every step of the development lifecycle*". It is important to address cybersecurity through embedded features of endpoints rather than only at the network level. In particular, in the cloud environment protecting endpoints from vulnerabilities and threats is crucial. It is all the more so important to carry out a security and safety assessment and embed even the most basic connected devices holding very limited processing capabilities (e.g. actuators, converters) with identification and authentication features as well as ensure compatibility with IAM class solutions.

In this respect, risk and threat analysis must be performed involving cybersecurity experts from the very early stages of the design process of DUET. Hence, for each design document, a chapter addressing the security of all the information and control systems must be included.

We hereby provide certain best practices in terms of security by design. Some of them are also foreseen under the EEEC, as analysed above. In the design and planning stage, the following should be guaranteed:

- **Strong cryptography of data at rest and at transit:** these encompass strong encryption, secure encryption keys, etc.
- **Enhance authentication capabilities:** these concern one time passwords, two-level authentication or biometrics use for authentication.
- **Ensuring authorisation capabilities' management:** This entails ensuring that the authorised parties have the adequate permissions to use data.
- **Secure updates of software:** This is necessary to do away with the abovementioned vulnerabilities**;**
- **Auditing and logging capabilities:** these require, for example, audit plans baseline and management;
- **Anti-tampering capabilities:** these entail ensuring procedures to avoid that data and systems are tampered with;
- **Accounts:** constant change of passwords, no backdoors, should be guaranteed;
- **Non-basic functionality should be disabled by default;**
- **Security and disaster management procedures** should be in place and there should also be
- **Secure configuration by default.**

Testing procedures should be foreseen. They concern plans on security requirements' compliance. System hardening. The role of certification schemes and the validation of security processes via certification is important as well. Similar protections should be ensured in the technology implementation phase and the operation and maintenance phase, this latter requiring monitoring, patching and regular auditing, as well as continuous assessment of intelligence against threats, classifying risks and having in place procedures which foresee recovery planning in case of compromised systems.

3.5 Cybersecurity risk management and controls: criteria to draw up sound risk management plans

Once risks have been identified, adequate controls must be put in place to adequately mitigate these risks. Industry-wide standards can help in this respect. Some of the areas concerning the risk mitigation that have been identified in the context of smart cities concern the following aspects:

- Audit and accountability: sound audit plans should be put in place.
- Awareness and training: risks should be continuously assessed and awareness raising should occur on aspects concerning security, including by means of periodical training.
- Business continuity management and operational resilience: the plan should encompass resilience procedures faced with operational risks. In addition, the plan should also foresee procedures in terms of incident responses that ensure secure continuity.
- Change control and configuration management.
- Data security and integrity: first, it is important to foresee the principle of least privilege, allowing only authorized accesses which are necessary to carry out tasks assigned. Encryption and Key Management Storage and Access should be secured. Platform and data-appropriate encryption should be required. In this context, keys shall not be stored in the cloud but maintained by trusted and secure key information providers.
- Information lifecycle management

➢ Identity and access management: The organization separates duties and ensures that there are authorisations to support the separation of duties. In addition, user access policies and procedures shall be established, which shall encompass these minimum requirements.

4. Ethics

4.1 Purpose: Ethics-related considerations on Big data, IoT and AI beyond privacy and (cyber)security

The purpose of this chapter is to identify and briefly present the ethics-related aspects of the technologies that DUET will operate on (machine learning and big data, IoT and Artificial intelligence (AI), and their intersection), beyond the privacy and data protection aspects and the cybersecurity-related aspects dealt with in the previous Chapters.

Throughout jurisdictions, there is now an increasing awareness that a responsible approach to AI and IoT, is needed to ensure the safe, beneficial and fair use of such technologies. This entails understanding the implications of moral decision-making by machines[130], in terms of accountability, responsibility and liability. Accountability relates to the "*need to explain and justify one's decisions and actions to its partners, users and others with whom the system interacts*". To ensure accountability, decisions must be derivable from, and explained by, the decision-making algorithms used[131]. Linked to accountability, responsibility refers "t*o the role of people themselves and to the capability of AI systems to answer for one's decision and identify errors or unexpected results*"[132]. Finally, liability relates to who bears legal liability when things go wrong. AI/IoT technologies raise the question of who is accountable for what, who is liable for what, and who is responsible for what. DUET's involvement in the project touches on all those aspects. The engineers must responsibly build a safe architecture through the involvement of the best data scientists (through design of safe systems) and DUET's accountability relates to the need to ensure that decisions are explained to the stakeholders involved in this ecosystem, while liability relates to tortious or contractual exposure of DUET towards potential victims in the event of accidents may occur.

In this respect, it is worth mentioning that two main risks for the project can be envisaged ethics-wise:
First, the law dealing with AI and IoT at EU level is not yet developed in hard law pieces of legislation, harmonized across the EU-27. Currently there is an array of soft law instruments (including at national level): While these instruments provide for some guidance - that will be described in this Chapter - at macro-level,

[130]Kavathatzopoulos, I., and Asai, R., Can Machines make ethical decisions, 2013.
[131] Dignum, V., López-Sánchez, M., Micalizio, R., Pavón, J., Slavkovik, M., Smakman, M., Steenbergen, M., Tedeschi, S., Toree, L., Villata, S., Wildt, T., Baldoni, M., Baroglio, C., Caon, M., Chatila, R., Dennis, L., Génova, G., Haim, G., Kließ, M., "Ethics by Design: Necessity or Curse?", AIES'18, February 2–3, (2018), New Orleans, LA, USA.
[132] Id. Also see Council of Europe Study, Responsibility and AI, 2019.

by no means do they exhaustively address all the ethics-related questions other than data privacy and security that may arise in the context of this project. Those questions deal with ethical issues of AI/IoT, including the product safety, contractual and extra-contractual liability issues which those technologies encompass, at the various layers where DUET will operate.
As there are no dedicated pieces of legislation yet tackling new technologies such as IoT at EU level, and as the Digital Data legislation announced by the European Commission in its Communication "A European Strategy for Data"[133] last February 2020, is still work in progress, it will be necessary to monitor EU-wide developments as the DUET project unfolds. It will also be necessary to acknowledge that soft law does not have answers for all the questions that may arise. Some of those questions will need to be addressed by reference to laws which were enacted prior to the technology having developed as much as it has nowadays.

The second risk is related to the fact that, given EU-wide legislation is yet unaccomplished, there may be a need to delve into contractual and non-contractual aspects of national legislation that may require, as the case may be, teams of national lawyers acquainted with the laws of national procedure, including civil procedure. While Grimaldi Studio Legale has a network of lawyers across the 27 Member States + UK whose expertise it can count on in short notice, a risk mitigation procedure must be put in place were DUET to be called upon in the context of national specific legislation already at the outset of the project. Given that well-established principles of civil liability law are still national, it is acknowledged that DUET may as well rely on the expertise of national lawyers, as the case may be, as going forward with the project, jurisdiction-specific questions may arise.

Against this background, it will be useful to think holistically about the ethics related questions. To do so, we will wear the same lens as that used in the previous Chapters tackling the data privacy and security aspects through the notion of "ethics-by-design". Ethics-by-design is defined by the European Commission as "the implementation, starting from the beginning of the design process, of ethical and legal principles"[134].

This Chapter will be structured as follows: Under Section 4.2, we will address some risk aspects cutting across the various technologies, highlighting – at macro level - concerns and risk mitigations for each of them. For better illustration, we will divide them into three separate main topic areas: Sub-Section 4.2.1 will deal with aspects concerning devices' safety and liability and how those aspects may raise issues for DUET. In particular, we will highlight the gaps under the existing EU legal framework, as well as briefly touch upon the complexity of national liability regimes. Under Sub-Section 4.2.2., and Section 4.2.3, we will delve into the potential liability for DUET of breaching third party rights (IP or trade secrets, under 4.2.3) or data subject rights when the data shared by DUET with its government partners, and eventually published, may accidentally cause harm, and entail the potential exposure of DUET in this respect (under 4.2.3).

Sub-Section 4.3 will delve into what the AI and IoT specific legal landscape is, highlighting that it - for the time being - consists of almost exclusively soft-law and policy papers. After an overview of the soft law instruments, we will subsequently map out, for each of those technologies, the areas where risks were identified, and will provide an overview of possible risk mitigations.
These high-level considerations will be further discussed and developed in the context of a separate ethics deliverable.

[133] See *supra*.
[134] European Commission, Communication to the European Parliament, the European Council, the Council, the European Economic and Social Committee, Coordinated Plan on Artificial Intelligence (COM(2018) 795 final).

4.2. Ethical risks for DUET beyond data privacy and data security: liability-related aspects

This Section aims at outlining various aspects of legal responsibility that DUET may be exposed to, or may assume at the various stages of the initiative implementation, based on the information available to the team at the time of the drafting. The revolving theme is the heavy reliance of systems and solutions on (big) data, devices made by third party manufacturers that may in part or wholly rely on data and their transmission over networks (IoT) and deployment of artificial intelligence and robots. Particular concerns arise with respect to liability for devices and services using data, accurateness of data, or classification of data as information or as property (IP rights and trade secrets implications). All these aspects challenge the traditional conceptions of attributing legal liability to various actors.

Subsection 4.2.1 explains the difference between rules aiming at prevention of harm (product safety rules) and rules that ensure that harm that has nevertheless occurred gets compensated (liability rules). The European Commission has recently looked at how the existing EU legislation on product safety and liability may (or may not) address issues specific to IoT, big data and AI. The subsection provides a high level overview of EU law framework and in particular highlights the need to focus down relevant national level liability regimes for cases where EU law does not apply, such as in cases of extra-contractual liability for tort, fraud, or misrepresentation.

Subsection 4.2.2 elaborates on aspects of legal liability where third parties are involved, such as parties claiming infringement of their IP rights and the overarching problems involved in data and databases ownership concepts. As DUET may potentially get exposed to IP infringement claims, a high level overview of IP enforcement framework is provided.

Subsection 4.2.3 flashes out liability concerns that may arise in connection with making data available to the general public or sharing it with businesses or governments (both open / public sector data, or private sector data). Such data may be misused, or their disclosure may harm third parties by divulging personal data, trade secrets or other commercially sensitive information. The subsection summarizes rules and good practices for responsible data sharing and publication by various actors, and considers what risk mitigation steps could be undertaken by DUET or its partners to limit any liability.

4.2.1 Sketching the issues: Safety and liability aspects

DUET may need to deal with legal concerns related to safety and liability of the systems, devices and services deployed in order to implement various aspects of Smart Cities. While certain safety and liability concerns can be addressed *ex ante* in the design and planning phases, some legal issues may arise in the course of operation of Smart Cities and may need to be dealt with on a case by case basis. The following subsection provides a high level overview of these aspects.

Safety vs. liability

While (product) safety rules aim at avoiding accidents from happening, rules on (civil) liability intervene in case accidents happen. Liability rules play a double role, on the one hand, they ensure victims of a damage caused by others get compensation, and, on the other hand, they provide economic incentives for the liable party to avoid causing such damage and thus also serve prevention.

Safety: Legal landscape

As stated in the **Commission Report on the safety and liability implications of AI, IoT and robotics**, "*safety in the current Union product safety legislation is a public policy objective*"[135]. The safety concept is linked to the use of the product and all kinds of risks, including mechanical, chemical, electrical but also cyber risks and risks related to the loss of connectivity of devices. The use of the product can encompass both the intended, the foreseeable and the reasonably foreseeable use.

The Union product safety framework already sets obligations for producers to take into account in the risk assessment the said use of the products throughout their lifetime. It also foresees that manufacturers must provide for instructions and safety information for users or warnings (Decision No 768/2008/EC of the European Parliament and of the Council of 9 July 2008 on a common framework for the marketing of products, and repealing Council Decision 93/465/EEC, OJ L 218, 13.8.2008. p. 82–128. Annex I, Art. R2.7 reads: "*Manufacturers shall ensure that the product is accompanied by instructions and safety information in a language which can be easily understood by consumers and other end-users, as determined by the Member State concerned*.").

As regards prevention, IoT products and systems, including their AI/big data elements, are currently subject to the existing general product safety legislation at the EU and national levels. The EU framework includes horizontally applicable provisions of the **General Product Safety Directive (Directive 2001/95/EC)[136]**, and several sectoral directives, including for Smart Cities potentially pertinent Radio Equipment Directive (see *supra* Section 3.2) or vehicle-type approval legislation. These instruments would have national-level equivalents implementing the specific rules.

The EU product safety legislation does not generally provide for specific mandatory essential requirements against cyber threats affecting the safety of users. However, there are provisions related to security aspects in the Regulation on Medical Devices[137], the Directive on measuring instruments[138], the Radio Equipment Directive, or the vehicle-type approval legislation. Differently from the Cybersecurity Act which sets up a voluntary cybersecurity certification framework for Information and communications technology (ICT) products, services and processes, the relevant EU product safety legislation lays down mandatory requirements.

The Commission has considered the extent to which the legislative framework is capable of tackling certain IoT/AI-related issues, with the help of, for instance, existing obligation of producers to carry out risk

[135] European Commission, Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, February 2020.
[136] Directive 2001/95/EC of the European Parliament and of the Council of 3 December 2001 on general product safety [2001], OJ L 11.
[137] Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC [2017], OJ L 117.
[138] Directive 2014/32/EU of the European Parliament and of the Council of 26 February 2014 on the harmonisation of the laws of the Member States relating to the making available on the market of measuring instruments (recast) [2014], OJ L 96.

assessment for their products' entire lifetime. Such assessment should therefore take into account any foreseeable future 'behaviour' of AI products as it evolves thanks to machine learning. In the Commission's view, unforeseen modifications in autonomous behaviour would require a new re-assessment of the self-learning product. Under existing legislation, producers have further notification obligations to competent authorities as regards risks having impact on users' safety, and must also ensure that a fault in a software integrated in a machinery product's control system does not lead to hazardous situations[139].

The Commission has, however, considered that some new issues arise by the nature of IoT/AI systems that may need addressing by new (or by extension of scope of the existing) product safety legislation, including the issue of possible mental health risks of users interacting with AI, IoT/AI systems dependency on accurate and relevant data, algorithmic opacity, increased complexity of products and systems, or issues related to loss of connectivity[140].

At this stage the following **risk mitigation** is recommended:

Ensure all devices, systems and applications used in the Smart City, and their suppliers/operators, conform to the applicable standards on product safety;

Monitor and adapt to legal and standardisation development in IoT/AI safety requirements.

Liability

Liability of a device or system manufacturer, seller or its operator, or a service provider, can either be contract-based (liability for damage that occurs as a consequence of breach of contract), or non-contractual (liability arises as a consequence of breach of an obligation imposed by law). Liability can be either fault-based (liability arises in case the person liable has breached a contract or another legal obligation by fault) or the liability can be strict – it will arise irrespective of the liable person's fault.

Legal landscape: As the Commission observes,[141] (civil) liability rules[142] are mainly provided by **non-harmonized national rules,** under which victims of damage can have several parallel compensation claims, based on fault or strict liability. These claims are directed often against different liable persons and have different conditions. National frameworks are complemented by the **EU Product Liability Directive (Directive 85/374/EC)[143]** providing an additional layer of user protection by means of a strict liability of the producer for physical or material damage caused by a defect in their product. Claims under the Directive may, but need not to be, based on a contract between the producer and the user. While the Directive will be applicable to many IoT elements, the Commission has considered that challenges may arise in connection with new technologies, e.g., whether software can be classified as a service, or rather a component of a product (and thus fall within the Directive's scope), how liability will be attributed in complex and integrated IoT environments (take algorithmic opacity, or who is responsible for data relied on by the systems) and when AI interacts with traditional technologies (such as traffic management), where partly automated AI systems will support human decision-making. The Commission has identified examples of AI devices and services that could have specific risk profiles by exposing the public at large to risk: devices that move in

[139] *Supra*.
[140] *Id*.
[141] European Commission, Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, February 2020; Expert Group on Liability and New Technologies New Technologies Formation: Liability For Artificial Intelligence And Other Emerging Digital Technologies, (2019).
[142] Lawyers call this "extra-contractual" or tortious liability, when damages occur from a civil wrong or a wrongful act, whether intentional or accidental, from which injury occurs to another.
[143] Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products [1985], OJ L 210.

public spaces (fully autonomous vehicles, drones, package delivery robots), or AI-based services such as traffic management services or power distribution management. The Commission is considering introducing strict liability for such cases and adopting a risk-based approach to regulation.

National liability regimes or, as the case may be, the EU Product Liability Directive would also apply in cases of liability for the accuracy and timeliness of data used in or relied on by IoT devices/systems. It is easier typically to establish liability where there is a contract between the liable party and the user (e.g., subscriber to a specific service that relies on the data at issue) and the damage occurs as a result of breach of contract (contracts typically include an obligation on the provider to sell product or provide a service free of defects). By contrast, claims for non-contractual liability in cases where a user relies on a generally available information that in turn has been derived from defective data may be more difficult to bring successfully, given the difficulty in establishing the information provider's duty of care and the chain of causation. Such cases typically involve complex questions of national liability regimes and procedure. The above considerations would also apply to cases of fraud/misrepresentation regarding the accurateness of the data when DUET and/or third parties rely on such data in their systems or provide services on its basis. While primary liability would lie with the original data vendor (or other party that committed the data fraud or misrepresentation), DUET could potentially be exposed to future liability claims from third parties damaged by the use of inaccurate data.

The abovementioned legal landscape has been complemented with some studies which are worth mentioning here. In particular, The 2018 'Study on emerging issues on data ownership, interoperability, access to data, and liability' by Deloitte[144] reviews the uncertainty around the suitability of current liability legislation in relation to the IoT and robotics data-driven applications. It concludes that there are many gaps in this respect.

The first gap that the study identified is that there is no universally applicable framework providing for liability rules. In addition, definitions of crucial liability concepts are also scattered across national liability regimes. The current EU-level legal framework is not suitable for these new technologies' complexities. Liability rules have been typically defined in relation to consumer protection without taking into account a business-to-business context. In addition, it is unclear to what extent the Product Liability Directive applies to 'data' or software. Thus, DUET is also potentially exposed to the issues of interpretation of the Product Liability Directive, as the case may be.

The problem arises since the Product Liability Directive addresses liability in relation to 'products', defined in the Directive as 'all movables' marketed in the EU. Such definition also encompasses any material products that incorporate digital data. What is still uncertain is whether it is applicable to purely digital 'products' that do not have any physical existence. Software or any digital data that has not been stored on a physical carrier do not unambiguously qualify or are disqualified as a 'product'. With IoTs, software meets a 'physical' product against a legal backdrop – both at EU and national level[145] – which has not caught up.

Where one would consider data as an intrinsic part of a product, e.g. a robot or IoT device, errors in the data could be seen as a defect in the product. Then, the concern would be addressed through the application of the rules of the Product Liability Directive which would provide consumers/end-users with redress for

[144] Deloitte, Impact Assessment support study on emerging aspects of data ownership, interoperability, re-usability and access to data and liability, Study Report for DG Connect, (2018).

[145] With exceptions. See UK government legislative initiative on IoT security at: https://www.gov.uk/government/consultations/consultation-on-regulatory-proposals-on-consumer-iot-security/outcome/government-response-to-the-regulatory-proposals-for-consumer-internet-of-things-iot-security-consultation.

damages caused by data errors, as is the case for data inaccuracy and untimeliness. This could help clarify, e.g., liability issues linked to the physical product, relying on data, that DUET is involved in making use of.

Problems arise if the data is provided by an external source other than the device or robot, or if a court holds that only the data was defective, but the source, device or robot was not. In this case, a consumer/end-user would not be protected by the said Directive. This is particularly true when the external source is assessed by national law as providing a service. This could be DUET's case. A person injured by a robot or IoT device which is used as a part of a service would have to demonstrate that the injury was a result of a defect in the product (i.e. in the robot or the device itself) in order for product liability law, partially harmonised at EU level through the above mentioned directive, to apply. The service provider might argue that the injury was a result of a problem with the service as a whole, including e.g. from errors in the software driving the service, and therefore that product liability law does not apply. In that case, the Deloitte study suggests that damages might be addressed under (potentially much more favourable) terms of service.

Further, it is noted that the functionality or fitness for a given purpose of a product does not fall within the purview of the Directive, unless the lack of functionality or fitness for purpose would create safety concerns. The Directive uses the criterion of the safety 'which a person is entitled to expect'. Yet, the Directive does not specify how the safety of a product must be assessed, therefore creating challenges on the determination of assurances a consumer is entitled to expect, and which tests a producer should be required to apply before bringing IoT devices and robots in the market. Particularly in the context of the IoT and robotics, it is unclear precisely what legitimate safety expectations the notion laid down under the Directive might entail. Software and robots are evolutionary, self- or quasi-autonomous, tools, which makes it difficult for both producers and users to predict their behavior and choices. Art. 7(b) of the Directive excludes liability if 'it is probable that the defect which caused the damage did not exist at the time when the product was put into circulation by him or that this defect came into being afterwards'. In the worst case scenario for the damaged person, these products will simply rule out any liability for the producer. Art. 7 (e) of said Directive exempts the producer from liability where he can prove that the state of scientific and technical knowledge at the time putting the product into circulation was not such as to enable the existence of the defect to be spotted. This is relevant as an exemption for rapidly evolving technologies, where it might be easier to argue that it was impossible for certain defects to be known to the producer. IoT products and software can be patched, updated or revised, by the producer or by third parties, in a way that can affect the safety over time.

Moreover, the Directive requires the injured person to prove the defect, damage and causality. Not only identifying the product and its producer is no easy task in this context but also proving a defect without expert assistance by an injured person is even more complicated. Finally, another hurdle is proving causality between the defect and the damage.

In addition, the Directive covers only specific types of damages, namely '(a) damage caused by death or by personal injuries; (b) damage to, or destruction of, any item of property other than the defective product itself, with a lower threshold of 500 EUR'. Thus, only material damage is covered. Hence, if a malfunctioning piece of software (including software embedded in an IoT device or robot) corrupts or destroys certain data but causes no other material harm, it is more likely than not that the Directive would not apply. This ambiguity should be resolved, along with the question of whether software and data as such qualify as a 'product'.

Another issue relates to the growing fragmentation of the national legal framework concerning the choices Member States have made in relation to liability under transposition of the abovementioned Directive

provisions, insofar as such national rules affect the data economy, robotics and the IoT. In this regard, no specific liability-related legislation has been enacted yet regarding autonomous devices, the IoT or robotics. One example, however, where there is an evolving scenario at national level is the UK, which in February 2020 announced its initiative to adopt an IoT law, which would incorporate notions of security-by-design into the legal framework. As it deals with security concerns, specifically tackled under Chapter 3, remains to be seen how such initiative will impact liability aspects of IoT. As such, a legal vacuum currently remains.

To sum up, the above mentioned study pinpoints the following characteristics and shortcomings liability rules insofar as the IoT and robotics technologies are concerned:

Extra-contractual liability rules depend on attribution of damage to a controller or custodian, provided that liabilities remain of a magnitude that the controller or custodian can manage and that identification is possible. If the robots obtain a degree of autonomy that could structurally create greater damage than the controller or custodian could assume, victims might not be able to obtain appropriate compensation. Extra-contractual liability rules do not provide for a defence against liability claims on the basis of the lack of foreseeability and or preventability of harmful behaviour. Significant autonomy (through machine learning or automated updates) could result in liability on the part of the controller or guardian, even though the behaviour causing the damage might not have been reasonably foreseeable for the controller or custodian. As regards evidentiary rules in relation to damage caused by robots, some Member States have rules in place, such as the rebuttable presumption of liability of the owner of the robot, or the application of hazard-based systems to robotics (creating again a presumption of liability for the users of hazardous devices). However, these are far from universal rules, resulting in an uneven and unpredictable landscape.

The study showed also that there is a misalignment between general extracontractual liability rules and product liability rules. As seen, there is absolutely non harmonisation of the former across the Member States, whereas the latter are more homogeneous. This can be problematic as an injured party who has no recourse to compensation on the basis of product liability law might be able to obtain compensation instead on the basis of general national extra-contractual liability laws (either from the producer or the owner/controller of the device), depending on which Member State is competent to hear the claim. All in all, current liability laws are unable to address liability challenges in relation to the IoT and robotics coherently.

Risk mitigation: scope, and where appropriate, case-study national liability regimes applicable to select DUET aspects in relevant jurisdictions with help of national lawyers/experts. At the project implementation/operational stage, responsible stakeholders will need to engage national lawyers for liability case management in the field. With regard to quality and accurateness of (supplied) data, pre-contractual (due diligence) and contractual measures (quality standards, liability limitation clauses) should be put in place that ensure that the data supplied and used in DUET systems or provided to third parties are in good order.

4.2.2 Sketching the issues: liability at data level when third party rights are involved

In May 2015, the European Commission announced the key actions to implement the digital single market[146]. In 2016, it announced a free flow data initiative, where it - *inter alia* - stated that it would address the restrictions on the free movement of data beyond personal data (Chapter 2), and look into the issue of data ownership. In particular, it affirmed that the free flow data initiative would tackle "*restrictions on the free movement of data for reasons other than the protection of personal data within the EU and unjustified restrictions on the location of data for storage or processing purposes*". It will "*address the emerging issues of ownership, interoperability, usability and access to data in situations such as business to business, business to consumers, machine generated and machine to machine data*"[147].

In 2017, the Commission published the Communication on Building a European Data Economy, accompanied by the Staff Working Document on the Free Flow of Data and emerging issues of the European Data Economy [148], where it acknowledged that the data economy requires, in order to thrive, access to large and diverse datasets, while ensuring that the protection of personal data is fully respected. The focus is on machine-generated data. Two aspects were highlighted: (a) data localisation restrictions; (b) barriers to data access and transfer in B2B relation. The law on data sharing (b) is not yet achieved in the EU.

In 2018, a study carried out for the European Commission's DG Connect highlighted that complexities in data handling can occur in each of the stages of the initial collection of data, the processing activities and the actuation of data[149]. Some of those complexities for data handling relate to liability when data hinges on third party rights, which does not touch upon the GDPR because that data is not personal. To what extent is DUET exposed?

Only certain of those aspects (in particular, data localisation restrictions) were addressed in the context of discussing the adopted Regulation on the free flow of non-personal data, which was analysed in Chapter 2. Above, under Section 4.2.1, we saw the liability aspects which may expose DUET when at stake are the physical devices, such as e.g. sensors, (which are data driven). We highlighted that the nature of IoT creates attribution challenges, such as to which entity is the behaviour of a device assigned, and who is required to bear the liability for any damage caused, considering that we are not necessarily talking about physical devices, but about services.

Here, we will look at what are the issues related to potential DUET exposure for breach of third party rights in data.

Against this background, it is important to highlight that data comes in many forms and it is often subject to third party rights which could restrict the way such data could be used. At this stage of the project, and due to the lack of visibility over the legal rights in data, it will be important to calibrate possible risks and mitigate accordingly. The questions that arise is to what extent do EU-wide instruments apply, what are the gaps and how to deal with them.

Typically, under EU law, data has been thought of as information, and less so as property, which has traditionally fallen outside the scope of the rights which apply to tangible property. This scenario could pose

[146] European Commission, Communication to the European Parliament, the European Council, the European and Social Committee and the Committee of the Regions, "A Digital Single Market Strategy for Europe", COM(2015) 192 final.
[147] *Supra* n. 27.
[148] European Commission, Building a European Data Economy, COM(2017) final, 10.1.2017; European Commission (2017), Staff Working Document, SWD(2017) 2 final.
[149] *Supra* n. 141.

liability concerns for DUET when the data is characterised by multiple overlapping legal rights which may affect their acquisition, use and disclosure: in this respect, the main rights in data, aside from data protection laws, under an EU perspective, are: (a) copyright (either in the data itself or an original database; (b) confidentiality/trade secrets; (c) EU database rights[150] (c) contractual rights.

The nature of the rights matters since an ongoing breach scenario potentially arises when data in which third party rights subsist end up being used by DUET without the required authorisation. There are hence different risks for different data, depending on how that data is protected.
Below, we highlight what regimes apply, depending on the nature of the rights protecting the third party data:

4.2.2.1 Potential liability for Third party Rights' infringements: copyright, trade secrets and right in databases

IP rights

Where data is protected by copyright or a database right, a competent court will have an array of remedies to penalise infringement of those rights. Protection of IP triggered by the Enforcement Directive (2004/48/EC) will come into play then. The transposition of the Directive varies from country to country. The Directive allows rights owners to seek the following:

(a) Corrective measures: in relation to goods found to have been infringing an intellectual property right;
(b) Injunctions: aimed at prohibiting the continuation of the infringement. Injunctions can be either preliminary or permanent. As a matter of fact, in some Member States courts have given broad injunctions, which is a risk that needs to be calibrated in the specific when third party rights on data will become apparent in the concrete;
(c) Damages (based on harm suffered, unlawful profits or a reasonable royalty);
(d) Legal costs and publication of judgments.

There are two likely scenarios: (a) first, the DUET architecture is not developed resting on the third party data and thus future use of DUET cannot result in infringement of those IP rights and no risk for DUET to be called upon as defendant and be subject to the remedies of the IP Enforcement Directive for breach of third party IP rights; (b) second, because third party data is relied on, and some of these data can be potentially infringed and DUET could be subject in a national court to a threat of injunction or demand for compensation. The second scenario is less likely. To mitigate this risk, it will be necessary to work in tandem with the legal compliance departments of IMEC and the technical DUET partners, to identify risks as they may materialise, and carry out the necessary due diligence when it comes to IP-related issues. It may also be necessary to foresee some form of insurance against potential breaches, if they cannot be easily and precisely quantified.

Breach of confidentiality and trade secrets

[150] This Section draws on Bond, T., Aries, N., "Forbidden Fruits: third party rights in AI training data, a European Perspective", Bird & Bird News Centre, (2019).

When the use of confidential data covered by a trade secret is unlawful, potential rights holders are subject to the protections offered by the EU's Trade Secrets Directive (2016/943/EU)[151]. Such a piece of minimum harmonization aims at harmonizing the national rules on trade secrets protection. Yet, from a substantive viewpoint it is unclear to what extent data produced by smart products are covered by a trade secrets protection and it can be opined that, given the developments of IoT, the Directive can be considered already as outdated.

Concerning its scope, Art. 2(1) mentions that for it to fall under the Directive's protection the "*know-how*" or the "*business information*" must be "*secret*" in the sense that it is not generally "*known*" among or readily accessible to persons within the circles that normally deal with the kind of information in question, the information must have commercial value because of its secrecy, and it has to be subject by the person lawfully in control of the information, to be kept secret. Some authors question how data created by sensors can fall within the scope of this Directive[152]. Among others, when data is generated in a network of different entities connected through the value network, it is unclear to a single person controlling the secret. The same authors consider that the Trade Secrets Directive only establishes a system of liability for tortious conduct (it does not protect against any use of the data, but against unlawful conduct, which can be regarded as "*contrary to commercial practices*")[153]. As such it does not solve the question of access to data in the era of IoT for other firms and public entities that may generate additional knowledge from that data through big data analysis. In addition, the protection under this piece of EU legislation is much narrower than an exclusive data use right[154].

When it comes to procedure, the same types of protection than those under the abovementioned IP Enforcement Directive apply. Namely, damages and preliminary/permanent injunctions prohibiting the unlawful use or disclosure of the confidential data can be sought against a defendant, and in the specific DUET, before a competent national court, for misuse of trade secrets. The EU's Trade Secrets Directive requires Member States to provide remedies related to infringing "*goods*" and defines this notion as encompassing "*the design, the characteristics, the functioning, the production process or marketing of which significantly benefits from trade secrets unlawfully acquired, used or disclosed*")[155]. Similarly to the IP Enforcement Directive, injunctions can be either temporary or permanent. Other remedies include damages and recall and destruction. Further to a potential risk materialising, it will be important to look at what national courts have said on how these provisions apply in the context of IoT/AI systems so as to precisely identify the risk and calibrate its mitigation. In this respect, it is important to highlight that the scope of the Trade Secrets Directive can be triggered when the AI is embodied in physical products, while it is less clear whether this also occurs when the applications are provided as a service, i.e. on a SaaS basis. As far as we are aware, the CJEU has never clarified how the Trade Secrets Directive provisions are to be applied in this latter case, so recourse to national law (and its potential fragmentation) must be had, regard having had to how the Trade Secrets Directive is transposed in national law. According to scholarship, indeed "*this may vary between European jurisdictions, although a survey of our colleagues in Germany, France, Italy, Spain and Finland suggests that courts in their jurisdictions are more likely than not to hold that the relevant national*

[151] Directive (EU) 2016/943 of the European Parliament and the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use or disclosure, 2016, OJ L 157/1.

[152] Drexl, J., "Designing competitive markets for industrial data – between propertisation and access", Journal of Intellectual Property, Information Technology and E-Commerce Law. (8), 4, (2017).

[153] *Id.*

[154] *Id.*

[155] Bond, Aries, cited.

implementation would cover an AI system provided as a service"[156]. In addition, since the Directive is a minimum harmonisation instrument, this means that in some Member States the national protection can be higher than that enshrined under the Directive[157].
By means of example, in the UK courts have a broad discretion to issue sanctions for misuse of confidential information by an AI trained system and would be allowed to issue springboard injunctions (which are limited to the time it would take someone starting from public domain sources to reverse engineer or compile the information).

As regards confidentiality of trade secrets in the form of data flowing in electronic communication networks and in provision of electronic communication services, the extant ePrivacy Directive as well as the proposed ePrivacy Regulation (both legal instruments discussed in Chapter 3 apply not only to personal data of natural persons, but to any data in electronic communications, including data relating to legal persons. To the extent the ePrivacy rules aim to guarantee confidentiality of electronic communications, they also serve to protect confidentiality of commercially sensitive information and other legal persons' legitimate interests in the data thus transmitted. In that context, the ePrivacy Regulation proposal makes an explicit reference to the EU Trade Secrets Directive.

EU database rights

As J. Drexl opines, while at first glance "*Database rights present an obvious property regime for controlling access to data*, *this kind of protection has limitations that explain why it will often fail to provide protection to data for the new business models of the data economy*"[158].

The EU legal regime for database protection is two-tiered: (a) copyright protection is granted to creative databases; (b) *sui generis* protection is granted to databases based on "*substantial investment*". Concerning (a), Art. 3(1) of the Database Directive clarifies that the character of a creative work defined as the author's own intellectual creation has to either relate to the selection or the arrangement of the database's contents. Under Art. 3(2) of this piece of legislation, the copyright protection for databases will not extend to the contents as such. Hence, even if the data is included in a copyrightable database, such copyright protection would not extend to that data.
Concerning (b), "*sui generis database protection may at first glance provide a better basis for protecting the data in the world of IoT*", but there are limitations from both the subject-matter of protection and the scope of protection. Authors clarify that the Database Directive is based on a database technology concept that no longer corresponds to the use of data in the era of IoT, because that concept is static and fails to adequately respond to the features of constantly changing datasets and real-time data services[159]. Authors identify limitations both with respect to the subject-matter and the scope of protection. They opine that the Database Directive is not it for this era[160].

Breach of the database rights triggers the protections under the abovementioned IP Enforcement Rights Directive. The same above mentioned considerations apply.

Contractual Rights and IP infringements

[156] *Id.*
[157] *Id.*
[158] Drexl, cited.
[159] *Id.*
[160] *Id.*

There is an additional risk that DUET may breach contractual rights which may give rise to damages claims: this is less likely to result in an injunction preventing commercialisation. The risk is related, in this sense, for DUET to be called upon in a national lawsuit with a third party seeking compensation. As seen, both the above mentioned IP Enforcement Directive and the Trade Secrets Directive allow for damages seeking before national courts.
In addition, injunctions for breach of contract are more likely when IP rights underpin the contract, such as for datasets used for AI training. Again, the approach to injunctions is national and, going forward with the project, it will be necessary to see what is the jurisdiction concerned by the contract as well as the national law and the courts' approach in that specific jurisdiction.

Further risks aspects:

(a) Rights holders' audits: when data is acquired from a rightsholder under a data licence, does the licence provide the rightsholder with audit rights. This can create an additional risk exposure since the audit allows the rightsholder to identify the use of data beyond the scope of the licence;
(b) Possible data sources: when the data has come from a sole provider, the risk of detection of the category of data being used beyond what is lawfully allowed becomes bigger;
(c) Identity of the rightsholder: commercial data suppliers are more aggressive in infringements.

4.2.3 Sketching the issues: Open data and data sharing liability aspects

EU law and national laws allow, and in some cases require, publication of data held by public entities. Also private entities may, for various reasons, wish to make data available publicly, share it among themselves, or share data with government entities (also called private sector data); they may do so to the extent they do not contravene any applicable laws. In some cases, private entities can also be obligated by law to share certain data with public authorities.

Making data available to the general public or sharing it with others may expose the publishing entity to legal liability in case the data gets misused or otherwise harm interests of others[161]. In the United Kingdom, for example, data regarding individual police recorded crimes have been made openly available to the public via the police.uk website since 2011. This had raised concerns that, by releasing data about burglaries in the UK, house-specific data enabled burglars to target households where new smart devices, computers, flat screen TVs could be stolen. In order to avoid such misuse, the police had to obfuscate the data using geo-masking techniques to reduce its spatial accuracy.[162]

[161] For limits of open data, see Delong de Rosnay, M., Janssen, K., Legal and Institutional Challenges for Opening Data across Public Sectors: Towards Common Policy Solutions, in Journal of Theoretical and Applied Electronic Commerce Research ISSN 0718–1876 Electronic VersionVOL 9 / ISSUE 3 / SEPTEMBER 2014 / 1-14 © 2014 Universidad de Talca – Chile.
[162] Tompson, L., Johnson, S., Ashby, M., Perkins, C., Edwards, P., "UK open source crime data: accuracy and possibilities for research", Cartography and Geographic Information Science, (2015).

As regards open data, Chapter 2 provided an overview of the existing EU legislation in the area[163]. The series of EU Open Data directives has aimed at unlocking the potential of big data (referred to also as Public Sector Information, or PSI) held and accumulated by public authorities and provide a set of rules for the data's re-use. The Open Data directives essentially mandate that public entities make PSI available subject to certain exceptions or subject to making access conditional on license limitations adhered to by the parties wishing to use the data.

Personal data is information especially protected in the EU, mainly via the GDPR (as explained in Chapter 2). To clarify the interplay with Open Data legislation, the GDPR allows the principle of public access to official data and even acknowledges that public access to such data may be considered in the public interest[164]. The GDPR thus allows such data to be made available on the basis of EU or national law (such as national law-level lists of categories of documents that must be published as open data by public entities), but these laws must reconcile public access to official documents and the re-use of public sector information with the right to the protection of personal data. The Open Data Directive, in turn, considers that rendering information anonymous is a means of reconciling the interests in making public sector information as re-usable as possible with the obligations under data protection law[165].

This principle of weighing interests in publication/data sharing with the interests of data subjects (persons or entities concerned by the data) should be the rule of thumb even where no personal data is involved. Open data may not only be misused, but its improper disclosure can harm legitimate interests such as when trade secrets or otherwise commercially sensitive information gets disseminated. The European Commission has in its working documents[166] identified a set principles for responsible sharing as regards private sector data sharing between businesses (B2B data sharing)[167] as well as businesses to government/public authorities sharing of private sector data (B2G)[168].

[163] Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information; Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information; Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information.

[164] Recital 154 of the GDPR.

[165] Recital 52 of the Open Data Directive.

[166] COMMISSION STAFF WORKING DOCUMENT Guidance on sharing private sector data in the European data economy SWD(2018) 125 final; Communication from the Commission to the European Parliament, the Council, the European economic and social Committee and the Committee of the Regions "Towards a common European data space" (COM(2018) 232 final).

[167] **Transparency** – to identify who have access to/and use the data generated by the product or service and specify the purposes of the data use; **Shared value creation** – to recognize when several parties have co-created the data; **Respect for each other's commercial interest** – to protect the commercial interests and secrets of data holders and users; **Ensure undistorted competition** – especially when exchanging commercially sensitive data.

[168] **Proportionality in the use of private sector data** – Any requests for supply of private sector data under preferential conditions shall be justified by clear and demonstrable public interest. The requests should be proportionate and the associated costs and efforts for the undertaking concerned should be reasonable compared with the expected public benefits. **Purpose limitation** - Purposes for the re-use of data by the public body must be specified, including a limited duration for use of the data. Additionally, specific assurances should be offered by the public body that the data will not be used for unrelated administrative or judicial procedures. **'Do no harm'** - Safeguards to ensure the protection of legitimate interests of the private party, notably the protection of its trade secrets and other commercially sensitive information. **Conditions for data re-use** - Seek to be mutually beneficial while acknowledging the public interest goal by giving the public sector body preferential treatment over other customers, particularly in terms of the agreed level of compensation. **Non-discrimination** - Ensure that the same public authorities performing the same functions are treated in a non-discriminatory way. **Mitigation of limitations of private sector data** - Ensure that companies supplying the data offer reasonable and proportionate support to help assess the quality of the data for the stated purposes, including through the possibility to audit or otherwise verify the data wherever appropriate. Companies should however not be

Public authorities may also limit the re-use of open data by imposing conditions in the **standard licenses**. This means that only the individuals or entities who agree to adhere to these conditions will be able to access and re-use the data. Art. 8 of the Open Data Directive specifies that such conditions must, however, be "*objective, proportionate, non-discriminatory and justified on grounds of a public interest objective*." These conditions should also not unnecessarily "*restrict possibilities for re-use and should not be used to restrict competition*" (e.g., by preferring certain economic operators over others without any objective justification). The Open Data legislation acknowledges that these licences may allow the data publishers (public sector body or public undertakings) to waive all liability with regards to the documents made available for re-use[169]. Businesses may share data among themselves by the help of so-called **data sharing agreements** with the aim to achieve adherence to the good principles of data sharing.

Rules and principles set out above aim to achieve a responsible sharing and publication of data and prevent any harm that might occur from use, misuse or mere dissemination of the data. Should harm to third parties nevertheless occur, the entity that made the data available might in principle be held liable for such damage. There is no harmonized EU framework on contractual or extra-contractual liability rules that would apply to such cases, except in a specific case where the open data could be considered part of a product in which case the EU Product Liability Directive could apply (see for more detail 4.2.1 above). Such cases typically involve complex questions of national liability regimes and procedure.

Risk mitigation: where a public or private entity decides to make data available or share it, the following best practices may be recommended to minimize impact on third party rights and legitimate interests:[170]

- **Pseudonymization** of personal data (see Chapter 2 for more detail);
- **Aggregation.** For example, weekly sales are summed and prices and promotions are averaged across stores within a market. Two types of aggregation are generally used:
  - Time aggregation = All data points for a single resource over a specified time period.
  - Spatial aggregation = All data points for a group of resources over a specified time period.
- **Geomasking**. Geographic masking is used to provide privacy protection for individual address information, while maintaining spatial resolution for mapping purposes. Such spatial resolution might get lost due to simple aggregation. Geo-masking is a class of methods for changing the geographic location of an individual in an unpredictable way to protect confidentiality, while trying to preserve the relationship between geocoded locations and occurrence of the mapped issue (e.g., disease, or occurrence of a burglary)[171].
- **Creating synthetic data**. When intense redaction is needed to protect data subjects' confidentiality, statistical agencies can release synthetic data, in which identifying or sensitive values are replaced with draws from statistical models estimated from the confidential data[172].

required to improve the quality of the data in question. **Transparency and societal participation** - Be transparent about the parties to the agreement and their objectives, and ensure that public bodies' insights and best practices of B2G collaboration will be disclosed to the public as long as those do not compromise the confidentiality of the data.

[169] Recital 44 of the Open Data Directive preamble

[170] For a summary of these best practices, see, e.g., Gupta, S., Schneider, M., Protecting Customers' Privacy Requires More than Anonymizing Their Data, Harvard Business Review Security & Privacy, June 01, 2018, available at: https://hbr.org/2018/06/protecting-customers-privacy-requires-more-than-anonymizing-their-data .

[171] E.g., Seidl, D.E., Jankowski, P., Clarke, K.C., "Privacy and False Identification Risk in Geomasking Techniques. Geographical Analysis", October 2017, available at https://onlinelibrary.wiley.com/doi/abs/10.1111/gean.12144

[172] E.g. J, "Jerome P. Reiter: An empirical evaluation of easily implemented, nonparametric methods for generating synthetic datasets", available at: https://ec.europa.eu/eurostat/cros/system/files/S5P1_0.pdf_en

- **Adding random noise**. For example, observations are grouped into deciles based on sales, and a random number is added to the sales in each decile.
- **Rounding**. For example, sales figures or other sensitive commercial information is rounded to the nearest hundred
- **Top coding**. For example, all sales above a threshold value, such as 100, are set equal to 100.
- **Swapping**. For example, observations are divided into groups and their sales data are exchanged.

Given that liability for potential damage resulting of use, misuse or dissemination of data by public or private entities would be governed by non-harmonized national liability regimes, there may be several contractual or extra-contractual measures available under national laws to limit or exclude such liability, such as use liability limitation contractual clauses and other risk-shifting agreements (e.g., insurance contracts), or unilateral liability disclaimers. National law experts should be consulted as regards potential use and effectiveness of such measures. In particular, scope exists for potentially mitigating risks through insurance, though this would require, on the one hand, quantification of such risks, which is hard to carry out at this preliminary stage, as well as potential budgeting of risk premiums, should such insurance coverage be considered.

In 2019, the Expert Group on Liability and New Technologies has suggested the introduction of an obligatory insurance scheme for certain categories of AI/robots as a possible solution to the problem of allocating liability for damage caused by such systems (sometimes combined with compensation funds for damage not covered by mandatory insurance policies)[173]. Currently, EU law requires obligatory liability (third-party) insurance, for instance, for the use of motor vehicles, air carriers and aircraft operators, or carriers of passengers by sea. Laws of the Member States require obligatory liability insurance in various other cases, mostly coupled with strict liability schemes, or for practising certain professions. The report also explains that new optional insurance policies (e.g. cyber-insurance) are provided for covering both first- and third-party risks and adds that "*overall, the insurance market is quite heterogeneous and can adapt to the requirements of all involved parties*". However, it points out at the downsides of such normative heterogeneity that, combined with the big number of actors involved in an insurance claim, "*can lead to high administrative costs both on the side of insurance companies and potential defendants, the lengthy processing of insurance claims, and unpredictability of the final result for the parties involved*".[174]

4.3. Artificial intelligence (AI) and Internet of things (IoT): specific ethical risks and risk mitigations

We herewith provide an overview of the pieces of legislation and the specific ethical risks connected to AI, as well as the high level risk mitigation. These will later on be elaborated on in the ethics deliverable.

4.3 Artificial intelligence: legal landscape

Below is an overview of the legal landscape covering the soft law instruments we identified as relevant for the purpose of this deliverable.

[173] Expert Group on Liability and New Technologies New Technologies Formation, LIABILITY FOR ARTIFICIAL INTELLIGENCE AND OTHER EMERGING DIGITAL TECHNOLOGIES (2019).
[174] *Id.*

First, we recap the legal instruments at International level:

G20 Ministerial Statement on Trade and Digital Economy (June 2019)

Here are some of the principles that were reaffirmed in this document.

- *Data Free Flow with Trust* - Digitalization gives the opportunity to promote inclusive and sustainable economic growth and also social and cultural progress and development, it fosters innovation, and empowers individuals and businesses, including micro, small, and medium-sized enterprises (Member StatesMEs) to benefit from emerging technologies and data. Cross-border flow of data, information, ideas and knowledge generates higher productivity, greater innovation, and improved sustainable development. At the same time, the free flow of data raises certain challenges. By continuing to address challenges related to privacy, data protection, intellectual property rights, and security, the free flow of data can be further facilitated, and consumer and business trust can be strengthened. In order to build trust and facilitate such free flow, it is necessary that legal frameworks both domestic and international should be respected. Interoperability of different frameworks will be encouraged.

- *Human-centered AI* - Recognizing the efforts undertaken so far by all stakeholders in their respective roles including governments, international organizations, academia, civil society and the private sector, and mindful of how technology impacts society, the G20 endeavoured to provide an enabling environment for human-centered AI that promotes innovation and investment, with a particular focus on digital entrepreneurship, research and development, scaling up of startups in this area, and adoption of AI by Member States MEs which face disproportionately higher costs to adopt AI technologies, which can help to promote inclusive economic growth, bring great benefits to society, and empower individuals. The responsible development and use of AI can be a driving force society, mitigating risks to wider societal values. The benefits brought by the responsible use of AI can improve the work environment and quality of life and create potential for realizing a human-centered future society with opportunities for everyone, including women and girls as well as vulnerable groups. At the same time, AI, like other emerging technologies, may present societal challenges, including the transitions in the labor market, privacy, security, ethical issues, new digital divides and the need for AI capacity building (see table *infra*). To foster public trust and confidence in AI technologies and fully realize their potential, a human-centered approach to AI is needed, guided by the G20 AI Principles drawn from the OECD Recommendation on AI, which are non-binding. Principles such as the ones of "*inclusive growth, sustainable development and well-being*", "*human-centered values and fairness*", "*transparency and explainability*", "*robustness, security and safety*" and "*accountability*". In pursuing human-centered AI, G20 members recognized the need to continue to promote the protection of privacy and personal data consistent with applicable frameworks. The G20 also recognized the need to promote AI capacity building and skills development.

- *Security in the Digital Economy* - Security in the digital economy is essential for strengthening public's confidence in digital technologies and the entire digital economy. It is important for governments and other stakeholders within their respective roles to address security gaps and vulnerabilities. These have a negative impact on digital innovations, and trust by consumers and businesses, and thus hinder from taking full advantage of the benefits of digitalization. Security in the digital economy is also important for governments in providing their services. Along with the rapid expansion of emerging technologies, including IoT, the value of an ongoing discussion on security in the digital economy is growing. It is recognised the global aspect of security in the digital economy

together with the need to develop localized and customized frameworks and methodologies. Industry-led and market-led global technical standards developed based upon principles of openness, transparency, and consensus help deliver interoperability. These promote trust, which is essential for enabling the benefits of the global digital economy. There is the need to raise awareness on the importance of actions to enhance security in the digital economy. It is recognised that the role played by stakeholders such as the private sector, the technical community and civil society, and relevant international organizations, is to further discuss those issues. There are relevant international organizations working on security in the digital economy within their existing mandates and efforts in security in the digital economy.

- Smart Cities - To contribute to sustainable and inclusive growth in urban areas where most of the world's population and energy consumption are concentrated, the G20 encourages networking and experience-sharing among cities for the development of smart cities. Implementations of smart cities should take into account transparency, resiliency, privacy, security, efficiency, and interoperability.

The OECD Recommendation on Artificial Intelligence (AI) (The OECD Principles) (May 2019)

Here we make a brief recap of this document:

- *Inclusive growth, sustainable development and well-being* - A responsible stewardship of trustworthy AI should be engaged in pursuit of beneficial outcomes for people and the planet, such as augmenting human capabilities and enhancing creativity, advancing inclusion of underrepresented populations, reducing economic, social, gender and other inequalities, and protecting natural environments, thus invigorating inclusive growth, sustainable development and well-being.

- *Human-centred values and fairness* - AI actors should respect the rule of law, human rights and democratic values, such as freedom, dignity and autonomy, privacy and data protection, non-discrimination and equality, diversity, fairness, social justice, and internationally recognised labour rights, throughout the AI system lifecycle. To this end, AI actors should implement mechanisms and safeguards, such as capacity for human determination, that are appropriate to the context and consistent with the state of Art.

- *Transparency and explainability* - AI Actors should commit to transparency and responsible disclosure regarding AI systems. To this end, they should provide meaningful information, appropriate to the context, and consistent with the state of art: to foster a general understanding of AI systems; to make stakeholders aware of their interactions with AI systems, including in the workplace; to enable those affected by an AI system to understand the outcome; to enable those adversely affected by an AI system to challenge its outcome based on plain and easy-to-understand information on the factors, and the logic that served as the basis for the prediction, recommendation or decision.

- *Robustness, security and safety* - AI systems should be robust, secure and safe throughout their entire lifecycle so that, in conditions of normal use, foreseeable use or misuse, or other adverse conditions, they function appropriately and do not pose unreasonable safety risk. To this end, AI actors should ensure traceability, including in relation to datasets, processes and decisions made during the AI system lifecycle, to enable analysis of the AI systems' outcomes and responses to inquiry, appropriate to the context and consistent with the state of Art. AI actors should, based on their roles, the context, and their ability to act, apply a systematic risk management approach to

each phase of the AI system lifecycle on a continuous basis to address risks related to AI systems, including privacy, digital security, safety and bias.

- *Investing in AI research and development* - Governments should consider public investment and encourage private investment in open datasets that are representative and respect privacy and data protection to support an environment for AI research and development that is free of inappropriate bias and to improve interoperability and use of standards.

- *Fostering a digital ecosystem for AI* - Governments should foster the development of, and access to, a digital ecosystem for trustworthy AI. Such an ecosystem includes in particular digital technologies and infrastructure, and mechanisms for sharing AI knowledge, as appropriate. In this regard, governments should consider promoting mechanisms, such as data trusts, to support the safe, fair, legal and ethical sharing of data.

- *Shaping an enabling policy environment for AI* - Governments should promote a policy environment that supports an agile transition from the research and development stage to the deployment and operation stage for trustworthy AI systems. To this effect, they should consider using experimentation to provide a controlled environment in which AI systems can be tested, and scaled-up, as appropriate. They also should encourage innovation and competition for trustworthy AI.

Council of Europe, Declaration by the Committee of Ministers on the manipulative capabilities of algorithmic processes (February 2019)

Digital services are used today as an essential tool of modern communication. This results in unprecedented amounts of new data that are constantly created with mounting speed and scale. Advanced technologies play a pivotal role in maintaining the efficiency and public service value of digitisation, in strengthening individual autonomy and self-determination, and in enhancing human flourishing by creating optimal conditions for the exercise of human rights. Technology is an ever growing presence in our daily lives and prompts users to disclose their relevant, including personal, data voluntarily and for comparatively small awards of personal convenience. Public awareness, however, remains limited regarding the extent to which everyday devices collect and generate vast amounts of data. These data are used to train machine-learning technologies to prioritise search results, to predict and shape personal preferences, to alter information flows, and, sometimes, to subject individuals to behavioural experimentation. The application and strengthening of data protection laws should consider the particular risks for and interests of those persons that may be especially unaware of the dangers of data exploitation. Increasingly, computational means make it possible to infer intimate and detailed information about individuals from readily available data. This supports the sorting of individuals into categories, thereby reinforcing different forms of social, cultural, religious, legal and economic segregation and discrimination. It also facilitates the micro-targeting of individuals based on profiles in ways that may profoundly affect their lives. The effects of the targeted use of constantly expanding volumes of aggregated data on the exercise of human rights in a broader sense, significantly beyond the current notions of personal data protection and privacy, remain understudied and require serious consideration. Member States should consider the need for additional protective frameworks related to data that go beyond current notions of personal data protection and privacy and address the significant impacts of the targeted use of data on societies and on the exercise of human rights more broadly. They also should acknowledge the need to consider, at both national and international levels, the growing onus on the industry across sectors to live up to their important functions and influence with commensurate

levels of increased fairness, transparency and accountability, in line with their responsibility to respect human rights and fundamental freedoms, and under the guidance of public institutions, drawing attention to the necessity of critically assessing the need for stronger regulatory or other measures to ensure adequate and democratically legitimated oversight over the design, development, deployment and use of algorithmic tools, with a view to ensuring that there is effective protection against unfair practices or abuse of position of market power.

EU level

Commission White Paper On Artificial Intelligence - A European approach to excellence and trust (February 2020)

AI is a strategic technology that offers many benefits for citizens, companies and society as a whole, provided it is human-centric, ethical, sustainable and respects fundamental rights and values. AI offers important efficiency and productivity gains that can strengthen the competitiveness of European industry and improve the wellbeing of citizens. It can also contribute to finding solutions to some of the most pressing societal challenges, including the fight against climate change and environmental degradation, the challenges linked to sustainability and demographic changes, and the protection of democracies and, where necessary and proportionate, the fight against crime. For Europe to seize fully the opportunities that AI offers, it must develop and reinforce the necessary industrial and technological capacities. As set out in the accompanying European strategy for data, this also requires measures that will enable the EU to become a global hub for data. The European approach for AI aims to promote Europe's innovation capacity in the area of AI while supporting the development and uptake of ethical and trustworthy AI across the EU economy. AI should work for people and be a force for good in society.

Commission Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics (February 2020)

The emergence of new digital technologies like AI, the IoT and robotics raise new challenges in terms of product safety and liability like connectivity, autonomy, data dependency, opacity, complexity of products and systems, software updates and more complex safety management and value chains. The current safety legislation contains a number of gaps that need to be addressed, in particular in the General Product Safety Directive, Machinery Directive, the Radio-Equipment Directive and the New Legislative Framework. [kn2] Future work on the adaptation of different pieces of legislation in this framework will be done in a consistent and harmonised manner. The new challenges in terms of safety also create new challenges in terms of liability. Those liability related challenges need to be addressed to ensure the same level of protection compared to victims of traditional technologies, while maintaining the balance with the needs of technological innovation. This will help create trust in these new emerging digital technologies and create investment stability. While in principle the existing Union and national liability laws are able to cope with emerging technologies, the dimension and combined effect of the challenges of AI could make it more difficult to offer victims compensation in all cases where this would be justified. Thus, the allocation of the cost when damage occurs may be unfair or inefficient under the current rules. To rectify this and address potential uncertainties in the existing framework, certain adjustments to the Product Liability Directive and national liability regimes through appropriate EU initiatives could be considered on a targeted, risk-based approach, i.e. taking into account that different AI applications pose different risks.

Commission staff working document: Liability for emerging digital technologies (April 2018)

The law of tort of EU Member States is largely non-harmonised, with the exception of product liability law under Directive 85/374/EC, some aspects of liability for infringing data protection law (Art. 82 of the General Data Protection Regulation (GDPR), and liability for infringing competition law (Directive 2014/104/EU). There is also a well-established regime governing liability insurance with regard to damage caused by the use of motor vehicles (Directive 2009/103/EC), although without touching upon liability for accidents itself. EU law also provides for a conflict of tort laws framework, in the form of the Rome II Regulation. On a national level, it can generally be observed that the laws of the Member States do not (yet) contain liability rules specifically applicable to damage resulting from the use of emerging digital technologies such as AI. By way of exception, those jurisdictions that already allow the experimental or regular use of highly or fully automated vehicles usually also provide for coverage of any damage caused, be it only by way of insurance or by reference to the general rules. Apart from this legislation, the harmful effects of the operation of emerging digital technologies can be compensated under existing ('traditional') laws on damages in contract and in tort in each Member State. This applies to all fields of application of AI and other emerging digital technologies the NTF of the Expert Group have analysed. In general, these domestic tort laws include a rule (or rules) introducing fault-based liability with a relatively broad scope of application, accompanied by several more specific rules which either modify the premises of fault-based liability (especially the distribution of the burden of proving fault) or establish liability that is independent of fault (usually called strict liability or risk-based liability), which also takes many forms that vary with regard to the scope of the rule, the conditions of liability and the burden of proof. Most liability regimes contain the notion of liability for others (often called vicarious liability). However, these regimes may not always lead to satisfactory and adequate results. Furthermore, given the significant differences between the tort laws of all Member States, the outcome of cases will often be different depending on which jurisdiction applies. As experience with the Product Liability Directive has shown, efforts to overcome such differences by harmonising only certain aspects of liability law may not always lead to the desired degree of uniformity of outcomes. It is possible to apply existing liability regimes to emerging digital technologies, but in light of a number of challenges and due to the limitations of existing regimes, doing so may leave victims under- or entirely uncompensated. The adequacy of existing liability rules may therefore be questionable, considering in particular that these rules were formulated decades or even centuries ago, based on even older concepts and incorporating a primarily anthropocentric and monocausal model of inflicting harm. Digitalisation brings fundamental changes to our environments, some of which have an impact on liability law. This affects, in particular: complexity, opacity, openness, autonomy, predictability, data-drivenness, and vulnerability of emerging digital technologies. Each of these changes may be gradual in nature, but the dimension of gradual change, the range and frequency of situations affected, and the combined effect, results in disruption.

Other relevant recommendations can be found in the following policy documents:

- High-Level Expert Group on Artificial Intelligence: Policy and Investment Recommendations for Trustworthy AI (June 2019)
- Commission Communication: Building Trust in Human-Centric Artificial Intelligence (April 2019)
- High-Level Expert Group on Artificial Intelligence: Ethics Guidelines for Trustworthy AI (April 2019)
- Commission Communication: Coordinated Plan on Artificial Intelligence (December 2018)
- Commission Communication: Artificial Intelligence for Europe (April 2018)

National level

The following paragraph provides an overview of national-level initiative in the field of AI and ethics in one of the relevant countries. **The Czech Republic** has been fairly proactive in the field of AI policy making. In

December 2018, the Government Office in cooperation with other stakeholders (the Czech Technical University, the Academy of Sciences, and the Czech Technology Centre) published an Analysis of the Development Potential of Artificial Intelligence, which included among its key findings the need for developing an adequate regulatory framework. An accompanying special report on legal-ethical aspects of AI development identified several key issues including AI definition, legal responsibility, personal data protection, confidentiality of electronic communications and treatment of non-personal data, cybersecurity, and ethical issues ("ethics for design", "ethics by design", "ethics in design"); it also provided short vertical analyses through specific sectors affected by AI such as healthcare, journalism, banking & finance, economic competition, or autonomous vehicles. A National AI Strategy, adopted in May 2019, edifies further these findings with a chapter setting out short-, medium- and long-run objectives regarding legal, ethical, and security issues. In addition, in late 2019 Czechia also issued a position non-paper discussing the EU regulatory framework on AI, taking views on several aspects of AI regulation including questions of various regulatory approaches and defining horizontal red lines in the areas of facial recognition systems, social credit scoring systems, recognition of users' emotions, limitations of liability, etc. The Member State is also a signatory to the 'Visegrad 4 countries' thoughts on Artificial Intelligence and maximising its benefits' paper published in April 2018, proposing a set of regulatory priorities in the field.

Relevant Member States' as well as city-level stakeholders' activities in the AI and ethics field will continue to be monitored and elaborated on in the ethics deliverable.

4.3.1 AI and potential ethics risks: An overview

We provide below an overview of the legal (soft-law) instruments when at stake is AI technology, dealing with ethical issues of AI, and an overview of risks for macro-areas, and risk mitigations.

Risk concern	Comment	Risk mitigation (high level)
Human-machine interaction/Lack of human-centric approach	AI enables machines to 'learn' and to take and implement decisions without human intervention. The 'loss of human touch' may impact the well-being of users and can cause distrust, physical or economic harm to, or mental safety risks of users.	Human oversight at appropriate stages of AI system design, operation, and output control. Effective redress process involving humans. Carrying out trustworthy AI impact assessment and/or stakeholder consultations. Users need to be informed that they are interacting with an AI system, not a human being (also relevant in the issue of 'transparency'). Take regard to Art. 13(2)(f) GDPR concerning disclosure of information about the existence of automated decision-making in personal data processing.

Bias/Discrimination/Unfairness	Bias, when present in AI systems, may have large scale negative effects. Risk of discriminatory profiling (e.g. predictive policing and crime prevention, social field, employment). Discriminatory or unfair effects may be exacerbated further by machine learning processes (data mining).	Ethics in/by design (prevent flaws in the original system design), prevent flaws in practical impacts of correlations or patterns that the system identifies in a large dataset. Need to: (a) involve different individuals from multiple disciplines to define the risks; (b) monitoring the implementation of the project; (c) reporting outward (ethics committee?) Training of AI systems (machine learning) on data that are sufficiently broad, representative and diverse. Compliance with ECtHR case law and the Racial Equality Directive 2000/43/EC. Compliance with the GDPR: algorithmic due process under the GDPR/ how could data protection law mitigate discrimination risk?
Lack of transparency	Lack of transparency in how algorithms work (opacity of AI, the 'black box effect') makes it difficult to identify and prove breaches of laws. Impact on users' trust in the technology.	Users need to be informed that they are interacting with an AI system, not a human being. Provide clear information on the AI system's capabilities and limitations. Anticipate, monitor and comply with developing rules on transparency requirements, keeping of records and data. Keep accurate records on training data, datasets themselves, documentation on programing and AI systems training methodologies. Transparency enhancing technologies: aims at making information flows more transparent through feedback and awareness, thus enabling individuals and collectives to better understand how information is collected, aggregated, analysed and used for decision-making.

Contractual Liability/Accountability/Safety	Safety risks when AI technologies are embedded in products and services (e.g., autonomous cars causing accidents to flaw in object recognition technology). New types of risks linked to cyber threats, personal security risks, loss of connectivity, etc. Difficulties in allocation of responsibilities between different economic operators. Some sectors may be considered 'high risk' by regulators (e.g., healthcare, transport, use of AI for remote biometric identification and other surveillance technologies). Possibility of regulation in these areas to shift from 'risk-based' to 'precautionary principle-based' (stricter) approach.	Ensure full compliance with existing rules on product liability and compliance (e.g., EU Product Liability Directive, the General Product Safety Directive, further sectorial product requirements legislation). Ensure full compliance with existing cybersecurity rules. Security-by-design approach. Anticipate, monitor and comply with developing rules on robustness and accuracy of AI systems and data quality. Implement human oversight from the product design and throughout the lifecycle of AI products/systems. Ensure auditability of AI systems.
Impact on privacy/personal data protection	AI increases possibilities to track and analyse daily habits of people, may be used to retrace and de-anonymise data, create data protection risks even in datasets that *per se* do not include personal data.	Ensure full GDPR (in particular, Art. 4(4) GDPR on automated decision making) and ePrivacy rules compliance; deploy advanced encryption and anonymization measures. Cooperate with data protection authorities.

4.4 IoT: Legal landscape

Similarly to the above, we provide herewith an overview of the legal (soft-law) instruments when at stake is IoT technology, dealing with ethical issues of IoT.

EU level:

The regulation of telecommunications and electrical equipment in the EU is a complex panorama, which has a direct impact on IoT developers. This Section offers a main overview of the most relevant legislation, including the **Net Neutrality Regulation 2015/2120**, the **New Legislative Framework** (composed by the

Radio Equipment Directive (2014/53/EU), the **Low Voltage Directive** and the **Electromagnetic Compatibility Directive**). We will look into those instruments in turn:

The Net Neutrality Regulation

The Net Neutrality Regulation (Regulation (EU) 2015/2120 of 25 November 2015 by the European Parliament and the Council): net neutrality is the principle that Internet Service Providers (ISPs) should treat all traffic on the Internet the same, not discriminating or charging differently by user, content, website, platform, application, type of attached equipment or mode of communication. Columbia University Professor Tim Wu coined the term in 2003.

EU rules on open internet access apply as of 30 April 2016, further to the Adoption of the Net Neutrality Regulation. The Regulation enshrines the principle of non-discriminatory traffic management and ensures that common rules on open Internet access apply throughout Europe. Their enforcement should be ensured by national regulatory authorities (NRAs), which should respect the guidelines adopted by BEREC in August 2016. It further clarifies the requirements concerning the provision of specialised services with specific quality requirements by Internet access providers and providers of content and applications. The BEREC guidelines help NRAs to assess - *inter alia* - agreements and commercial practices and "*specialised services"* against a common benchmark, and to reach consistent decisions. Finally, it creates an enforceable right for end-users in the EU to access and distribute internet content and services of their choice.

The New Legislative Framework

To improve the internal market for goods and strengthen the conditions for placing a wide range of products on the EU market, the New Legislative Framework was adopted in 2008. It *inter alia* improves market surveillance rules, sets clear and transparent rules for the accreditation of conformity assessment bodies and boosts the quality and confidence in the conformity assessment of products through stronger and clearer rules. It also clarifies the meaning of CE marking and lays down a common legal framework for industrial products through a toolbox of rules for use in future legislation. Of relevance for IoT are the following pieces of legislation, adopted in 2014, and which have applied in the EU since 2016:

Radio Equipment Directive (2014/53/EU)[175]

The Directive was adopted in order to provide a harmonized framework across Member States related to making products defined as "*radio equipment*" (including wireless modules), to meet the world's demand for IoT, smart cities and wireless technologies. Radio equipment is "e*quipment which intentionally emits or receives radio waves for the purpose of radio communication or radio-determination and makes systematic use of radio spectrum*": all such equipment fall within the scope of this Directive. RED will apply to most IoT devices, as these tend to have some form of radio connectivity. Other IoT devices which do not have radio connectivity will be instead subject to the other two pieces of legislation discussed below.

In a nutshell, the directive "*defines essential requirements for health, safety, electromagnetic compatibility and the efficient use of the radio spectrum to avoid interference. It applies to all products using the radio frequency spectrum (even for secondary functions such as location positioning, including GPS, Wi-Fi), including many IoT devices*", Manufacturers or importers must carry out a conformity assessment that will

[175] Directive 2014/53/EU of the European Parliament and of the Council of 16 April 2014 on the harmonisation of the laws of the Member States relating to the making available on the market of radio equipment and repealing Directive 1999/5/EC [2014] *OJ L 153.*

include safety and risks, taking into account the reasonably foreseeable usage conditions. In terms of risks, the manufacturer must gouge the potential misuse of the equipment and carry out safety checks. The Directive allows for self-certification and it gives the possibility to obtain a certification of quality from a recognized body among a list of European technical organisations. A "*grey area is the fact that the directive does not cover kits used only for research and development*". Another grey area is "*software compliance: if the operation of devices includes open source software, manufacturers need to test for this possibility*".

In order to ensure that radio equipment uses the radio spectrum effectively, radio equipment is constructed so that:

- in the case of a transmitter, when the transmitter is properly installed, maintained and used for its intended purpose, it generates radio waves emissions that do not create harmful interference, while unwanted radio waves emissions generated by the transmitter (e.g. in adjacent channels) with a potential negative impact on the goals of radio spectrum policy should be limited to such a level that, according to the state of the art, harmful interference is avoided; and
- in the case of a receiver, it has a level of performance that allows it to operate as intended and protects it against the risk of harmful interference, in particular from shared or adjacent channels, and, in so doing, supports improvements in the efficient use of shared or adjacent channels.

Although receivers do not themselves cause harmful interference, reception capabilities are an increasingly important factor in ensuring the efficient use of radio spectrum by way of an increased resilience of receivers against harmful interference and unwanted signals on the basis of the relevant essential requirements of Union harmonisation legislation.

Interworking via networks with other radio equipment and connection with interfaces of the appropriate type throughout the Union is necessary in some cases. Among others, protection from fraud may be enhanced by particular features of radio equipment. Radio equipment is, therefore. in appropriate cases designed in such a way that it supports those features. Radio equipment can be instrumental in providing access to emergency services and is therefore, in appropriate cases, designed in such a way that it supports the features required for access to those services.

All economic operators intervening in the supply and distribution chain take appropriate measures to ensure that they only make available on the market radio equipment which is in conformity with this Directive. In particular, the following must be ensured: a high level of protection of health and safety of persons and of domestic animals, and the protection of property, an adequate level of electromagnetic compatibility, an effective and efficient use of radio spectrum and, where necessary, a high level of protection of other public interests, and to guarantee fair competition on the EU market. It is necessary to provide for a clear and proportionate distribution of obligations which correspond to the role of each economic operator in the supply and distribution chain.

The Low Voltage Directive (Directive 2014/35/EU)

The Directive, applicable since 20 April 2016, covers health and safety risks on electrical equipment operating with an input or output voltage of between 50 and 1000 V for alternating current, 75 and 1500 V for direct current. National authorities are responsible for implementing and enforcing the LVD.

The Electromagnetic Capability Directive (Directive 2014/30/EU)

The Directive, which applies since 20 April 2016, ensures that all electrical and electronic equipment, placed on the EU market, comply with the allowed adequate level of electromagnetic compatibility. The electrical products falling under its scope shall not generate or be affected by any electromagnetic disturbance. To this end, the Directive sets out mandatory essential requirements that all equipment within its scope need to comply with. The technical specifications are not spelled out, but the results that need to be attained are specified.

4.4.1 IoT: An overview of potential ethics risks

We provide herewith an overview of the legal (soft-law) instruments when at stake is IoT technology, dealing with ethical issues of IoT, and an overview of risks for macro-areas, and risk mitigations.

Risk concern	Comment	Risk mitigation (high level)
Risk of user-related vendor-lock in/risk of anti-competitive behaviour	Lock-in to certain providers is a real likelihood. When it is difficult to wish to switch to a different provider, concerns of anti-competitive behaviour may arise.	Data portability and interoperability Importance of standardisation.
Product liability risks	Analyse liability under the Product Liability Directive. Potential gap: does it apply? Existing liability regimes in Member States provide answers to the questions of whether the victim of any risk that materialises can seek compensation from another and if so under what conditions: gaps. Discuss liability at different levels: operators' strict liability, producer's strict liability, fault liability and duties of care. Redress between multiple tortfeasors.	Contractual risk mitigation: see above. This presupposes a contractual relationship (user-DUET, DUET-third party provider). National level: discuss extra-contractual (tort) liability general principles.
False advertising, unfair and deceptive trade practices and fraud	Litigation risk for alleged misrepresentation, false advertising, failure of proper disclosures.	National legislation.
Sustainability		Factor in risk mitigation in decision making during the whole lifecycle World Economic Forum: IoT, guidelines for sustainability

Security (safety)	Issues in terms of cyber-security (Chapter 3). Discuss potential product safety risks (malfunction by defect or update, loss of connectivity, data quality and integrity concerns, physical dangers)	Security by default. Discuss safety implications: gaps of Product Liability Directive?
Loss of user control	Loss of agency: strong mediation inherent to IoT developments will lead to shifting or delegation of human autonomy and agency to the objects of IoT. -Right to integrity of a person (Art. 3 of the ECHR)	Human centred approach by design Federated architectures for greater control and agency. Ensure data management when data is released to another party (means of control). Record data flows. Enable audits. Accountability
Potential violation of specific human rights (Art. 21 ECHR)	Profiling, targeted advertising which could create/perpetrate bias and discrimination Large data stores could lead to discriminatory impact through codified and inferential discrimination	Transparency is an important tool which provides data subjects with the possibility to make informed decisions as well as to ascertain the basis on which decisions about them are taken, thereby reducing the risk for discrimination.
Opacity/Lack of transparency	See above on AI (black boxes)	Vendor certification to reduce risk of opacity.. IoT databox model.
Damage to data and related liability	Can there be liability in tort where the relevant data was protected by IP law or trade secret protection?	National law relevant question: basis for tort liability. Infringement of product safety legislation. Insurance and compensation funds.
Impact on privacy/personal data protection	Issues with loss, violation of individuals' privacy (traceability, profiling, unlawful processing) Challenges to notion of consent under the GDPR. Reduction of de-identification possibilities Repurposing and mission creep (Chapter 2)	Ensure full GDPR and ePrivacy rules compliance; deploy advanced encryption and anonymization measures without tampering with usefulness of the data. Cooperate with data protection authorities. Have DPIAs in place. Have contingency plans in place for cases of risk of data breaches.

5. Conclusion

This deliverable 1.1 is a starting point for building a legal roadmap for envisaged activities of the DUET Digital Twins consortium. It addresses the legal and ethical aspects of using the advanced capabilities of the Cloud and high-performance computing (HPC) through the use of (big) data, IoT and AI technology as an integral part of policy and decision-making processes in the project's planning and implementation phases. While the use of data and technology is capable of enabling many improvements for the public good, it may also generate risks vis-à-vis specific stakeholders or the general public. Such risks should be addressed, where possible, by building a good understanding of legal and ethical requirements and designing a set of adequate safeguards. In the course of upcoming project design and implementation milestones, applicable legal requirements need continuous monitoring, and result in adjustment of compliance and feasibility assessment where necessary.

Chapter 2 provided an overview of the current EU legal landscape related to data governance. The chapter preliminarily identified smart city-related privacy concerns and suggested an early set of possible risk mitigation procedures. It provided a high-level mapping of the relevant legislative and policy initiatives with regard to data governance, highlighting the aspects that are prima facie most relevant to the DUET project.

The chapter took close account of the main principles of the GDPR, the primary piece of EU data protection legislation. Processing large amounts of personal data may be central to many envisaged DUET's activities; for example, we estimate that DUET may use extensively so-called a mixed datasets, a blend of personal and non-personal data to feed smart city systems and processes. Chapter 2 showed that the dividing line between what constitutes personal and non-personal data causes application problems, further exacerbated by the fact that data can be rather dynamic (change their classification as personal or non-personal data over time). The GDPR along with the EU Regulation on the Free Flow of non-personal data and the European Commission Guidance on the interplay between these two pieces of legislation clarifies the difference. However, the GDPR has the limit of taking a snapshot and static approach on what personal and non-personal data, while also not dealing with data sharing. Sharing of data with private stakeholders, public authorities, or general public (open data) is a key issue for the DUET project. The EU Open Data legislation allows (or indeed, requires) data sharing and its re-use between public and private entities, and the abovementioned Regulation on the Free Flow of non-personal data establishes the principle of data availability which ensures that competent authorities are able to access non-personal data - including anonymized personal data - for supervisory control wherever it is stored or processed in the EU. Chapter 2 explained further particularities with regard to these issues, whereas Chapter 4 shed light on potential liability aspect of data sharing.

Chapter 2 further outlined the ePrivacy Directive, which imposes additional data processing requirements on actors who provide electronic communication services, which may encompass IoT/machine-to-machine flow of data, and to those who wish to use information on, or communicate with, users' terminal equipment (mobile phones, connected vehicles, etc.). These requirements may soon change with the adoption of the e-Privacy Regulation currently pending in the EU legislature, and Section 2.2.3 provides an overview of selected major amendments. In the same vein, as announced by the European Commission in its Data Strategy, a future EU Data Act deemed to harmonize the scenario on data governance, access and reuse is equally high in the agenda and worth of being continuously monitored.

Soft law helps to fill gaps of the non-existing or patchy legislation by guidance. Chapter 2 further provides an overview of relevant guidelines issued by the European Data Protection Supervisor (EDPS) and the European Data Protection Board (EDPB), which clarify certain central GDPR-related notions, such as who are data controllers and data processors and their respective obligations, or the necessity and proportionality of measures involving personal data processing. Additionally, the guidelines explain regulators' position on certain specific data processing scenarios such as data collection through video devices, mobility applications, or the use of location data and contact tracing tools. The emerging legal roadmap should ensure that DUET takes note of the applicable hard and soft law in order to achieve an adequate level of compliance in all sensitive areas.

Chapter 2 further dived deeper into certain data governance risks that appear highly relevant for smart cities:

- A feeling of surveillance and 'dataveillance' may decrease trust in data governance and system integrity, thus possibly creating a barrier to a smooth implementation of smart city projects. Stakeholders should consider to help build trust (of public authorities, and the general public alike) in the DUET Digital Twins by means of transparent communication and educational initiatives. DUET partners should devise not only compliant, but also sound, transparent and participated data governance policy, and consider awareness raising campaigns or digital literacy programs to enhance public response.

- The distinction between personal and non-personal data, and the issue whether personal data has been irreversibly de-identified (and thus can be considered non-personal going forward) pose difficulties in practice. When personal data is in play (whether or not combined with non-personal data in a mixed dataset), it will be necessary to identify an appropriate legal basis for their processing before DUET engages in such activities. Obtaining a meaningful informed prior consent may be difficult given the ubiquitous computing on which DUET's activities may depend. Further legal bases may be available, e.g., collection and processing of data in the public interest. Section 2.3.2 explored several practical solutions and best practices in this regard.

- Big data can be put to an initially unintended or unexpected use, which depart from the use for which an initial lawful ground for processing has been provided. In a practical example, data generated with the intention beneficially to personalize a particular smart city service may be utilized for that purpose, but should not be used to profile customers with the effect of discriminating people unlawfully, or to make profit by selling it to advertisers. In the light of the GDPR principles of data minimization, purpose limitation, storage limitation, integrity, accuracy, transparency and fairness of data, Section 2.3 on the privacy risks and Section 2.4 on the risk mitigation plan elaborated on those risks and suggested technical as well the organizational and logistical risk mitigation measures that DUET could adopt in order to comply with the applicable law.

The answer to many privacy concerns can be summarized by the privacy-by-design (PbD) approach, i.e., protecting privacy by embedding it into the design specifications of technologies, business/organizational practices, and physical infrastructures. The last section of Chapter 2 illustrates the possible PbD solutions capable of mitigating privacy risks, such as restricting the amount of data applications collect to the necessary minimum, encrypting data flows as default, anonymization and pseudonymization of personal data at source, use of participatory Privacy Impact Assessments (PIAs), approved certification mechanisms,

embedding privacy notices systems in user-friendly ways at appropriate times and places, specific risk mitigation procedures for specific categories of data, and others.

Chapter 3 presented the legal landscape dealing with the (cyber-)security aspects of the technologies that the DUET Digital Twin will potentially make use of and discussed the security issues that may arise during the projects' design, implementation and maintenance phases. It further laid down guidelines, drawing from the best practices upheld by ENISA (the European Union Agency for Cybersecurity) as well as the industry as to the security solutions that need to be implemented to keep the infrastructures secure. Section 3.2 provided an overview of the EU legislation on the matter and made clear that, when cloud computing is at stake, various laws may apply concurrently, according to the location of the cloud provider, the cloud user, the data subject(s), the servers, the legal jurisdiction of the contract between the parties and any common legal frameworks in force in the relevant locations. Hard law legislative instruments of relevance were taken into consideration and shortly explained, namely the Directive on security of network and information systems (NIS Directive), the Regulation 2019/881 on ENISA and ICT Cybersecurity Certification (Cybersecurity Act), the Regulation on the Free flow of Non-Personal Data and, with specific concern to the telecoms sector, the European Electronic Communications Code (EECC). As well, several soft instruments adopted by ENISA, the International Organization for Standardization and the International Electrotechnical Commission Joint Technical Committee 1, Information Technology are relevant in that they advance state-of-the-art recommendations and shape best practices necessary for tackling security issues throughout DUET Digital Twins' various phases.

Furthermore, Chapter 3 briefly illustrated the main concerns raised by the literature as well as the best practices emerging from the business sector, and the Cloud Security Alliance. As seen, vulnerabilities may arise throughout all the stages of the DUET's project, from the design and planning stages through the implementation, the operational and maintenance phases. In the design and planning stages, the vulnerabilities identified were several-fold, such as - for instance - concerns with cryptographic protection (e.g. encryption), on rest and on transit, loss of encryption keys by the data controller, issues of authentication capabilities and authorization, secure configuration, issues with updates of software as well as tampering by unauthorized sources, phishing, system malfunctions and crashes at software and hardware level. A list of vulnerabilities arising during the technology implementation phase and the operation and maintenance phase was also provided.

Starting by the security by design approach adopted by ENISA - from the viewpoint of the devices and infrastructure at every step of the development lifecycle- , Chapter 3 further sketched a risk mitigation plan. In this respect, risk and threat analysis must be continuously performed, involving cybersecurity experts from the very early stages of the design process of DUET. It was therefore recommended to devote a chapter of each DUET's design document to the security of all the information and control systems. A number of best practices in terms of security by design and security during both the implementation, operation and maintenance phases were also suggested. Those include, just to mention a few, strong cryptography of data, enhanced authentication procedures, secure configuration by default, testing procedures concerning security requirements' compliance, system hardening, use of certification scheme, monitoring, patching and regular auditing, recovery planning in case of compromised systems. Finally, Chapter 3 drew up criteria for a sound management risk plan which takes into account industry-wise standards applicable to the smart city context.

Moreover, IoT and AI and big data raise a number of ethical issues beyond data privacy and cybersecurity, including those of potential DUET's legal liability for economic, physical or mental harm caused by such

systems or datasets. Chapter 4 explained that the legal framework for regulating these is not yet fully developed at the EU or national level but that there is an increasing awareness throughout jurisdictions that a responsible approach is needed to ensure safe and beneficial use of the technologies and data.

The EU has achieved important milestones recently. “A European Strategy for Data”; the Commission Report on the safety and liability implications of AI, the IoT and robotics; and the Commission White Paper on AI, are pivotal policymaking documents adopted in February 2020 that will shape the institutional thinking about these issues going forward, and to some extent already outline the playing field for actors such as DUET. Developments should be closely monitored.

The use of AI and IoT systems and data pose old questions of liability for faulty products, service, data, or for intellectual property (IP) breaches subject to existing regulation (e.g., EU safety or product liability legislation) and national civil liability frameworks, but often with a new twist. For instance, the sheer scale of data may cause exponentially higher or wider-spread damage when misused. Novel questions arise also with regard to quality of data used (accurateness, timeliness), attribution of liability in complex systems (e.g., mixed data sets or a complex infrastructure), suitability of current IP enforcement framework for databases, or the manner in which data are made available or shared (open data). DUET can preventively achieve much by designing its activities with these liability aspects in mind (ethics-by-design), but also needs to understand that the level of harmonization between EU Member States’ liability laws is low and that a tailored approach for relevant jurisdictions may be necessary.

In addition, AI, machine learning, and IoT have such a disruptive potential that they raise new questions by their nature: algorithmic bias and system opacity, lack of human interaction, vendor lock-in risks, or safety risks are aspects that may call for brand new ways of regulating. It is advisable to monitor regulatory developments in these areas closely and where appropriate, approach responsible regulatory bodies and other stakeholders proactively.

6. Annex I - Matrix of data governance and privacy topics - Scoping the legal issues through the Sidewalk Toronto Project

When scoping potential legal issues, we have looked for inspiration at other Smart Cities initiatives, including and in more detail the Sidewalk Toronto project[176]. Sidewalk Toronto is an urban development project operated until May 2020 by Sidewalk Labs[177], an Alphabet (Google) subsidiary, at Quayside, a waterfront area in Toronto, Ontario, Canada. In charge of steering the project in line with public interest is Waterfront Toronto[178], a body created by the governments of Canada, Ontario, and the City of Toronto.

After winning a request for proposals in October 2017, Sidewalk Labs committed USD 50 million to test pilot projects and in June 2019 published the Master Innovation Development Plan, a detailed set of project documentation. Sidewalk Labs have withdrawn from the project as of 7 May 2020. From the outset, data governance and privacy had raised concerns of public and privacy experts, and eventually became threshold issues for deciding whether the project will move forward at all. Even if Sidewalk Labs' official statement quotes the "*unprecedented economic uncertainty*" around Covid-19 as the reason for withdrawal[179], the interested public has not failed to note that the relative lack of Sidewalk Labs' experience with urban development and its bullish and secretive approach to data governance issues caused the project to become "*an obvious mess*"[180]. Prior to the withdrawal, one of the key pillars for reaching an agreement on data governance and privacy management was the envisaged adherence by all stakeholders (including Sidewalk Labs) to emerging but not yet published "*Intelligent Community Guidelines*", a set of rules combining input from government stakeholders, industry and the broader community on digital governance issues and privacy, which was supposed to be enforceable against private parties (including Sidewalk Labs) through contract.

In this Annex, we provide a high level overview of data governance and privacy issues that have emerged in the Sidewalk Toronto project documentation made publicly available with relevant excerpts or our brief commentary. Even though regulatory frameworks, including but not limited to personal data protection laws, may be to a various degrees different in Canada than they are in the EU or individual Member States, we suggest that the Sidewalk Toronto project is monitored closely by DUET to anticipate current and future trends in Smart Cities approach to, among others, data governance and privacy issues, and that lessons are learnt from Sidewalk Labs' failure. The further following country/city-level projects could similarly be explored for inspiration: Estonia (X-Road; Digital ID), Montreal (AI ethics; integrated mobility, CivicInnovation Lab for Regulatory Testing), New York City (Guidelines for IoT, the AutomatedDecision Task Force; Open Data);

[176] https://www.sidewalktoronto.ca/ .
[177] https://www.sidewalklabs.com/.
[178] https://www.waterfrontoronto.ca/nbe/portal/waterfront/Home.
[179] https://medium.com/sidewalk-talk/why-were-no-longer-pursuing-the-quayside-project-and-what-s-next-for-sidewalk-labs-9a61de3fee3a .
[180] https://theconversation.com/sidewalk-labs-smart-city-plans-for-toronto-are-dead-whats-next-138175, or https://blaynehaggart.com/2020/05/08/no-longer-liveblogging-sidewalk-labs-midp-entry-49-a-letter-from-sidewalk-labs-ceo-dan-doctoroff/ .

Chicago (Tech Plan; Array of Things), Amsterdam (Data Sharing; IoT Registry; Data Exchange; TADA Manifesto), Barcelona (Ethical Digital Standards; Barcelona DigitalCity; Decidim), and others.

Table 3 - Matrix of data governance and privacy topics

Issue	Sub issue	Comments/excerpts from Sidewalks Toronto documents or other relevant sources
Data categories	Consent data ("transaction data")	How to gather data subject consent in public space. User Agreements: Where a digital service relies on a user agreement, how will information related to data collection and use be communicated? How will user agreements related to in-home services – 'pay-as-you-go' waste disposal, unit-level energy monitoring, etc. – be enforced, and what is the consequence for non-compliance? Can users lose access to these services?
	Non-consent data ("urban data")	Suggestion to invent a new type of data, dubbed "urban data," that sits outside the existing regulations of the Canadian government. Sidewalks Labs suggested that these data are managed by an independent trust (Urban Data Trust). "*Urban data would be broader than the definition of personal information and include personal, non-personal, aggregate, or de-identified data collected and used in physical or community spaces where meaningful consent prior to collection and use is hard, if not impossible, to obtain*," states the planning documents. "I*n that sense, urban data would be distinct from more traditional forms of data, termed here 'transaction data,' in which individuals affirmatively — albeit with varying levels of understanding — provide information about themselves through websites, mobile phones, or paper documents.*" (https://skift.com/2019/07/03/google-parent-alphabets-smart-city-vision-in-toronto-poses-privacy-concerns/) After criticism, Sidewalk Labs agreed that all personal information will be stored in Canada, and it has eliminated the Urban Data Trust proposal, as well as the term 'urban data'. It will comply with all existing and future legislative and regulatory frameworks. Note: 'urban data' term now politically charged, we suggest avoiding its use for DUET's purposes.

	Data types	Data types need to be identified to justify their collection and processing. E.g. (taken from Sidewalks Toronto project, mobility issues) location of streetcars, current availability of curb space, volume of pedestrians, cyclists, vehicles, how long do people have to wait to cross the street?, real-time alerts to autonomous vehicles of pedestrians and cyclists around the corner. The four types of data collected by the proposed services: **Non-personal data** is data that does not identify an individual and can include other types of non-identifying data that is not about people. Some examples of non-personal data are aggregated data sets, machine-generated data (such as weather and temperature data), or data on maintenance needs for industrial machines. **Aggregate data** is data that is about people in the aggregate and not about a particular individual. Aggregate-level data is useful for answering research questions about populations or groups of people. For example, aggregate counts of people in an office space can be used in combination with other data, such as weather data, to create an energy-efficiency program so consumption is controlled, with the goal of saving money and reducing energy use. **De-identified data** is data about an individual that was identifiable when collected but has subsequently been made non-identifiable. Third-party apps and services may wish to use properly de-identified data for research purposes, such as comparing neighbourhood energy usage across a city. **Personal information** has a legal definition in Canada and is the subject of privacy laws, including the Personal Information Protection and Electronic Documents Act (PIPEDA). The broad legal definition of personal information includes any information that could be used, alone or in combination with other information, to identify an individual or that is associated with an identifiable individual.
	Anonymisation/pseudonymisation	Some Panelists felt that Sidewalk overly relies on de-identification at source as a sufficient basis for making personal data open for re-use. While de-identification can help protect personal information, a panelist noted that it does not remove it from Canada's data protection regimes or put it beyond the oversight of a privacy commissioner (though another disputed that assertion).

Legal basis for processing	Public benefit considerations	Public benefits need to be identified to help justify data collection and processing. E.g. (taken from Sidewalks Toronto project, mobility issues), Reduce congestion, Increase safety, Adapt to new forms of transport from self-driving cars to scooters, Personal Rapid Transit, Responsive Traffic Signals

	'opt-in customer analytics'	https://skift.com/2019/07/03/google-parent-alphabets-smart-city-vision-in-toronto-poses-privacy-concerns/
	Privacy by design / Safe sharing sites	De-identification (anonymization) at source. Obtain consent otherwise where personal information is needed to provide the service. When data collection is minimized and 'de-identified' (the process of anonymizing data) at source, the risk of privacy violations can be dramatically reduced. And when no personal information is collected, consent is not required. https://qz.com/1756852/smart-cities-will-always-have-a-data-privacy-problem/ The trick they propose is to encrypt personal information in a way that preserves the ability to run queries on the encrypted data. Analysts can ask questions that link together personal data, but they only ever see anonymized, aggregated results. All the questions and answers are recorded, creating an audit trail that allows regulators and courts to inspect how the data has been used and to penalize misuse.
Data ownership	Primary data ownership / processors	
	Independent governance	E.g. independent trust proposed in Sidewalks Toronto. Rejected by public interest stakeholders. DIA Reference: p. 5] The DIA states "*the clear feedback was that a new standalone entity for these functions was not a preferred path for this project*". This is misleading. The main objections to the UDT were that SWL was setting the terms when this should be done by WT, and that these terms did not comport well with public interest requirements. The idea of a new digital stewardship body still has much merit, and should be part of the digital governance discussion. As noted, working this out will take considerable time.
	IP rights	The parties recognize that data ownership is an issue that will be resolved through implementation agreements and that prevailing ownership and partnership models will serve as baseline expectations for the City of Toronto, Province of Ontario, and Government of Canada.

IP rights	Attribution of acquired IP rights may be difficult in complex cooperative structures (such as public-private partnerships)	Waterfront Toronto and Sidewalk Labs are committed to working in good faith to design an intellectual property framework that not only recognizes the value of Waterfront Toronto's contribution to catalyzing innovation but also creates a foundation for Canadian-based companies to innovate in Canada and compete on a global scale. This includes but is not limited to: - A revenue stream on products and services piloted - An expanded patent pledge that allows innovators to leverage Sidewalk Labs' hardware and software digital innovation patents. The patent pledge will provide Canadian innovators operating globally with the right to use all Sidewalk Labs' Canadian and foreign patents covering hardware and software digital innovations. - Resolve issues relating to the role and obligations of Sidewalk Labs' affiliates that Sidewalk Labs has working in a Waterfront Toronto-facilitated testbed area on their associated products and services, including without limitation, in relation to revenue share arrangements and remedies in the event of default. - - Sidewalk Labs will provide Waterfront Toronto with an irrevocable, perpetual license to use the Site-Specific IP. - A mechanism that will provide appropriate recognition for Waterfront Toronto's contributions to co-created IP. - A reporting and audit structure which is transparent and manageable. Waterfront Toronto will work with Sidewalk Labs and other innovators to provide meaningful support and enable the testing, piloting, and development of products and services that serve the Innovation Plan. Given that the new services will necessarily create new IP the ability to exploit that IP becomes critical. Whoever is procuring services, there needs to be an arrangement that enables contracted companies to continue to exploit that IP without restriction, so that it is not just Sidewalk or WT or other public sector actors who commissioned the work that benefit from the IP that has been created.

Proportionality of collecting and processing (data minimisation)		Sidewalk Labs has committed to collecting the minimum amount of data needed to achieve the beneficial purpose of proposed services, and to using the least invasive technology available to achieve the beneficial purpose. The Digitally Enabled Services List provides early information on the anticipated data to be collected and the technologies used for collection, based on the planning performed to date. It was flagged that the first question in Sidewalk Labs' RDUA relates to "*beneficial purpose*" – that is, whether there is a clear purpose and value to any proposed collection or use of data. However, this is only one side of the equation – a beneficial purpose must be weighed against potential or known negative impacts. It was recommended, then, that a necessity and proportionality test might be a more appropriate starting point for the RDUA (and/or the Intelligent Community Guidelines) [noting that these are two of the four elements of the privacy regulators in considering the appropriateness of a technology or service - the other two being Effectiveness and Minimization]. Panelists were split on whether a digital governance framework (e.g. the Intelligent Community Guidelines) should include specific "no-go zones" (such as an outright ban on facial recognition and other forms of biometric capture), or whether a necessity and proportionality test would be effective while allowing for individual choice and/or democratic decision-making – particularly for technologies that impact individuals, rather than whole populations, and assuming appropriate transparency. The first question in the RDUA relates to beneficial purposes. It asks whether there is a "*clear purpose and value to any proposed use of data*" as well as a clear connection to benefits to individuals or the community. However, as I noted with some of my comments about specific technologies in the previous sections, while there may be a clear purpose and value to proposed uses of data, there may also be negative impacts and effects that outweigh these values. I wonder whether a necessity/proportionality test might be a better starting point. Certainly, in the public sector context (when we are talking about, for example, data collection for public services/programs) necessity and proportionality are guiding considerations. Do the benefits of suite level electricity metering outweigh the potential harms? In my view, this is a better question to ask.
Recognition of the limitations of the data used		
Incomplete or inaccurate data, rectification mechanisms		

Precautionary approach		What happens if the digital innovations don't work? What is the revert-to-normal plan, and what is "normal" in an advanced community? Again, this will have to be made clear for each proposed innovation.
Transparency and accountability in collection, processing, and accessing data /		Minimum Technology Used - standardised icons on devices/buildings. Records of processing activities. Perhaps, SWL should create a digitally enabled service that allows individuals to use their phone to point at a sensor and get information about that sensor immediately. I believe SWL suggested scanning a QR code for the same purpose elsewhere in the document (pg 315). We should also create an IoT registry. If only for the reasons that Amsterdam created it: "*to eliminate the duplication of data collection [and sensor clutter] and provide a back door to data sharing among entrepreneurs*".
Repeat use / Use for a different purpose / secondary use		The 'RDUA in practice' document (pp. 251-269) demonstrates how the section on 'secondary purposes' operates. The party proposing to collect the data must indicate the purpose of collection as well as any secondary purposes. In the example provided, it states that "T*here are no secondary purposes with respect to the data collected in this pilot*". But the same document indicates that the collected data "*will be made publicly available in some format*". So, while SWL may not use the data for secondary purposes, the data will be available to others to use for who knows what purposes. It seems to me that the 'secondary purposes' category of evaluation is meaningless (and/or misleading) if the data will be shared with others, as open data or otherwise, for other purposes. If the data is shared through some form of data governance body, this might be addressed in that process, but if it is made available as open data, then it is open to all manner of secondary uses.
Combining and merging of data		I note that the entire RDUA and guidance document seem to assume that all data used will be collected directly from individuals. I would assume that in some cases, AI will be trained on data acquired in other contexts and from other sources, or that data collected directly from individuals may be combined with data acquired from other sources. Is there some process for assessing the quality/suitability/ethical nature of data acquired from other sources?

Localization/storage/retention		With respect to the operations of digitally enabled solutions in Quayside, Sidewalk Labs agreed (i) that personal information will be stored and processed in Canada; and (ii) to use commercially reasonable efforts to store and process non-personal data in Canada. Should exceptions be required, they will be determined on a case-by-case basis through a review process. As in the Preliminary Commentary, the issue of data localization was raised. One panelist clarified that the principle of Canadian data residency should include not just storage but also transmission, as data which transmits through the US is subject to NSA surveillance. Another recommended that the mission criticality of data should factor into any decision in which a lack of redundancy forces non-personal data to be stored outside of Canada; for example, Sidewalk Labs' Numina pilot – which involves 3-cameras measuring movement of de-identified individuals within Sidewalk's 307 Lakeshore exhibit space – would not seem to be negatively impacted by a brief loss of data in the event of a region failure, but nonetheless this was the reason given that data was stored outside of Canada. This was considered to be a concerning precedent, which should be addressed within the digital governance framework as it is developed.
Data sharing		Sidewalk Labs' commitments: 1. Sidewalk Labs will not sell personal information. 2. Sidewalk Labs will not use personal information for advertising. 3. Sidewalk Labs will not share personal information with third parties, including other Alphabet companies, without explicit consent. (NOTE this does not cover non-personal information) Cross-sectoral data sharing is part of the discussion in this section as well, and it is proposed that a hub for data collaboration be created. While data sharing remains an important part of the overall proposal, the plans here are very vague and general (not really surprising given the need to back away from the Urban Data Trust). While Waterfront Toronto has indicated that it will play more of a role in relation to data governance for data sharing, these details need to be worked out - and this is not something that can be left to the last minute. The data collaboration hub being proposed is an interesting idea - in many ways it is smaller scale and more modular than the Urban Data Trust, and it also seems more oriented towards private rather than public sector data (it does not resolve the public/private sector issues relating to data collected within the development). It could allow for the development of smaller-scale, case-specific forms of data sharing; data sharing between specific entities rather than more global data sharing; and even more general data sharing. The complexity of developing data governance for data sharing means that it might be more manageable to proceed in this way than to create a large, overarching and all-inclusive infrastructure for data sharing - but this needs thought and discussion. If it is experimental, case-specific, and

		not clearly mandated, it might also not amount to much. So there are interesting ideas here, but they need to be further developed. One of the ideas from the previous concept of 'urban data' and the UDT was that data about urban residents collected from 'public' spaces was data in which the broader community had an interest, and therefore it should be governed in the public interest. This concept is somewhat lost in the discussion of the data collaboration hub. Part of the challenge with 'urban data' was the role of the public sector in relation to the governance of data in which there is a strong public interest. The role of the public sector still needs to be clarified in this regard.
Data monetization	Use of data for advertising etc.	Sidewalk Labs has pledged not to sell advertisers the personal data collected to serve residents and visitors. Privacy activists have insisted that Sidewalk Labs must guarantee that personal data used to run the project remains anonymous.
Open data and access	Access to data, also regarding law enforcement authorities' access to data collected by Smart Cities	E.g. New York's use of global positioning systems (GPS) on buses to prioritize public transport at intersections. As has been stated elsewhere public security agency and police access to data are not discussed at all in the document. This should be a matter for Waterfront Toronto to document in its Intelligent Community Guidelines for SWL and all others to conform to. The assumption seems to be that any Police surveillance systems such as red-light cameras or video cameras are separate systems. It is not clear if they could take advantage of proposed infrastructure such as Koala if they chose to do so. Publicly available - SWL indicates that it is committed to non-personally identifiable data being publicly accessible by default. I would like some more clarification of this. Does it mean open data? Or data governed by an entity set up to oversee data sharing (or one or the other depending on the circumstances). When is open data appropriate? When is more controlled sharing appropriate? I realize that some of this might fall to be determined by the data governance scheme that Waterfront may now be committed to developing - but these are important questions and there should be some sense of the answers going into this project. If personal information is not made available by default, what about aggregate or de-identified data? What protocols will be in place to ensure that data is properly deidentified? Again, this may be for the data governance body, but this highlights the need for movement in this area. I note that this commitment talks about publicly accessible by default with the exclusion of personal data - but presumably confidential commercial information will also be excluded. How will this be determined? Who gets to decide what information is commercially sensitive or confidential? Does anyone get to review such determinations? This is potentially a huge loophole in the data sharing

		commitment. What role, if any, will any data governance body have in overseeing decisions about what data should be shared?

Combining and merging of data		I note that the entire RDUA and guidance document seem to assume that all data used will be collected directly from individuals. I would assume that in some cases, AI will be trained on data acquired in other contexts and from other sources, or that data collected directly from individuals may be combined with data acquired from other sources. Is there some process for assessing the quality/suitability/ethical nature of data acquired from other sources?
Localization/storage/retention		With respect to the operations of digitally enabled solutions in Quayside, Sidewalk Labs agreed (i) that personal information will be stored and processed in Canada; and (ii) to use commercially reasonable efforts to store and process non-personal data in Canada. Should exceptions be required, they will be determined on a case-by-case basis through a review process. As in the Preliminary Commentary, the issue of data localization was raised. One panelist clarified that the principle of Canadian data residency should include not just storage but also transmission, as data which transmits through the US is subject to NSA surveillance. Another recommended that the mission criticality of data should factor into any decision in which a lack of redundancy forces non-personal data to be stored outside of Canada; for example, Sidewalk Labs' Numina pilot – which involves 3-cameras measuring movement of de-identified individuals within Sidewalk's 307 Lakeshore exhibit space – would not seem to be negatively impacted by a brief loss of data in the event of a region failure, but nonetheless this was the reason given that data was stored outside of Canada. This was considered to be a concerning precedent, which should be addressed within the digital governance framework as it is developed.

Data sharing		Sidewalk Labs' commitments: 1. Sidewalk Labs will not sell personal information. 2. Sidewalk Labs will not use personal information for advertising. 3. Sidewalk Labs will not share personal information with third parties, including other Alphabet companies, without explicit consent. (NOTE this does not cover non-personal information) Cross-sectoral data sharing is part of the discussion in this section as well, and it is proposed that a hub for data collaboration be created. While data sharing remains an important part of the overall proposal, the plans here are very vague and general (not really surprising given the need to back away from the Urban Data Trust). While Waterfront Toronto has indicated that it will play more of a role in relation to data governance for data sharing, these details need to be worked out - and this is not something that can be left to the last minute. The data collaboration hub being proposed is an interesting idea - in many ways it is smaller scale and more modular than the Urban Data Trust, and it also seems more oriented towards private rather than public sector data (it does not resolve the public/private sector issues relating to data collected within the development). It could allow for the development of smaller-scale, case-specific forms of data sharing; data sharing between specific entities rather than more global data sharing; and even more general data sharing. The complexity of developing data governance for data sharing means that it might be more manageable to proceed in this way than to create a large, overarching and all-inclusive infrastructure for data sharing - but this needs thought and discussion. If it is experimental, case-specific, and not clearly mandated, it might also not amount to much. So there are interesting ideas here, but they need to be further developed. One of the ideas from the previous concept of 'urban data' and the UDT was that data about urban residents collected from 'public' spaces was data in which the broader community had an interest, and therefore it should be governed in the public interest. This concept is somewhat lost in the discussion of the data collaboration hub. Part of the challenge with 'urban data' was the role of the public sector in relation to the governance of data in which there is a strong public interest. The role of the public sector still needs to be clarified in this regard.
Data monetization	Use of data for advertising etc.	Sidewalk Labs has pledged not to sell advertisers the personal data collected to serve residents and visitors. Privacy activists have insisted that Sidewalk Labs must guarantee that personal data used to run the project remains anonymous.

Open data and access	Access to data, also regarding law enforcement authorities' access to data collected by Smart Cities	E.g. New York's use of global positioning systems (GPS) on buses to prioritize public transport at intersections. As has been stated elsewhere public security agency and police access to data are not discussed at all in the document. This should be a matter for Waterfront Toronto to document in its Intelligent Community Guidelines for SWL and all others to conform to. The assumption seems to be that any Police surveillance systems such as red-light cameras or video cameras are separate systems. It is not clear if they could take advantage of proposed infrastructure such as Koala if they chose to do so. Publicly available - SWL indicates that it is committed to non-personally identifiable data being publicly accessible by default. I would like some more clarification of this. Does it mean open data? Or data governed by an entity set up to oversee data sharing (or one or the other depending on the circumstances). When is open data appropriate? When is more controlled sharing appropriate? I realize that some of this might fall to be determined by the data governance scheme that Waterfront may now be committed to developing - but these are important questions and there should be some sense of the answers going into this project. If personal information is not made available by default, what about aggregate or de-identified data? What protocols will be in place to ensure that data is properly deidentified? Again, this may be for the data governance body, but this highlights the need for movement in this area. I note that this commitment talks about publicly accessible by default with the exclusion of personal data - but presumably confidential commercial information will also be excluded. How will this be determined? Who gets to decide what information is commercially sensitive or confidential? Does anyone get to review such determinations? This is potentially a huge loophole in the data sharing commitment. What role, if any, will any data governance body have in overseeing decisions about what data should be shared?

Data breaches/Security / Data impact assessment		Sidewalk Labs applies best practices to prevent network and data breaches before they occur and also recognizes its obligations under the Personal Information Protection and Electronic Documents Act (PIPEDA) to maintain appropriate safeguards that include physical, organizational, and technical measures to ensure the security of networks and data that it controls. These measures include: • Implementing internationally recognized information security standards, such as the ISO 27000 series of standards. • Conducting Threat Risk Assessments/Vulnerability Assessments and penetration testing. • Developing, implementing, and maintaining an information security program to proactively assess risks and implement safeguards. • Rigorously updating and patching operating systems, firmware, and software. • Continuous monitoring for unusual network activity. • Physical measures to limit physical access to digital infrastructure. • Administrative measures to limit system and data access. • Security procedures and regular training. • End-to-end encryption, as applicable. • Contractual requirements with vendors that provide appropriate safeguards consistent with those above, and notification of network or data breaches. Additionally, Sidewalk Labs believes in applying best practices to address any network or data breaches that may occur, including having a cyber-incident response plan in place, which includes: • Detecting incidents and escalating to the appropriate level within the organization. • Investigating the characteristics of an incident and its impact. • Containing the scope and severity of incidents. • Coordinating and managing recovery activities. • Assessing and managing risks. • Preserving information associated with the incident, as appropriate. • Providing notification to insurers, affected individuals, affected third parties, and authorities, as applicable. • Analyzing the incident after the fact to prevent future incidents. In the event of a network or data breach, Sidewalk Labs will diligently execute requirements under PIPEDA, other applicable legislation, and contractual obligations. This includes: • Reporting breaches of security safeguards involving personal information that Sidewalk Labs controls to the Privacy Commissioner of Canada when the breach involves personal information and it is reasonable in the circumstances to believe that the breach creates a real risk of significant harm to individuals. • Notifying affected individuals about such breaches and notifying any third parties that may be able to reduce or mitigate harm, such as other organizations or government agencies. • Maintaining records of all breaches.

Algorithms / issue of algorithmic bias /AI		Responsible AI. Development of algorithms according to principles that include fairness, accountability, inclusiveness, and reduced bias. Sidewalk Labs agrees to work with Waterfront Toronto and its government stakeholders in good faith to ensure each digitally enabled solution will not impede (and where feasible, will foster) accessibility in Quayside, freedom of association, freedom of expression, equitable treatment of marginalized groups, public engagement and participation and other fundamental rights and freedoms, as applicable. To assess the benefits and risks of smart designs, we need to go beyond what data is being gathered for what purpose, and how the data is secured. We need to know what 'smart' decisions are being made, based on what data, integrating what assumptions, and by whom. The criterion for replacing human decision-making with AI-based decision-making should not be based purely on equivalent percentage of accuracy, but also the pattern of accuracy. How accurate is the decision when it comes to minorities and outliers? Sidewalk Labs' responsible AI framework is guided by six overarching principles that are contextual, progressive, and technology neutral. 1. Fairness and equity: All projects involving AI systems should be designed and developed responsibly from the start and should consider an individual's reasonable expectations and the original purposes of data collection. 2. Accountability: Sidewalk Labs commits to completing RDUAs for all projects and products that involve AI and compiling an archive of all automated decision systems. 3. Transparency and explainability: Individuals should be informed when they are interacting directly with an automated system and when their personal data is being used to make consequential decisions about them. All systems should be designed with the ability to explain and debug their output in terms people can understand. 4. Relevance: Sidewalk Labs commits to high standards of scientific excellence and a multidisciplinary approach that includes sharing research and best practices with regard to AI. 5. Value alignment: AI systems should be designed, developed, and used in line with international human rights and local community values. 6. Respect for human dignity and safety: Individual autonomy and agency should be upheld through a diverse and multidisciplinary design process. AI systems should be used to empower individuals and communities and enhance public engagement.

Privacy legislation	How to approach the fact that privacy and data governance legislation is constantly evolving	Sidewalk Labs reaffirms its commitment to comply with all existing and future privacy legislation, regulations and policy frameworks (e.g., Canada's Digital Charter and Ontario Digital Principles). This includes an understanding that data governance, in particular, personal information, varies for public and private activities and actors. Data governance will be determined by the municipal, provincial and federal laws applicable to access and protection of data in the Project. These laws apply to Sidewalk Labs as they do to any private sector organization. Sidewalk Labs will not condition implementation agreements on the requirement for new or amended privacy laws or other new laws or regulations in order to achieve a digital governance structure. This includes removing the expectation for the creation of the proposed 'Urban Data Trust'.
Trust in tech firms/suppliers – general questions		Sidewalk Labs has pledged not to sell advertisers the personal data collected to serve residents and visitors. Privacy activists have insisted that Sidewalk Labs must guarantee that personal data used to run the project remains anonymous. CEO Dan Doctoroff said that Sidewalk Labs will not disclose personal information to third parties without explicit consent and will not sell personal information.https://www.businessinsider.com/alphabet-commits-to-data-privacy-in-toronto-smart-city-master-plan-2019-6
Self-regulated checks on tech firms/suppliers	Form of self-regulated pledges/guarantees In which documents they should appear (cooperation agreements etc.) Enforcement Approved codes of conduct	At the core of the Sidewalk Labs project is the goal of tracking people without their consent. While the data will be protected as a 'public asset', Alphabet will still have access to all this data because its systems, of course, will be collecting it. The company calls for a public data trust to collect the information that is collected, but Sidewalk Labs' systems will be interlocked with almost every aspect of not just infrastructure but city life. Sidewalk Labs has already committed publicly that it would not sell personal information to third parties or use it for advertising purposes. It also commits to not share personal information with third parties, including other Alphabet companies, without explicit consent. https://skift.com/2019/07/03/google-parent-alphabets-smart-city-vision-in-toronto-poses-privacy-concerns/ Sidewalk Labs have agreed to adhere to Waterfront Toronto's Digital Principles (found on Waterfront Toronto's website – www.waterfrontoronto.ca), which have been developed through consultation with industry, academia, government stakeholders, and the broader community. Emerging Intelligent Community Guidelines will be similarly enforced through contract.

		https://waterfrontoronto.ca/nbe/wcm/connect/waterfront/86d92f81-20be-4029-a616-00522abbd34a/Threshold+Issues+Resolution+Documents.pdf?MOD=AJPERES Sidewalk Labs has created a set of Responsible Data Use Guidelines to address data ethics, access to information, and the ways that aggregate or de-identified data can impact individuals and groups of people through the use of advanced analytics such as artificial intelligence. This is in addition to the areas covered by standard tools such as a Privacy Impact Assessment.
Community engagement		Sidewalk Labs has agreed that there will be further documented, facilitated consultation with community stakeholders, with an emphasis on ensuring engagement with groups most impacted by a particular technology, during the development process. Digital proposals may be required to go through a public meeting process and approval by governments Tantoco also said that community engagement was crucial in getting ideas across to the public. "[It's] absolutely, absolutely essential, especially if you're putting technology out on the city streets. People are very concerned about privacy and security," she said. https://skift.com/2017/06/15/smart-cities-need-open-data-and-a-willingness-to-test-and-learn/
Establishment/engagement of advisory bodies		The Digital Strategy Advisory Panel has been established to provide peer review and advice to Waterfront Toronto digital proposals, including privacy concerns. https://waterfrontoronto.ca/nbe/wcm/connect/waterfront/7a14158a-be29-4e60-a420-b99fcb2d4382/20180427+Digital_Strategy_Advisory_Panel_Mandate.pdf?MOD=AJPERES
Enforcement		Emerging Intelligent Community Guidelines, comprising rules on data governance and privacy, will be enforceable through contract. Enforceability and remedies for breach are not easy to develop for an Alphabet company. Certainly, protections against corporate restructuring and parental guarantees need to be in place. Additionally, traditional financial incentives / penalties may not be enough. As far as I can tell Google's total EU antitrust bill now stands at €8.2 billion. There has been some impact for these fines as this article reports. Perhaps personal liability of SWL officers is an approach to be considered.