COMPUTER LAW & SECURITY REVIEW 40 (2021) 105486

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/CLSR

Computer Law & Security Review

Data protection law beyond identifiability? Atmospheric profiles, nudging and the Stratumseind Living Lab

Maša Galič[a,*], *Raphaël Gellert*[b]

[a] *Faculty of Law, VU University Amsterdam, the Netherlands*
[b] *Radboud Business Law Institute, and Interdisciplinary Hub for Security, Privacy, and Data Governance (iHub), Radboud University, the Netherlands*

ARTICLE INFO

Keywords:
Data protection
Personal data
Smart city
Profiling
Nudging
Stratumseind

ABSTRACT

The deployment of pervasive information and communication technologies (ICTs) within smart city initiatives transforms cities into extraordinary apparatuses of data capture. ICTs such as smart cameras, sound sensors and lighting technology are trying to infer and affect persons' interests, preferences, emotional states, and behaviour. It should be no surprise then that contemporary legal and policy debates on privacy in smart cities are dominated by a debate focused on data and, therefore, on data protection law. In other words, data protection law is the go-to legal framework to regulate data processing activities within smart cities and similar initiatives. While this may seem obvious, a number of important hurdles might prevent data protection law to be (successfully) applied to such initiatives. In this contribution, we examine one such hurdle: whether the data processed in the context of smart cities actually qualifies as personal data, thus falling within the scope of data protection law. This question is explored not only through a theoretical discussion but also by taking an illustrative example of a smart city-type initiative – the Stratumseind 2.0 project and its living lab in the Netherlands (the Stratumseind Living Lab; SLL). Our analysis shows that the requirement of 'identifiability' might be difficult to satisfy in the SLL and similar initiatives. This is so for two main reasons. First, a large amount of the data at stake do not qualify as personal data, at least at first blush. Most of it relates to the environment, such as, data about the weather, air quality, sound and crowding levels, rather than to identified or even likely identifiable individuals. This is connected to the second reason, according to which, the aim of many smart city initiatives (including the SLL) is not to identify and target specific individuals but to manage or nudge them as a multiplicity – a combination of the environment, persons and all of their interactions. This is done by trying to affect the 'atmosphere' on the street. We thus argue that a novel type of profiling operations is at stake; rather than relying on individual or group profiling, the SLL and similar initiatives rely upon what we

* Corresponding author: Maša Galič, Faculty of Law, VU University Amsterdam,
E-mail addresses: m.galic@vu.nl (M. Galič), r.gellert@jur.ru.nl (R. Gellert).

https://doi.org/10.1016/j.clsr.2020.105486

have called 'atmospheric profiling'. We conclude that it remains highly uncertain, whether smart city initiatives like the SLL actually process personal data. Yet, they still pose risks for a wide variety of rights and freedoms, which data protection law is meant to protect, and a need for regulation remains.

1. Introduction

Smart city initiatives – generally referring to the extensive embedding of software-enabled technologies into the city environment – are by now an obligatory feature of almost any city or town in the developed world, including Europe. The contemporary city is increasingly built of pervasive information and communication technologies (ICTs), including digital cameras, sound, smell and other sensors, adaptive lighting, and wifi tracking technology. Within the smart city discourse, such technologies serve two broad purposes: improving urban management and the quality of life in the city (such as, lowering gas emissions, traffic congestion and crime) and stimulating the economic development of the city.[1] However, beyond the grand promises of the smart city discourse, data-driven monitoring practices of ICTs transform cities into extraordinary apparatuses of data capture, trying to infer and affect persons' interests, preferences, emotional states, and behaviour. People are thus subjected to much greater levels of scrutiny and control as increasing aspects of their daily lives are captured as data.[2]

It should be no surprise then that contemporary legal and policy debates on privacy in smart cities and similar initiatives, such as living labs, are dominated by a debate focused on data and, therefore, on data protection law.[3] In other words, data protection law is the *go-to* legal framework to regulate data processing activities within smart cities and living labs. While data protection law seems to be the obvious legal framework to regulate smart cities, it might actually have serious trouble with regulating smart city-type initiatives due to several normative and practical arguments. For instance, Hildebrandt has argued that data protection law is not equipped to deal with real-time and continuous profiling in ambient environments (which smart cities and living labs aspire to be), as it does not offer real-time protection.[4] Furthermore, smart cities are commonly formed in complex public-private partnerships (PPPs) including numerous projects, goals and actors, where data is processed both for commercial and law enforcement purposes. Yet, the dual regime of the General Data Protection Regulation and the Law Enforcement Directive has been said to result in a muddled relationship, complicating the attribution of responsibility and thus hindering the capacity of data protection law to effectively regulate data processing within such partnerships.[5] Finally, following the advances in data-driven technology, there has recently been a renewed academic interest in the question of what constitutes personal data. Some scholars have convincingly argued that the material scope of EU data protection law has grown so broad that 'literally any data can plausibly be argued to be personal'.[6] Others have pointed to the lack of clarity and the inconsistency surrounding the notion of anonymous data, and hence the extremely uncertain and probabilistic nature of the concept of identifiability.[7] As Edwards put it, what constitutes personal data is still one of the central causes of doubt in the current data protection regime,[8] an issue that is exacerbated in the smart city context.

In this contribution, we further these discussions and critiques of data protection law by exploring whether and to what extent data protection law could apply to smart city-types of initiatives, particularly those that include a focus on safety and public order types of issues. In fact, public safety and security are often a key driver when considering investment in particular smart city technologies, since *safe*

[1] Rob Kitchin, 'The promise and perils of smart cities' (2015) *Society for Computers Law*; 'Barcelona Smart City'; 'Amsterdam Smart City'; European Commission, 'Smart Cities: Cities Using Technological Solutions to Improve the Management and Efficiency of the Urban Environment'.

[2] Rob Kitchin, 'The Ethics of Smart Cities and Urban Science' (2016) 374 Philosophical Transactions A.

[3] Lorenzo Dalla Corte, Bastiaan van Loenen and Colette Cuijpers, 'Personal Data Protection as a Nonfunctional Requirement in the Smart City's Development', *13th International conference on internet, law & politics, Universitat Oberta de Catalunya, Barcelona, 29-30 June 2017* (2017); Lilian Edwards, 'Privacy, Security and Data Protection in Smart Cities: A Critical EU Law Perspective' (2016) 2 European data protection law review 28.

[4] Mireille Hildebrandt, 'A Vision of Ambient Law' in Roger Brownsword and Karen Yeung (eds), *Regulating Technologies* (Bloomsbury 2008); Mireille Hildebrandt and Bert-Jaap Koops, 'The Challenges of Ambient Law and Legal Protection in the Profiling Era' (2010) 73 Modern Law Review 428.

[5] Nadezhda Purtova, 'Between the GDPR and the Police Directive: Navigating through the Maze of Information Sharing in Public-Private Partnerships' (2018) 8 International Data Privacy Law 52; Orla Lynskey, 'Criminal Justice Profiling and EU Data Protection Law: Precarious Protection from Predictive Policing' (2019) 15 International Journal of Law in Context 162; Catherine Jasserand, 'Law Enforcement Access to Personal Data Originally Collected by Private Parties: Missing Data Subjects' Safeguards in Directive 2016/680?' (2018) 34 Computer Law and Security Review 154.

[6] Nadezhda Purtova, 'The Law of Everything. Broad Concept of Personal Data and Future of EU Data Protection Law' (2018) 10 Law, Innovation and Technology 40.

[7] Michèle Finck and Frank Pallas, 'They Who Must Not Be Identified — Distinguishing Personal from Non-Personal Data under the GDPR' (2020) 10 International Data Privacy Law 1.

[8] Lilian Edwards, 'Data Protection: Enter the General Data Protection Regulation' in Lilian Edwards (ed), *Law, Policy and the Internet* (Hart Publishing 2018).

(smart) cities are worth investing in.[9] We thus ask the question, whether the data processed in this context could qualify as personal so as to trigger the application of data protection law. We posit that it remains highly uncertain, whether a large part of such smart city operations could actually fall within the scope of the data protection. This is so for two main reasons. First, a large amount of the data collected within smart city initiatives does not qualify as personal data, at least at first blush. Most of it relates to the environment, such as, data about the weather, air quality, sound and crowding levels, rather than to identified or (likely) identifiable individuals.[10] Moreover, the aim of many smart city initiatives, particularly in Europe, is not to identify and target specific individuals but to manage or nudge them as a multiplicity – a combination of the environment, persons and all of their interactions.[11] Following real-time algorithmic decisions, smart environments namely alter themselves so as to indirectly influence people's behaviour.[12] Consequently, persons do not need to be identified or identifiable at all.

In order to substantiate these theoretical claims, we offer a context-specific analysis of the possibilities for application of EU data protection law to a concrete example of a smarty city-type initiative with a focus on safety and security – the *Stratumseind 2.0* project and its living lab (*Stratumseind Living Lab; SLL*) in the Netherlands. Touted as a success by national and supra-national authorities,[13] this longer-lasting mid-sized project, which includes both multinational and local technology companies, is an illustrative example of a European smart city-type initiative. Despite all of the scholarly attention, the concept of personal data has as yet not been studied in connection to a concrete example of a smart city initiative.[14] Given the increasing role that smart cities play in shaping our experience in urban public space – including limiting our privacy and other fundamental rights and freedoms – there is a pressing need for exposing the contours of the concept of personal data in this particular context.

The structure of the paper is the following. First, we describe the *Stratumseind 2.0* initiative, its living lab (*SLL*), and two of its sub-projects relating to predictive policing and nudging: *CityPulse* and *De-escalate*. In the third part, we address the scope of personal data by focusing on the two key elements of its definition: 'relating to' and 'identifiability'. We explore the meaning of information 'relating to' by applying it to the *SLL* example. We then investigate the notion of identifiability from a legal and technical perspective, followed by an analysis of the notion of identifiability through the lens of profiling. Based on this discussion, the last section examines the type of profiling and nudging found in the *SLL*, which we refer to as 'atmospheric profiling'. We conclude that it remains difficult to give a clear-cut answer as to whether the *SLL* example processes data, particularly in relation to the goal of affecting the atmosphere. Even more so, the notion of atmospheric profiling puts an additional strain to the concept of personal data, and therefore renders the application of data protection law to smart cities all the more uncertain. Furthermore, beyond these issues related to the material scope of data protection, there are other structural hurdles that stand in the way of its proper application to contemporary smart cities environments organised within complex public-private partnerships.

2. Stratumseind 2.0 and its living lab[15]

Stratumseind is a busy nightlife street in the centre of Eindhoven, promoting itself as the longest nightlife street in the Netherlands. It is about four-hundred metres long and houses around fifty establishments, such as pubs, cafés, snack bars, a nightclub and a coffee shop, where marihuana is sold for personal consumption. It has been a popular nightlife destination, attracting a diverse and young public, for several decades. According to Eindhoven municipality, however, the number of visitors has been significantly dropping since 2010, arguably due to the rising criminality and vandalism on the street, giving Stratumseind a bad image.[16] In 2013, the municipality thus initiated the *Stratumseind 2.0* project, which officially ran until mid-2018 (although some of the projects seem to be continuing with several of the same parties still in 2020).[17] *Stratumseind 2.0* was an umbrella project in the form of a public-private partnership (PPP) between Eindhoven municipality, the police and a range of private parties, including the Stratumseind establishments association, real-property owners on the Stratumseind street, and Dutch breweries.[18] While

[9] Brunilda Pali and Marc Schuilenburg, 'Fear and Fantasy in the Smart City' (2019) Critical Criminology; Maroš Lacinák and Jozef Ristvej, 'Smart City, Safety and Security' (2017) 192 Procedia Engineering 522.

[10] Purtova (n 6).

[11] Marc Schuilenburg and Rik Peeters, 'Smart Cities and the Architecture of Security: Pastoral Power and the Scripted Design of Public Space' (2018) 5 City, Territory and Architecture; Dorine Duives, Stephvan Beffers and Maurits van Hövell, 'Crowdmonitoring Systeem Amsterdam' (2016) <https://www.noord-holland.nl/Actueel/Tijdelijk/Symposium_Samen_slimmer_reizen_in_de_MRA/Documenten/deelsessie_12_Crowdmonitoring_systeem_Amsterdam.pdf> accessed 6 May 2020.

[12] Sofia Ranchordás, 'Nudging Citizens through Technology in Smart Cities' (2019) 0 International Review of Law, Computers and Technology 1.

[13] 'Eindhovense Innovatie Is Voorbeeld Voor de Rest van Het Land, Staatssecretaris Brengt Bezoek Aan Stratumseind' (*Studio 040*, 2018) <https://archief.studio040.nl/eindhovense-innovatie-is-voorbeeld-voor-de-rest-van-het-land-staatssecretaris-brengt-bezoek-aan-stratumseind/content/item?1109883> accessed 6 May 2020; 'Context Broker's Smart Services Are Making the City of Eindhoven a Safer Place' <https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/2019/06/07/Context+Broker%27s+smart+services+are+making+the+city+of+Eindhoven+a+safer+place> accessed 6 May 2020.

[14] Although some issues connected to personal data have been discussed in relation to a concrete example; e.g. Liesbet van Zoonen, 'Privacy Concerns in Smart Cities' (2016) 33 Government information quarterly 472.

[15] Description based on Maša Galič, *Surveillance and privacy in smart cities and living labs: conceptualising privacy for public space* (doctoral dissertation; Optima Grafische Communicatie 2019).

[16] Eugene van Gerwen, 'Stratumseind 2.0: Plan van Aanpak' (2013) <https://eindhoven.raadsinformatie.nl/document/1047462/1/document> accessed 20 August 2020.

[17] 'The Stratumseind Pilot' (2020) < https://oddity.ai/blog/stratumseind-pilot/> accessed on 16 September 2020.

[18] Grolsch, Heineken, Bavaria and the InBev conglomerate.

the notion of PPP has been defined in many ways, we use the broad definition by Savas, understanding a PPP as 'any arrangement between government and the private sector in which partially or traditionally public activities are performed by the public sector'.[19] This definition includes not only long-term cooperation between public and private parties but also short-term and *ad hoc* cooperation, which is the case in the *Stratumseind 2.0* project.

The main objective of *Stratumseind 2.0* was to 'long-lastingly improve the street from an economic as well as a social point of view'.[20] More concretely, the project wanted to attract more visitors, make them stay longer and spend more money in the establishments, lower the vandalism, police and health-related costs in connection to Stratumseind, increase the income related to Stratumseind and Eindhoven as a whole and create positive value of direct marketing.[21] These ambitious goals were planned to be achieved through a variety of means and initiatives, especially through 'a 365 days, 24/7 scan of *all data on Stratumseind*'.[22] The key element and the main sub-project of Stratumseind 2.0 was thus its living lab – the *Stratumseind Living Lab* (*SLL*).

The *SLL* has been described as a field lab with a variety of sensors mostly located on the Stratumseind street and numerous actors, with the intention of measuring, analysing and stimulating the behaviour of people in public places.[23] Its main goal was to gain insight into the ways in which external stimuli can (substantially) influence the escalating as well as de-escalating behaviour of visitors of the street.[24] Like *Stratumseind 2.0*, the *SLL* was also organised in the form of a PPP, involving a large and growing number of actors. On the one hand, public actors included Eindhoven municipality, the police and local higher education institutions, including the Technical University Eindhoven and Tilburg University. On the other hand, various technology companies, global and large (multinationals such as Philips, Atos and Cisco) as well as local and small businesses (like Sorama and Vinotion), were the main private actors involved.

The *SLL* entailed a 'basecamp' – a physical location on the Stratumseind street, where big screens visualised the real-time collection and analysis of various types of data. The *SLL* included numerous sensors and actuators that were mostly positioned on the Stratumseind street, including: video and sound cameras with embedded analytical capabilities, special lighting and olfactory (emitting smells) technology, CityBeacons,[25] wifi tracking, technology for social media sentiment analysis, and a weather station.[26] These technologies continuously captured and generated data. Other types of data were also collected and stored, including crime statistics, the amount of beer sold, and trash collected per week. All of these data were (and likely still are) stored in a database that can be utilised through data analysis techniques. Data that was captured and collected included: image data from the video feed (with blurred faces), the number of people approaching and moving away from Stratumseind, people density on the street, persons' walking patterns, the total number of people, stress levels in people's voices, the nationality and hometown of the visitors (based on smartphone subscription data received from Vodaphone), the average sound level on the street, petty, moderate and serious crime levels on the street (based on police statistics), the number of tweets with data relating to Stratumseind (e.g. a bar on the street), the percentage difference of beer ordered by the Stratumseind establishments (weekly), volume of garbage collected from Stratumseind (weekly), tons of glass from the street collected (weekly), the number of cars parked in certain car parks in the centre of Eindhoven, the temperature, wind speed and direction, and the amount of rainfall per hour. Note that data pertaining to wifi tracking, tweet sentiment analysis, or visitors' nationality and hometown is only collected at an aggregate level (and the unique identifiers captured through wifi tracking are said to be anonymised). On this basis, the *SLL* actors conclude that the level of aggregation is sufficient for these data to be considered anonymous.[27] In fact, throughout the lifespan of the *SLL* sub-projects, there was no sign on the street that would notify citizens of the various activities happening there.[28]

The *SLL* consisted of a growing number of sub-projects with diverse but intertwined actors and goals, ranging from crime prediction (*CityPulse*), community policing (*Trillion*) and community building (*Stratumsepoort*) to de-escalating people's behaviour via light (*De-escalate*).[29] *CityPulse* and *De-escalate* were the two biggest and most long-lasting sub-projects with potentially the highest impact on privacy and, as such, we will examine them further.

2.1. *CityPulse: predictive policing*

CityPulse was a project developing a system for detection of deviant behaviour and atmosphere on the street, taking place from 2015 to the end of 2017. Some of the world's biggest ICT companies, such as Atos (funded by IBM for this project) and Intel, were parties to it. Other actors included the Eindhoven municipality, the police, and a few smaller, local technology companies. Employing a variety of sensors on the Stratum-

[19] ES Savas, *Privatization and Public-Private Partnerships* (Chatham House 2000).

[20] van Gerwen (n 16).

[21] Tinus Kanters, 'Living Lab, Onderdeel van Stratumseind 2.0, Smart Sensors, Smart Interfaces, Smart Actors, Smart Lights, Smart Data, Smart Design, Augmented Reality, Gaming'.

[22] Ibid.

[23] Pieter Ballon, 'Living Labs' in Robin Mansell and Peng Hwa Ang (eds), *The international encyclopedia of digital communication and society* (John Wiley & Sons 2015).

[24] Clara Kuindersma, 'De Openbare Ruimte Als Proeflab Voor Nudging' (2018) *Stadszaken*.

[25] Large objects combining the functions of cameras, information signs, signposts, antennas, advertising spaces and video screens.

[26] Several of these technologies, including the video cameras, sound sensors and lighting technology, were developed for the *SLL*, either by multinational technology companies (Atos, Intel, Philips) or smaller local businesses (Vinotion, Sorama), sometimes in co-operation with Dutch universities (Technical University Eindhoven).

[27] Peter de Graaf, 'Een biertje met Big Brother erbij op Stratumseind' *De Volkskrant* (23 November 2015).

[28] At the moment that the authors last checked, in August 2019, there were still no signs notifying the data capture taking place on the Stratumseind street.

[29] 'The Stratumseind Pilot' (n 17).

seind street and social media analysis, the *CityPulse* system aimed to create 'a powerful picture of the street and help authorities better predict and react to situations and de-escalate them before they develop.'[30]

In particular, this project employed video cameras (with blurred faces), sound sensors and sound cameras with embedded analytical capabilities, as well as data created by other sensors, such as temperature, wind speed and rainfall, and data gathered from weekly statistics. The *CityPulse* system thus supposedly knew how many people were walking or biking around on Stratumseind at any time of the day, which bar is the most crowded, how fast persons are moving and who has a suspicious walking pattern.[31] The system also performed sentiment analysis of *tweets* that related to the Stratumseind street – for example, a tweet that was posted from the physical location on the street or mentioned chosen 'Stratumseind keywords', such as the street name, a bar on the street or a famous bartender there.

The system was designed to analyse all of these types of data, looking for anomalies in data patterns, which could then be cross-referenced against other gathered data sources. For example, the video cameras of the *CityPulse* system had an embedded capability of tracking walking patterns. As such, the software could single out an individual with a 'suspicious walking pattern' on the street. Such a suspicious walking pattern could refer to someone walking up and down the street numerous times at a slow pace, indicating the possibility of a thief.[32] If additional data sources would confirm the likelihood of an incident, the system would alert the regional police control room (*regionale toezicht ruimte*), giving the police an opportunity to make better informed decisions on any action on the Stratumseind street that might be required.[33] Another example would be, when the *CityPulse* system would detect an 'escalated atmosphere' potentially soon requiring police presence. The police would be alerted through an application (app), first requiring human authorisation but acting autonomously when fully developed, representing a direct link between the *SLL* and the police. The app would have four possible notifications or 'buttons': 'nothing wrong', 'everything alright', 'backup needed' and 'high risk situation'.[34] The *CityPulse* system could also adapt the lighting colour and levels on the street (a technology developed in the *De-escalate* project, which is discussed in the following paragraph). As such, it can be considered a type of predictive policing (that is, a 'policing strategy or tactic that develops and uses information and advanced analysis to inform forward-thinking crime prevention'),[35] with a focus on making aggregate predictions relating to times and places of crimes.[36] The system was thus seen as a predictive, preventative and an ancillary tool for the police in its role for crime and order maintenance in the city.

2.2. *De-escalate: nudging the atmosphere through light*

De-escalate was a project developing a special lighting system with the purpose of influencing and diffusing 'escalated' mood and behaviour through 'dynamic lighting scenarios' in public space, such as on the Stratumseind street, or in small-scale indoor settings, such as prisons, psychiatric wards and help desks.[37] It ran from 2014 to 2018 and was led by researchers from the Technical University of Eindhoven (TU/e) and Philips, a large Dutch technology company from Eindhoven, which designed and provided the lighting system. Other parties to the project include Eindhoven municipality, the police, and smaller local technology companies.

The *De-escalate* project has been described as an intelligent lighting system to control emotion.[38] It experimented with the effect of interactive lighting design in 'de-escalation' of aggressive behaviour, based on psychological pathways through which exposure to dynamic lighting could defuse escalating behaviour. In other words, the goal was to produce more 'social' and less aggressive behaviour, particularly through the use of light in relation to attention and (self-)awareness.[39] Literature in the field of psychology shows, for instance, that dim and warmer colour light is associated with lower arousal and that exposing people to pulsating orange light at slow frequencies leads to relaxing breathing rhythms.[40] Directed or bright light can, according to de Kort, heighten self-awareness, whereas darkness can trigger feelings of anonymity.[41] The awareness of the loss of anonymity, when one is in spotlight,

[30] Atos, 'CityPulse - Using Big Data for Real Time Incident Response Management' (2015) <https://atos.net/wp-content/uploads/2016/06/atos-ph-eindhoven-city-pulse-case-study.pdf> accessed 23 March 2019.

[31] Reinier Kist and Wouter van Noort, 'Het Misdrijf Is Al Ontdekt Voor Het Gepleegd Is' *NRC* (22 August 2015).

[32] Albert H Seubers, 'CityPulse Eindhoven - Netherlands' (2015) <https://www.euroforum.nl/media/filer_public/2015/04/20/albert_seubers_v13_atos_city_pulse_presentation_april_2015.pdf> accessed 16 September 2020.

[33] Atos, 'Intelligent City Management' (2015) <https://atos.net/wp-content/uploads/2016/06/atos-ph-eindhoven-city-pulse-case-study.pdf> accessed 23 March 2015.

[34] It is unclear to the authors what the difference between 'nothing wrong' and 'everything alright' is.

[35] Craig D Uchida, 'A National Discussion on Predictive Policing: Defining Our Terms and Mapping Successful Implementation Strategies' (2009).

[36] Rosamunde van Brakel, 'Pre-Emptive Big Data Surveillance and Its (Dis)Empowering Consequences: The Case of Predictive Policing' in Bart van der Sloot, Dennis Broeders and Erik Schrijvers (eds), *Exploring the Boundaries of Big Data* (Amsterdam University Press 2016) This type of predictive policing has been called 'predictive mapping' – identifying when and where crimes may take place based on aggregate-level analysis. PredPol is the most famous example, making predictions about possible future crime hotspots.

[37] This project run from 2014 to 2018 and was led by researchers from the Technical University of Eindhoven (TU/e) and Philips, a large Dutch technology company from Eindhoven, which designed and provided the lighting system. Other parties to the project include Eindhoven municipality, the police, DITTS and smaller technology companies.

[38] Yvonne AW de Kort, 'Spotlight on Aggression' (2014) ILI 2014 10.

[39] Ibid.

[40] Karin CHJ Smolders, Yvonne AW de Kort and Pierre JM Cluitmans, 'A Higher Illuminance Induces Alertness Even during Office Hours: Findings on Subjective Measures, Task Performance and Heart Rate Measures' (2012) 107 Physiology & Behavior 7.

[41] Yvonne AW de Kort, 'Light on and in Context' (2015) <https://pure.tue.nl/ws/files/8575906/Kort2015.pdf> accessed on 16 September 2020.

may turn a person's attention to their inner states and traits, and prompt them to examine their personal norms and engage in better self-regulation.[42]

Within this project the Stratumseind environment was explored, Stratumseind visitors were profiled in order to find out what causes aggression and escalation on Stratumseind and how such behaviour can be measured, and, furthermore, to explore the possibilities for dynamic light designs, which could reduce these escalation/aggression levels.[43] The project made use of analysis of the data of incidents in the past, 'open data' and social media data, trying to find correlations with all kinds of potentially influencing factors on people's stress levels (such as weather or results of a football match).[44] With the use of such data analysis predictions were made about the stress levels that put the lighting system in motion, aiming to proactively keep stress levels at 'acceptable levels'.[45] What these 'acceptable levels' are was not further specified.

The project thus aimed to provide insights into human behaviour and deliver lighting schemes applicable and effective in real-time conditions in order to de-escalate behaviour on the street. In this sense, it can be seen as a nudging tool or, more broadly, a tool that influences people's behaviour and, potentially, impairs persons' autonomy.

There were two key terms used throughout the project: escalation and atmosphere. The first term, 'escalated behaviour', was used in a very broad sense within this sub-project and the *SLL* in general, referring to all types of behaviour of persons who in some way lose self-control, including screaming, getting abusive, aggressive or crossing other behavioural boundaries that a person would otherwise not cross.[46] Escalation, then, refers to 'an increase in severity of aggressive means used in a given conflict'.[47] The second key term used was atmosphere. This term was seen as being of value, as the police and the security staff on the Stratumseind street often refer to it and use it in order to evaluate the general situation on the street and anticipate people's behaviour. Since aggression (connected to the broader concept of escalated behaviour) is behaviour that is strongly dependant on context (including crowding, noise and temperature), socio-physical characteristics of environments can coerce behaviour in dynamic but patterned, and thus predictable ways.[48] Most often, aggression and escalation occur not because they were intentionally planned, but because people respond to perceived stress, become ignited by autonomic arousal and anger and in this process break personally held norms and revert to things they might not do otherwise. *De-escalate* researchers thus wanted to affect the 'atmosphere' on the Stratumseind street in order to de-escalate aggression. They defined 'atmosphere' as 'people's attitudes, mood, behaviours and interactions with one another as well as with their immediate environment'.[49] It was thus seen as a characteristic of social context that could be transformed into measurable data, coloured through interactions with other visitors.[50] In other words, atmosphere was seen as an important indicator of the risk of incidents[51] as well as a proxy for influencing persons' behaviour.

3. The perennial question: personal or non-personal data?

Based on this description of the two sub-projects, we will attempt to answer the question, whether data protection law applies to the processing of data within the *SLL*. In other words, do the actors within the *SLL* project process personal or non-personal data?

As seen from the description of the initiative, most of the data collected and processed by the living lab seem – at first blush – non-personal. This is connected to the fact that the *SLL* and similar types of initiatives try to 'datafy' public space and the people there through its sensors.[52] For this reason, they processes a lot of data that first and foremost relates to the environment (i.e. 'environmental data'), such as data relating to the weather, amount of beer sold, sound and crowding levels. In this section, we will examine whether data collected and processed within the *SLL*, which we will call '*SLL* data', can nevertheless be considered personal.

The GDPR defines personal data as any information that relates to someone who is identified or identifiable on the basis of that data (Art. 4(1) GDPR). 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' (ibid.). This definition is further explicated by the Article 29 Working Party (Art. 29 WP or WP29)[53] in Opinion 4/2007 on the concept of personal data[54] and, more recently, in the case

[42] Ibid.

[43] Indrė Kalinauskaitė, 'Measuring Stratumseind Experience: De-Escalate Stratumseind' (PDEng rapport 2014).

[44] H den Ouden and AC Valkenburg, 'Smart Urban Lighting' in A. Nighten (ed), *Real projects for real people* (Volume 3, The Patching Zone 2013).

[45] Ibid.

[46] de Kort (n 38).

[47] Zeev Winstok, Zvi Eisikovits and Gideon Fishman, 'Towards the Development of a Conflict Escalation Model: The Case of Israeli Youth' (2004) 33 Journal of Youth and Adolescence 283.

[48] Indrė Kalinauskaitė and others, 'Atmosphere in an Urban Nightlife Setting: A Case Study of the Relationship between the Socio-Physical Context and Aggressive Behavior' (2018) 59 Scandinavian Journal of Psychology 223.

[49] Ibid. 228.

[50] Ibid.

[51] Ibid. 229.

[52] Somia Belaidouni and Moeiz Miraoui, 'Machine Learning Technologies in Smart Spaces', *Ubicomm 2016: The Tenth International Conference on Mobile Ubiquitous Computing Systems, Services and Technologies* (2016).

[53] In the wake of the GDPR, the Art. 29 WP has been replaced by the European Data Protection Board (EDPB). The latter has endorsed all the documents issues by the Art. 29 WP. We continue to refer to the Art. 29 WP for the documents it has authored. See, EDPB, 'Endorsement 1/2018' (2018).

[54] Even though the Opinion concerns the concept of personal data in the Data Protection Directive (DPD), it is still significant after the adoption of the GDPR, as the new regime does not affect the concept of personal data (see Case C-434/16 Peter Nowak v. Data Protection Commissioner (2017) ECLI:EU:C:2017:994, Opinion of Advocate General Kokkot, p. 3).

law of the Court of Justice of the European Union (CJEU).[55] According to WP29, the notion of personal data is very broad, covering all information which *may be linked* to an individual. While the CJEU has not been unequivocal in endorsing such a broad approach,[56] it implicitly confirmed it in the 2017 Nowak Judgment.[57]

According to WP29[58] – and endorsed by the CJEU in Nowak – the notion of personal data has three elements: (1) any information; (2) relating to; (3) an identified or identifiable natural person.[59] 'Any information' is used to reflect the aim of data protection law to 'assign a wide scope to that concept',[60] potentially encompassing all kinds of information (e.g. objective and subjective, sensitive and not sensitive).[61] The more interesting question then concerns the second and third element. In the following sections we will examine both the 'relating to' and 'identifiability' elements of personal data, applying them to the data collected within the *SLL*.

3.1. *Information 'relating to' a person*

According to WP29, information can 'relate to' a person in content, purpose, or result.[62] In some situations, the relationship between the information and the individual is quite obvious. This is generally the case, when the information relates to a person in *content* – that is, it is *about* a particular person, such as one's Twitter account name, one's walking pattern and the level of one's intoxication. In other cases, where the relationship with the person is indirect, the link may not be as self-evident. This is often the case when information relates to a person in *purpose*, that is 'when the data are used or are likely to be used … with the purpose to evaluate, treat in a certain way or influence that status or behaviour of an individual'.[63] In this case, the information might relate to an object (e.g. the value of one's bicycle), a process or an event (e.g. one's participation in a fight). Finally, information can relate to a person in *result*, when its 'use is likely to have an impact on a certain person's rights and interests'.[64] This impact does not need to be 'major'. In fact, it is 'sufficient if the individual may be treated differently from other persons as a result of the processing of such data'.[65] As such, it is 'not necessary that the data "focuses" on someone in order to consider that it relates to [her]'.[66]

Furthermore, the relationship in regard to purpose and result will occur not only when the data is already used, but also where it is *likely to be used* with the purpose or effect of impacting people 'taking into account all the circumstances surrounding the precise case'.[67] The intended and unintended impact or likelihood of impact of data processing thus need to be taken into account too. Consequently, whether a particular information relates to an individual is context-specific and cannot be answered in an absolute manner.[68] The same piece of data can be considered as relating to a person in one case and not relating to a person in another case depending upon a number of factors (e.g. the entity in possession of the data, purposes of processing, current and future technological and organizational context of processing). This analysis of 'relatability' can be said to have been upheld in the CJEU Nowak judgement.[69]

Let us now apply the above considerations to the data collected within the *SLL*. For instance, in the case of trying to detect a thief, where the goal is to single them out and alert police officers to them, data on walking patterns and 'image data' with blurred faces from the video feed will likely be analysed. Both of these types of data relate to the alleged thief in content. However, considering the fact that the surveillance within the *SLL* is aimed at all visitors, most of which will not be (alleged) thieves, we consider that this type of monitoring and data collection will be more of an exception rather than a rule.

Following Purtova, we consider that most if not all of the data processed in a smart environment, such as the *SLL*, relates to a person at least in purpose or impact.[70] This is so because one of the key goals of the *SLL* – a common goal of many smart city and living lab projects – is to process data with the purpose of adapting the environment and influencing persons' behaviour there. As such, *SLL* data can be seen as constituting information that is used or is likely to be used with the *purpose* to evaluate and influence the behaviour of persons on the street.

For instance, if Bart and Marloes – a couple – are having a fight that gets out of hand, resulting in yelling and some bro-

[55] E.g. Case C-582/14, Patrick Breyer v. Bundesrepublik Deutschland, [2016] ECLI:EU:C:2016:779 (hereinafter Breyer); Joined Cases C-141/12 and C-372/12 YS and MS v. Minister voor Immigratie, Integratie en Asiel, (2104), ECLI:EU:C:2014:208 (hereinafter YS and MS).
[56] See YS and MS (n 55).
[57] Case C-434/16, Peter Nowak v. Data Protection Commissioner (2017), ECLI:EU:C:2017:994 (hereinafter Nowak).
[58] While the Art. WP 29 Opinions are not formally binding, its Opinions hold much authority in member states and provide comprehensive guidelines for data controllers as to how they should apply the concept of persona data in practice.
[59] The WP 29 in its Opinion 2007 actually distinguished 'natural person' as a fourth element but as this is a rather straightforward legal term, we will not consider it as a separate element.
[60] Nowak (n 57), para. 34.
[61] WP 29 Opinion 2007. Nowak, para. 34. See also, Case C-131/12, Google Spain SL, Google Inc. v. Agencia Española de Protección de Datos (AEPD), Mario Costeja González (2014), ECLI:EU:C:2014:317 (hereinafter Google Spain).
[62] Art. 29WP, Opinion 4/2007, p. 11.
[63] Ibid. 10.
[64] Ibid. 11.
[65] Ibid.
[66] Ibid.
[67] Ibid.
[68] Ibid.
[69] Nowak (n 57), pp. 35, 56; in the 2016 judgement YS and others, the CJEU did not follow the Opinion of WP29 on the matter. In this judgment concerning exams transcripts, the CJEU interpreted 'information relating to' narrowly, as information about an individual. However, the Court later reversed its stand on the matter in the 2017 Nowak judgment. In Nowak, the CJEU stated that the notion 'personal data' potentially encompasses any information, as long as it *relates* to the data subject, that is where the information is linked to a particular person 'by reason of its content, purpose or effect'. The Court explicitly noted the contradiction with YS and others, thus de facto overruling the older judgment. The legally binding position of the CJEU, at least since Nowak, is thus in line with the position of the WP29.
[70] Purtova (n 6).

ken glass, the *SLL* surveillance system would capture the image (with blurred faces) and the sound of them yelling and breaking glass. The *CityPulse* and *De-escalate* systems might determine that this is a 'low risk' situation on the street and would only turn on the special lighting technologies with the aim of de-escalating their behaviour, rather than notifying the police about it. While 'image data' from the video feed do relate to Bart and Marloes in *content*, this is not necessarily the case for the broken glass noise, general stress level of voices and crowding levels that together trigger the nudging response. These latter data do not relate to the couple in content, but they do in its *purpose*, which is to 'de-escalate' the couple's behaviour. Further, these same data also relate to other bystanders who will be affected by the nudging measures. These data relate to them in *effect*. Thus, while the *SLL* might also process data relating to persons in content, the focus can be said to be on data that relates to persons in purpose and, potentially, in effect.

3.2. *Identifiability: a broad but not unlimited notion*

3.2.1. The Stratumseind Living Lab and the limits of identifiability? The second key part of the definition of personal data is the notion of 'identified or identifiable' (Art. 4(1) GDPR). Whereas 'identified' refers to a person who is known (that is, distinguished in a group), 'identifiable' relates to a person who is not identified yet, but where identification is possible. Moreover, identification can be direct or indirect. A person is identified directly by, for instance, reference to a name, sometimes in combination with additional information if the name is not unique. A person can be identified indirectly through 'unique combinations' of not unique identifiers that allow the individual to be singled out in a group.[71] The notion of 'identified or identifiable' is thus very broad.

Recital 26 GDPR adopts a test of *reasonable likelihood* of identification by the controller or by another person, referring to objective factors, such as the costs of and the amount of time required for identification and taking into account the state of art of technology at the time of processing. According to WP29, a 'purely hypothetical possibility' of identification is insufficient to meet the standard of reasonable likelihood.[72] This test thus has a higher threshold than the 'relating to' element of the definition, where mere likelihood suffices. Besides the costs and the amount of time required for identification, the WP29 expanded the factors that should be taken into consideration by including:

- the intended explicit or implied purpose of processing: when 'the processing … only makes sense if it allows identification of specific individuals and treatment of them in a certain way',[73] the availability of tools of identification should be presumed reasonably likely;
- the risk of organisational dysfunctions (e.g. breaches of confidentiality duties) and technical breaches (e.g. data breaches);
- measures to prevent data identification (i.e. to maintain anonymity) are of importance as a means of avoiding processing of personal data altogether, rather than a fulfilment of data security obligations under the now repealed Data Protection Directive.[74]

The CJEU also had an opportunity to consider the meaning of 'indirect' identification involving 'all the means likely reasonably to be used' in Breyer, a case concerning dynamic IP addresses.[75] In this case, the Court examined whether the possibility for the German state to combine the dynamic IP address with additional identifying information held by the internet service provider (ISP) constituted a means likely reasonably to be used.[76] It determined that even though German law does not generally allow the transmission of such information between the ISP and the state, such transmission is allowed in the case of issues such as cyber-attacks.[77] Furthermore, it considered that such a combination was not practically impossible, in the sense that it would require a disproportionate effort in terms of time, cost and man-power.[78] It concluded that identification was reasonably likely, so that dynamic IP addresses constituted personal data. Consequently, both the WP29 and CJEU interpret 'identifiability' very broadly, allowing for rather onerous steps to constitute 'means reasonably likely to be used'.[79]

So, does the *SLL* process data of identified or identifiable persons? Getting back to the lovers' quarrel between Bart and Marloes, one can argue that images (with blurred faces) and the sound of them yelling and breaking glass are not direct identifiers. The same is true *a fortiori* for other environmental data, that is data primarily relating to the environment, such as weather conditions, general sound and crowding levels, even though indirect identifiability based on these data is likely more difficult than based on video footage. However, and crucially, the point here is that identifiability is not needed for the *SLL*'s purposes. In the case of the *De-escalate* project (i.e., de-escalating behaviour through light), even though its light-based nudging system is primarily aimed at Bart and Marloes it does not target them directly (and hence does not identify them); instead it affects everyone present on the street. In fact, the whole nudging system makes perfect sense without identification or the need for identifiability. That is, even if individuals are not identifiable, they may be nudged in this or that way for this or that purpose. If Bart and Marloes are indeed calmed down by the lighting scenario and end up kissing in a secluded alley – in other words, if the nudging works – then their identifiability is not needed at all; neither now nor later.[80]

A similar assessment can be made in regard to a higher risk situation detected by the *SLL* system, such as a fight likely to break out, which would merit the deployment of the police on

[71] Art. 29WP, Opinion 4/2007, p. 13-14.
[72] Art. 29WP, Opinion 4/2007, p. 15.
[73] Ibid. 16.
[74] Ibid.
[75] Breyer (n 55).
[76] Lynskey (n 5).
[77] Breyer, para. 47.
[78] Ibid., para. 46.
[79] Lynskey (n 5).
[80] Cf. Schreurs and others who comes to the same conclusion in regard to behavioural biometric profiling; Wim Schreurs and others, 'Cogitas, Ergo Sum. The Role of Data Protection Law and Non-Discrimination Law in Group Profiling in the Private Sector' in Mireille Hildebrandt and Serge Gutwirth (eds), *Profiling the European Citizen: Cross-Disciplinary Perspectives* (Springer 2008) 243.

the street. In this case, the purpose of the *CityPulse* project will be to detect a risky situation that might otherwise remain undetected through the regular ('deaf and dumb') CCTV operated by the police. The goal here is to detect a 'bad atmosphere', in which a fight is likely to break out, so that the *CityPulse* system can alert the police ahead of time, allowing them to arrive on the street more quickly and have a pre-emptive effect on persons ready to pick a fight. In order to achieve this goal, the *CityPulse* would generally rely upon the same types of data as mentioned in the above example of the lovers' quarrel and the *De-escalate* project. Consequently, that in order to pre-empt a fight through atmosphere detection, identification is also not needed.[81]

In both of these cases no direct identifiers are at play, meaning that no personal data seems to be processed. However, is it possible to argue otherwise? Could individuals nonetheless be considered identifiable?

3.2.2. *Auxiliary or additional identifying information* One way to render the individual identifiable is by adding additional information so as to be able to single them out, what has been called auxiliary information.[82] Such auxiliary information depends on a number of technical factors. For instance, the level of granularity of the profiles used for nudging purposes, such as the number of data points they rely upon. As Purtova argues, the type of advanced machine learning technology commonly at play nowadays relies upon high-dimensional databases – databases that construct objects through an important number of features (or data points), which can lead to the identifiability of individuals through the matching of the features between the data points or databases at play.[83] Such cross-referencing allows for the identifiability of individuals: since individuals are constructed on the basis of many features, one can compare two or three databases and rapidly spot these unique constructions of features that stand out.[84] One should keep in mind however that comparing databases also relies upon additional social or organisational factors such as the extent of sharing taking place among the various databases.

Does the *SLL* context provide for such additional identifying information? On a general level, one can argue that this is the case *insofar* as the functioning of the *SLL* is predicated upon the interoperability of the databases of the various actors involved (cf. PPP), which in principle enables their cross-referencing.[85] Even if the databases used by the *SLL* do not technically qualify as high-dimensional, they certainly do operate on a high number of data points, and such cross-referencing could allow for the identifiability of individuals in the databases. This is particularly the case relating to wifi tracking or phone subscription data mentioned in Section 2. Even though these data points are collected at an aggregate level, when the various databases wherein they are contained are cross-referenced, the singling out of the individuals at stake cannot be excluded. Furthermore, in certain specific cases, the *SLL* might make use of additional identifying data, which have a closer link with individuals. For instance, the *SLL* features the possibility of sentiment analysis of *tweets* that relate to the Stratumseind street (e.g. tweets that mention the word 'Stratumseind' or that mention a bar on the street), as well as the collection of various behavioural biometric data (such as walking patterns). One can imagine a case in which Marloes would post a tweet with a picture of herself and Freek (tagging themselves in the post) making up after a quarrel, either mentioning the street by name or having the location embedded in the metadata of photo. Our quarrelling lovers would thus be quickly and easily individually identifiable (possibly by name, if they would use their real names on Twitter).

Another way of adding information is by inferring information. In the case of the video footage where people's faces are blurred, it might nonetheless be possible to infer the identity of the people (at least, in the sense of singling out). In order to infer such information, the SLL would need to rely upon the interoperability of databases, according to which inference is possible because of the existence of additional information.

These considerations point to the possibility that the data processed by the *SLL* is personal, however, they hinge upon a number of hypothetical factors that cannot be answered in advance. In any case, it is worth re-emphasising that the *SLL* does not need personal data in order to function properly; instead, it is meant to operate on the basis of non-personal data.

3.2.3. *Identifiability and profiling* Another way to approach the issue of identifiability is through a broader, socio-technical perspective, by looking at the purpose of processing at stake in any particular example.[86] As a type of smart environment, the functioning of the *SLL* relies upon a number of data mining algorithms, which create 'risk profiles'.[87] Think of the four 'risk situations' of the *CityPulse* system (discussed in 2.1). For instance, in the example of the lovers' quarrel, the *SLL* algorithms classify Bart and Marloes as being part of a 'low risk profile', which warrants nudging actions through light. In other words, the *SLL* is predicated on profiling operations. Before we can discuss the particular type of profiling taking place within the *SLL*, the concept of profiling merits some discussion.

[81] The situation would be different if a fight would nevertheless break out on the street and the perpetrators would not be apprehended immediately. Then the police would resort to its own high-resolution, non-blurred video feed from the CCTV (and eventually also post these images in the news and on social media, asking other citizens to help them identify them, as is common practice). However, this is a different matter and a distinct data processing situation, with which the present paper is not concerned.

[82] Lorenzo Dalla Corte, 'Scoping Personal Data: Towards a Nuanced Interpretaton of the Material Scope of EU Data Protection Law' (2019) 10 European Journal of Law and Technology 1.

[83] Luc Rocher, Julien M Hendrickx and Yves-Alexandre De Montjoye, 'Estimating the Success of Re-Identifications in Incomplete Datasets Using Generative Models' (2019) 10 Nature Communications; Purtova (n 7).

[84] Rocher, Hendrickx and Montjoye (n 83).

[85] See, Maša Galič, 'Surveillance, Privacy and Public Space in the Stratumseind Living Lab: The Smart City Debate, beyond Data' (2019) Ars Aequi 570.

[86] On the social aspect of identifiability, see Miranda Mourby and others, 'Are "Pseudonymised" Data Always Personal Data? Implications of the GDPR for Administrative Data Research in the UK' (2018) 34 Computer Law and Security Review 222.

[87] Belaidouni and Miraoui (n 52).

The process of profiling relies among other things upon data mining (i.e., the extraction of knowledge), prediction, and decision-making.[88] Profiling can thus be said to coincide with data analytics. In other words, profiling is a partly automated process used to find correlations in large data sets in order to build classes or categories of characteristics, which can then be used to generate profiles of individuals and groups, but also of places, events or whatever is of interest.[89] An individualised (or personalised) profile is built entirely on the basis of data pertaining to one individual. In practice, group profiles, which represent an individual only insofar as she is part of a group (i.e. she shares the characteristics of the group), are far more common.[90] Individual or group profiles are thus a way of reconstructing an individual through a number of features (such as walking speed or routines), rendering her *knowable* through data mining algorithms.[91] This serves to predict individuals' future behaviours and to take decisions affecting them on this basis.[92] Finally, applying a profile can be defined as 'the process of identifying and representing a specific individual or group as fitting a profile and of taking some form of decision based on this identification and representation.'[93]

While the question as to whether profiling amounts to the processing of personal data has been hotly debated in data protection scholarship, it has as yet not been definitively settled.

On the one hand, Schreurs et al. have argued that profiling does not involve the processing of personal data, when the data used to build the profile does not relate to an identifiable individual.[94] In this sense, profiling can be distinguished in three steps: (1) processing (personal and/or non-personal) data; (2) creating a profile; and (3) applying the profile. According to this position, if the first step does not process personal data, the rest of the steps in the profiling operation cannot be considered to involve (the processing of) personal data either. Schreurs et al. take the case of behavioural biometric data such as the way that a shopping trolley is driven in a supermarket as a means to infer the type of customer at play (e.g. hurried, higher purchasing power). This type of data does not allow for the identifiability of the individuals pushing the trolley so that the profiling operation will escape the reach of data protection law.[95] The Irish data protection Commissioner has adopted a similar decision in the case of facial detection technology for advertising purposes (i.e., inferring mood, age, gender, etc. on the basis of facial features without actually identifying people), considering it non-personal data.[96] In the context of affinity profiles, Wachter has also defended a similar view by arguing that an affinity profile built on the basis of anonymous data will not constitute personal data even though sensitive information can be inferred from it.[97] These views leave out the fact that profiling is a continuous process based on the creation and subsequent application of profiles, and that 'while the distinction between the two may be analytically salient, in practice the two phenomena intermingle.'[98] In other words, these views argue that it is possible to artificially distinguish between the creation and application of a profile and that on this basis the creation of profiles based on non-personal data would escape the scope of data protection law.

On the other hand, others have adopted a holistic view of profiling in order to argue that even on the basis of profiles not containing identifying information, profiling does amount to the processing of personal data. This is the position of the Council of Europe (CoE) Recommendation on profiling.[99] The main argument here is that even if based on an anonymous profile, the *application* of the profile to specific individuals entails *per se* that these individuals are identifiable. Put simply, one needs to be able to single out a person in order to apply the profile at all. Similarly, Bosco et al. define the application of the profile to individuals as 'the process of *identifying* and representing a specific individual or group as fitting a profile and of taking some form of decision based on this identification and representation.'[100] These authors argue that a strict distinction between the three profiling steps is artificial. On the contrary, all of the steps should be seen as 'inseparable, [and as such] they must all be considered part of personal data processing'.[101] This is very much in line with the reasoning of WP29 concerning the purpose of the processing operation as a key parameter of identifiability. If the processing only makes sense insofar as it allows for the treatment of a data subject in a certain way (which is precisely what is at stake with profil-

[88] Mireille Hildebrandt, 'From Data to Knowledge: The Challenges of a Crucial Technology' (2006) 30 Datenschutz und Datensicherheit 548; Francesca Bosco and others, 'Profiling Technologies and Fundamental Rights. An Introduction', *Profiling technologies in practice: Applications and Impact on Fundamental Rights and Values* (Wolf Legal Publishers 2015); Bart Custers, *The Power of Knowledge: Ethical, Legal, and Technological Aspects of Data Mining and Group Profiling in Epidemiology* (Wolf Legal Publishers 2004).

[89] Bosco and others (n 88).

[90] Hildebrandt (n 88).

[91] David-Olivier Jaquet-Chiffelle, 'Reply: Direct and Indirect Profiling in the Light of Virtual Persons', *Profiling the European citizen: cross-disciplinary perspectives* (Springer 2008).

[92] Mireille Hildebrandt, 'Defining Profiling: A New Type of Knowledge?' in Mireille Hildebrandt and Serge Gutwirth (eds), *Profiling the European citizen: cross-disciplinary perspectives* (Springer 2008).

[93] Bosco and others (n 88).

[94] Wim Schreurs and others (n 80) 243.

[95] Ibid.

[96] Irish Data Protection Commissioner, 'Annual Report' (2017) 16. See, however, the opinion of the Dutch AP concerning advertisement columns with cameras capable of facial detection, which states that personal data is processed, although without much discussion on how identifiability is established; Autoriteit Persoonsgegevens, 'AP Informeert Branche over Norm Camera's in Reclamezuilen' (2018) <https://www.autoriteitpersoonsgegevens.nl/nl/nieuws/ap-informeert-branche-over-norm-camera's-reclamezuilen#subtopic-1727> accessed 27 March 2019.

[97] Sandra Wachter, 'Affinity Profiling and Discrimination by Association in Online Behavioural Advertising' (2020) 35 Berkeley Technology Law Journal 1, 54–55.

[98] Hildebrandt (n 92) 19.

[99] Council of Europe, 'The Protection of Individuals with Regard to Automatic Processing of Personal Data in the Context with Regard to Automatic Processing. Recommendation CM/Rec(2010)13 and Explanatory Memorandum' (2010).

[100] Bosco and others (n 88).

[101] Council of Europe (n 99) para. 57; similarly, Hildebrandt has also argued that while the construction and application of the profile can be distinguished for the sake of analytical clarity, in practice this distinction collapses (see Hildebrandt (n 92)).

ing), then the identifiability of the individuals is implied by its very purpose.[102] Nevertheless, there is as yet no binding legal decision confirming this position.

4. Profiling and nudging in the Stratumseind Living Lab: from individuals and groups to atmospheres

Following the above discussion on issues of identifiability and profiling, we will now examine the specific type of profiling in the *SLL* in more detail. On the one hand, it can be argued that the profiling performed in the *SLL* does not process personal data, since the data used to build the profiles relates to the environment rather than to identifiable individuals. On the other hand, it can be argued that the *SLL* processes data relating to identifiable individuals, as identification is implied by the very purpose of the profiling operation – to affect persons' behaviour on the Stratumseind street. However, in the case of the *SLL*, persons are not to be affected as specific individuals or even as algorithmic groups, but only indirectly as a part of the general atmosphere on the street. We argue that a new type of profiling operations in smart environments is found in the *SLL*, what we call 'atmospheric profiling', which puts an additional strain on the notion of personal data.

The ideas behind the *SLL* are deeply rooted in the notion of atmosphere. While attempts to affect the atmosphere, especially for purposes of boosting sales, are nothing new,[103] trying to affect it via sophisticated digital technologies in the (partially) public sector is a more recent development. The goal of the *De-escalate* project was to create 'good atmospheres' on the Stratumseind street in order to de-escalate potential aggression. Similarly, a part of the goal of the *CityPulse* project can be described as detecting 'bad atmospheres' on the street in order to deploy the police in a more efficient way. Atmosphere was seen as being constituted, among others, from data relating to people's attitudes, mood, behaviours and interactions with one another as well as with their immediate environment (see Section 2.2).[104] The *SLL* is therefore based on the detection of a positive or negative *atmosphere*, with the intention of directly affecting this atmosphere – rather than any particular individuals – so as to reduce aggression and violence. In other words, the *SLL* is based on the creation of profiles of atmospheres – *atmospheric profiles* – which are then translated into 'everything alright' or 'high risk' profiles within the *CityPulse* project. Atmosphere can thus be described as a proxy to only indirectly affect and nudge people, who are reduced to a constitutive element of the atmosphere on the Stratumseind street.

By shifting the focus away from individuals to the broader environment and atmosphere, atmospheric profiling can be said to put an additional burden on the notion of identifiability and, consequently on the notion of personal data. If the target of the atmospheric profile is to influence the atmosphere on the street, this seems to refute arguments in favour of identifiability based on the purpose of the processing; that is, processing, which only makes sense, insofar as it allows for the treatment of a particular data subject in a certain way.[105] Of course, while atmospheric profiling directly affects atmospheres, its underlying goal is to indirectly affect (that is, nudge) people – their mood, behaviour, and interaction – in a particular place. After all, the term atmosphere in contemporary vernacular use refers to the distinctive 'influence' of a place on persons.[106] However, even if we accept that the purpose of this profiling is to indirectly affect persons' behaviour in a certain way, no particular individuals are or need to be singled out, so that no individuals are identified. Consequently, identifiability through purpose does not apply. In other words, whereas usual types of profiling practices lead to identification in terms of purpose (at least under certain interpretations), the same cannot be held in relation to this new type of profiling practice is concerned. From the perspective of profiling, data protection law would not apply to the majority of the data processing taking place within the *SLL* and similar smart city initiatives,[107] confirming what the *SLL* actors have been claiming. As Kanters, the *SLL* project manager, put it: 'In the absence of legislation in the Netherlands, we have drawn up our own data principles. If you want to build a house in the Netherlands, books full of rules apply before any stone has been laid. *There is nothing that applies to the use of data*. You can reason then, it is not forbidden, so just go ahead.'[108] Yet, should this indeed be so?

As already mentioned, in order for persons to be affected, they do not need to be identified by name or another unique identifier, or even singled out from the group, confirming once again that '[e]ven when individuals are not "identifiable", they may still be "reachable"'.[109] Perhaps luckily then, nudging based on atmospheric profiles, which does not single out persons, does not seem to work very well. At least, this seems to be the case within the *SLL*.[110] This is likely – at least par-

[102] Art. 29WP, Opinion 4/2007, p. 16.

[103] See the literature on 'atmospherics'; e.g. Kotler Philip, 'Atmospherics as a Marketing Tool' 48; Salomão Alencar de Farias, Edvan Cruz Aguiar and Francisco Vicente Sales Melo, 'Store Atmospherics and Experiential Marketing: A Conceptual Framework and Research Propositions for An Extraordinary Customer Experience' (2014) 7 International Business Research 87.

[104] Kalinauskaitė and others (n 48).

[105] Art. 29WP, Opinion 4/2007, p. 16.

[106] 'Atmosphere' (*Merriam-Webster online dictionary*) <https://www.merriam-webster.com/dictionary/atmosphere>.

[107] With the accepted exception of wifi-tracking and facial recognition.

[108] 'In Eindhoven Herkent Een Algoritme Vechtpartijen' (*Privacy Web*, 2019) <https://www.privacy-web.nl/artikelen/in-eindhoven-herkent-een-algoritme-vechtpartijen> accessed 22 May 2020. Emphasis added.

[109] Solon Barocas and Helen Nissenbaum, 'Big Data's End Run around Anonymity and Consent' in Julia Lane and others (eds), *Privacy, Big Data, and the public good: frameworks for engagement* (Cambridge University Press 2014).

[110] This is at least the case in De-escalate project, which has not resulted in any tangible de-escalating effects on aggressive behaviour. Tinus Kanters, the *SLL* project leader, stated: 'We thought that that the atmosphere could be influenced in this way [with dynamic lighting scenarios] but this was not the case in practice.

tially – connected to the way in which (manipulative) nudging works: trying to covertly subvert another person's decision-making power through exploitation of the person's cognitive weaknesses and vulnerabilities.[111] Consequently, nudges that are applied to everyone in the same way (such as the lighting scenarios in *De-escalate*), lack the key characteristic required to exploit someone's cognitive vulnerabilities: knowing what they are and how to leverage them. However, it should be noted that ICTs are well-suited to facilitate nudging that would allow for 'fine-grained microtargeting', that is *hypernudging*,[112] which target and exploit individual vulnerabilities, making them much more difficult to resist. In the case of hypernudging, one could thus speak of affecting individuals, meaning that data protection law would much more likely apply.

One could thus imagine that in order to improve the functioning of the SLL, profiling might move towards creating more specific and fine-grained atmospheres. For instance, creating a profile of the atmosphere of a specific corner on the Stratumseind street, where a smaller or more specific group of persons would gather, so that more granular data would be collected, making the atmospheric profile applicable to a set of persons in more narrowly delineated place. There are good chances that this type of more granular atmospheric profiling would allow for the identifiability of individuals. In any case, the distinction between nudging that targets individual vulnerabilities and nudging that does not, is certainly not clear cut. And even where nudging does not target identifiable persons, there are nonetheless important risks for a wide variety of rights and freedoms, which data protection law is meant to protect.[113] The most obvious risk here concerns the right to privacy or private life as found in Article 8 of the European Convention on Human Rights. This is so because nudging (hyper-nudging even more so) poses risks for autonomy, a key argument for why we value privacy,[114] one that is recognised by the European Court of Human Rights. For instance, when the environment is intentionally designed so as to covertly subvert individual decision-making in particular directions, we can speak of manipulative nudging, which bypasses autonomous decision-making.[115] Furthermore, in view of the scarce information concerning the specific functioning of smart city projects provided to the public, it remains unclear how the public could tell (and thus object), when the nudging might be targeting individuals, thus invoking the protection of data protection law. The reach of data protection law, as the most obvious legal framework to regulate smart city-types of initiatives, is thus very limited in smart city initiatives such as the *SLL*, where increasing amounts of environmental data are collected for the purpose of atmospheric profiling.

5. Conclusion

The goal of this contribution was to explore whether and to what extent data protection law could apply to smart city-types of initiatives focused on safety and security, more specifically, whether the data processed within them can be considered personal. It did so by examining a concrete example of a smarty city-type of initiative – the *Stratumseind Living Lab* (*SLL*) in the Netherlands. This is a worthwhile endeavour, as smart city initiatives (try to) affect and reshape both places and persons, thus posing important risks to the enjoyment of our fundamental rights and freedoms, such as privacy and data protection. A clear regulatory framework that would regulate such initiatives is thus needed.

Applying the four-folded definition of personal data (defined as any information relating to an identified or identifiable natural person) to the *SLL* shows that the crux of the problem revolves around the issue of identifiability. First, the majority of the data collected within the *SLL* does not concern the individuals as such. Rather, it relates to algorithmic groups of persons (for instance, walking patterns suggesting thieves) and, predominantly, to the environment itself (e.g. data about the weather, air quality, sound and crowding levels). Looking at identifiability from a technical perspective, this raises the question, whether there is sufficient additional data available that would render the individuals identifiable. Second, in relation to smart city-goals concerning nudging based on profiling, we can conclude that the *SLL* is generally not interested in identifying and targeting specific individuals. Rather, the focus is on the management and nudging of individuals conceived as a multiplicity – a combination of the environment, persons and all of their interactions. This points to a two-fold issue concerning profiling. On the one hand, it adds to and further complicates the discussions around the question of whether profiling constitutes a form of personal data processing simply because of its purpose to affect individuals (in the case of the *SLL*, non-identified persons). This issue, which has its proponents and opponents, has not yet been settled. On the other hand, it also implies a novel type of profiling – *atmospheric profiling* – which tries to indirectly affect persons by affecting the general atmosphere on the street (rather than singling out individuals). As such, this type of profiling does not seem to constitute a type of personal data processing. This might explain why in practice many actors of smart city projects consider that their data processing operation involve non-personal data thereby creating a legal vacuum.

So, what is the way forward? Surely, these pervasive and risky data-driven technologies can't be left unregulated.

A valid consideration at this point would be to question, whether the notion of personal data should be stretched so

---

In any case, it is hardly measurable.' Diede Hoekstra, 'Netwerk van Hypermoderne Camera's Op Stratumseind in Eindhoven Gaat Politie Helpen' (*ED*, December 2017) <https://www.ed.nl/eindhoven/netwerk-van-hypermoderne-camera-s-op-stratumseind-in-eindhoven-gaat-politie-helpen%7B~%7Da1e8acee/>.

[111] Daniel Susser, Beate Roessler and Helen Nissenbaum, 'Online Manipulation: Hidden Influences in a Digital World' (2019) 4 Georgtown Law Technology Review 1.

[112] Karen Yeung, '"Hypernudge": Big Data as a Mode of Regulation by Design' (2017) 20 Information, Communication and Society 118.

[113] See Arts. 24, 35 and Recital 75 GDPR.

[114] Galič (n 85); Marjolein Lanzing, '"Strongly Recommended" Revisiting Decisional Privacy to Judge Hypernudging in Self-Tracking Technologies' (2019) 32 Philosophy and Technology 549.

[115] Not all nudging is manipulative; for an excellent discussion of nudging, manipulation, persuasion and coercion see Susser and others (n 111).

far as to cover situations of data processing, where the identification of individuals (at least by singling them out) is not a goal at all, but where the intention is to manage or nudge people as a multiplicity. This is not a novel proposal. More than a decade ago already, scholars had argued for a shift from *personal* data protection towards data protection *tout court*.[116] That is, the application of data protection law to each processing of data that has potential negative consequences for our rights and freedoms irrespective of whether the data processed qualify as personal. Similarly, some scholars have casted doubt on the actual level of protection that the requirement of identifiability affords to data subjects in practice,[117] whereas others have proposed to replace the requirement of identifiability with that of 'reachability'.[118] Such perspectives are consistent with the overall objective of data protection law, which is to protect individuals' rights and freedoms in the context of the processing of data.[119] Yet, others have warned against such 'over-stretching' of data protection law, positing that it should not try to become a law of everything and thus of nothing.[120] But if data protection were to be discarded in such cases, it still remains unclear which type of (secondary) law should take on this role. Should it be left to municipal ordinances or another type of administrative law having to do with public order and safety? Another possibility could be to employ self-regulatory instruments going beyond the bounds of data protection law; that is, instruments that can assess the acceptability of these types of projects without being bound by the distinction between personal and non-personal data. A notable example are impact assessments, in particular surveillance impact assessments,[121] or in the context of increasingly smart data driven environments, so-called 'care robot impact assessments'.[122] However, the pitfalls of such self-regulatory initiatives have long been underscored, especially concerning the large amount of discretion that they afford.[123]

Finally, to echo Lynskey, even if data protection were to undoubtedly apply to smart environments, it may not be of much assistance to those whose fate is affected by activities within and connected to smart cities and living labs.[124] Indeed, beyond the issue of personal data, a number of other hurdles remain, which stand in the way of smooth implementation of data protection law. In particular, one can point to a number of dualisms that have been sustained rather than replaced during the GDPR adoption process. Beyond the distinction between personal and non-personal data, these include the distinction between data controller and data processor, and between processing for private or law enforcement purposes just to name a few.[125] Critically, these distinctions appear to be out of touch with the organisational and technological reality of contemporary smart cities organised within complex public-private partnerships, potentially preventing a successful application of data protection law to an increasing number of *loci*, where it is meant to apply.

Funding

The research for this Article was made possible by a grant from the Netherlands Organisation for Scientific Research (NWO, project number 453-14-004) and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (INFO-LEG project, grant agreement No 716971).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank Bert-Jaap Koops for commenting on an early version of this paper. They would also like to thank the anonymous reviewer for their helpful comments.

[116] Paul De Hert and Serge Gutwirth, 'Regulating Profiling in a Democratic Constitutional State' in Mireille Hildebrandt and Serge Gutwirth (eds), *Profiling the European Citizen: Cross-Disciplinary Perspectives* (Springer 2008) 289.

[117] Worku Gedefa Urgessa, 'The Protective Capacity of the Criterion of 'Identifiability Under EU Data Protection Law' (2016) 2 European Data Protection Law Review.

[118] Barocas and Nissenbaum (n 109).

[119] This was stated more clearly in the data protection Directive than in the GDPR. Art. 1(1) of the former stated that its goal is to "protect the fundamental rights and freedoms of natural persons, and in particular their right to privacy with respect to the processing of personal data."2.

[120] Purtova points specific attention to the resource intensive compliance regime of the GDPR. See, Purtova (n 6).

[121] David Wright, Michael Friedewald and Raphaël Gellert, 'Developing and Testing a Surveillance Impact Assessment Methodology' (2015) 5 International Data Privacy Law 40.

[122] E Fosch-Villaronga, 'Creation of a Care Robot Impact Assessment' (2015) 9 International Journal of Humanities and Social Sciences 1913.

[123] Concerning these criticisms in the context of data protection impact assessments and the GDPR, see Raphaël Gellert, 'The Article 29 Working Party's Provisional Guidelines on Data Protection Impact Assessment' (2017) 2 European Data Protection Law Review 1, 216–217.

[124] Lynskey (n 5).

[125] See for instance Paul De Hert and Vagelis Papakonstantinou, 'The New General Data Protection Regulation: Still a Sound System for the Protection of Individuals?' (2016) 32 Computer Law and Security Review 179.