Satellite Imagery, Very High-Resolution and Processing-Intensive Image Analysis: Potential Risks Under the GDPR

Cristiana SANTOS* & Lucien RAPP**

Two main trends are currently developing in the satellite imagery industry: the increasing availability of very high spatial and temporal resolution satellite imagery, and the outsourcing of processing-intensive image analysis. Alongside the foreseeable improvements of facial recognition technology and other image recognition software, such synergies carry the potential for identification of individuals, and thus for privacy, data protection and ethical risks. The intent of this article is to discuss the possibility of identification of individuals through high resolution images under the broad definition provided by the general data protection regulation, and to explain the risks therein. We further suggest risk-mitigation approaches for incoming space data policies.

1 INTRODUCTION

Within the satellite data value chain, data is collected remotely from sensors placed on satellites orbiting the planet. Once the process for collecting the imagery is set up, access to the data is quite fast, and the analysis can be largely automated with machine learning algorithms.

The advent of free open-source satellite imagery and technological developments in data infrastructure are producing a culture of shared and increasingly open information available to both public and private industry,[1] making access and use of space-based data more democratized.

Such democratization of satellite imagery is aligned with the Space 4.0[2] era. The latter represents the current space level playing field, where an increased number of space actors around the world (including private companies, academia, industry and citizens) are pursuing both disruption and opportunity, made possible by the acceleration of technology, where trends include space big data, predictive

* Research Associate, Chaire SIRIUS. Email: cristiana.santos@ut-capitole.fr.
** Professor, Scientific Director, Chaire SIRIUS. Email: lucien.rapp@ut-capitole.fr.

1 https://insuranceday.maritimeintelligence.informa.com/ID004586/Eyes-in-the-sky-paint-clearer-picture-of-world-of-risk.

2 http://www.esa.int/About_Us/Ministerial_Council_2016/What_is_space_4.0.

Santos, Cristiana & Rapp, Lucien. 'Satellite Imagery, Very High-Resolution and Processing-Intensive Image Analysis: Potential Risks Under the GDPR'. *Air & Space Law* 44, no. 3 (2019): 275–296.

analytics, imagery geospatial analytics, and data convergence. '*Space-as-a-service industry*'[3] is becoming a reality due to the existing demand for services, such as imagery-related services. According to Northern Sky Research (NSR), big data analytics via satellite will generate close to USD 18.1 billion in cumulative revenues by 2027. The highest increase will be in satellite imagery for data analytics applications, which are predicted to grow 23.5% through to 2027.[4]

As of today, commercially available imagery supplied by providers like DigitalGlobe's WorldView-3 satellite constellation, have resolution in which each pixel in a captured image corresponds to approximately 31 cm.[5] Notably, there is a tendency in the industry to push for the resolution restrictions threshold to be lowered to 10 cm and if such featuring resolution laws are approved, DigitalGlobe will be able to sell commercial high-resolution images to any companies willing to pay for it.

Yet, this does not suffice to be able to distinguish individuals or their features. Indeed, there is a need to debunk popular misconceptions about satellite imagery and its powers. At the present moment, it is not possible to directly identify an individual using today's satellites.

Moreover, processing higher-resolution imagery often creates considerable challenges for data analysis. The complexity in turning raw remote sensing data into post-processing imagery (or meaningful insights therefrom), often has several requirements: data science expertise, computer time necessary for processing, computing processing power, and the necessary budget. The higher the resolution, the greater the need for computer power analysis, and for in-house data science expertise. This means that even at 30 cm resolution, an uncompressed image of a big city would require millions of pixels of data [2].

Furthermore, location imagery data to be accurately used needs to surpass six dimensions of granularity that, cumulated, foster its amenable usage for further applications: frequency, latitude, longitude, altitude, time, and precision [33]. However, to gather data with respect to each of these dimensions for a single specific location is not always feasible.

Regardless of these requirements, organizations (such as Google) have recently sought to fill this need for imagery data by sharing intelligence analysis resources with other players in the data-sharing field.

---

[3] Vaibhav Sharma, *Mini Satellites, Maximum Possibilities*,https://www.livemint.com/Leisure/yEXAKO6k0UWRLtV6rzdQaP/Mini-satellites-maximum-possibilities.html (2018).

[4] The report identifies seven vertical markets as growth areas, and more than 70% of the share is held by the Transportation, Government & Military (Gov/Mil) and Energy markets throughout the forecast period, NSR's Big Data Analytics via Satellite, 2nd Edition, https://www.nsr.com/research/big-data-analytics-via-satellite-2nd-edition/.

[5] http://worldview3.digitalglobe.com/.

The demand for fast and accurate information worldwide is leading to the growth of space-generated and distributed data. As massive constellations of small satellites[6] are about to be launched in LEO in the next years, significantly increased data, observation capabilities, and high-quality imagery from EO satellites[7] are expected to become more widely available on a timely basis. EO massive constellations may provide more frequent image capture and updates (capturing a single point several times a day) at a much lower cost. Users can plan both the target and frequency, allowing for a more specific analysis in a particular tracking. This approach is revolutionizing the field, providing greater access to data and intelligence gathered from satellite imaging.

Given the growing commercial market of high-resolution imaging, and the advancements in satellite technology and sensor resolutions, it is likely that high-resolution space-based data collection for commercial purposes will improve.

Additionally, the availability of large data storage space and the advances in computing capabilities has resulted in an abundance of high-quality satellite data. As low-cost, highly responsive commercial satellite systems become operational, high-resolution imagery is expected to become a regular attribute of end-user products and information services.

The combined synergy of improvements in satellite imagery resolution, facial recognition technology (and other image recognition software), real-time imaging, and big data analytical software will enable end-users to observe industrial activity and the environment in far greater detail, which will assist in making more informed business decisions.

Conjectures revolve around satellite imagery discerning license plates, individuals, and 'manholes and mailboxes'.[8] It is claimed that such granular location-based information would only occur within secondary-use cases [33] (e.g. data analysis for smart cities, marketing profiles), and not undertaken by first-party uses (raw-data to analyse trends, user behaviour, detect security threats, improve a geo-aware service; and geo-fencing).

The ITU-T Study Group 17 (SG17),[9] EO experts and legal scholars foresee that, in concomitance to the growing resolution of remote sensing images, the likelihood of privacy, data protection and ethical issues also grow [1][16][17][18][19][34][35], demanding protection therefrom.

---

[6] The EO constellation will be centred at 600 km, which spans a large range of altitudes. It comprises 300 non-manoeuvrable 3U cubesats so is much smaller in both total areal cross-section and aggregate mass [14].

[7] G. Popkin, *Technology and Satellite Companies Open up a World of Data*, https://www.nature.com/articles/d41586-018-05268-w.

[8] *See* US lifts restrictions on more detailed satellite images, BBC, http://www.bbc.com/news/technology-27868703.

[9] https://www.itu.int/en/ITU-T/about/groups/Pages/sg17.aspx.

While these developments can potentially enable governments to better track down criminals and monitor suspicious behaviours,[10] they also bring privacy concerns. Such disruption carries the likelihood of invasive uses and identification risks,[11] as Ray Purdy claims [11]. Privacy and data protection concerns, raised by satellite imagery, are similar to those related to drone photography and closed-circuit television, and will become heightened as technology improves. Indeed, if high-resolution images fall into the hands of the wrong people, we shall face an ever-growing inventory of privacy and security problems. Analysing privacy and data protection risks is necessary to inform future regulations and satellite data policies towards General Data Protection Regulation[12] compliance (henceforth GDPR), in force since last May [12].

The article is organized as follows. Section 2 provides for the legal framework conveyed by space law and its relation to privacy. Section 3 discusses how big data analytics and space data relate. Section 4 defines what personal data in big space data is, and it depicts some of the privacy, data protection and ethical risks. Section 5 refers to the democratization of space data. Section 6 suggests mitigation risk approaches. Finally, section 7 concludes the article by summarizing the findings and proposals for the future development of the law.

2 LEGAL BACKGROUND: OUTER SPACE LAW PRINCIPLES AND PRIVACY

Space imagery is subject to an evolving legal regime, which has been built up over time in step with the progress of technology and the growth of its economic importance.

This legal regime is based on the principles of *freedom of access* to space and *freedom of use*, enshrined in the Outer Space Treaty (OST) of 27 January 1967 (Article I). These principles are supported by a third one: that of *non-appropriation* (OST, Article II), which prevents a state from making a claim or exercising jurisdiction over all or part of outer space, even if the provisions of Article II relate mostly to celestial bodies rather than to data produced by satellites. This results in the freedom for the operator of an earth observation satellite to observe the latter from space, whatever the technique used and market them as needed. An earth observation satellite is certainly under the jurisdiction of a State – the

[10] http://thescienceexplorer.com/technology/new-satellites-will-detect-your-face-and-phone-space.

[11] http://www.digitalethics.org/essays/high-resolution-satellites-are-our-privacy-expectations-too-high; https://www.forbes.com/sites/patrickwwatson/2018/04/26/this-is-the-end-of-privacy-as-we-know-it/#27d88ee96875.

[12] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 Apr. 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data. All the articles mentioned in this paper relate to the GDPR.

launching State which has registered it – but its functioning cannot come up against a claim of sovereignty on the part of a State.

It follows from Article VI that each State has an obligation to supervise space-related activities of commercial actors and is held directly responsible for any activities of its agencies in space. As such, the commercial market of high-resolution imaging will ultimately rely upon the supervision and responsibility of the respective States.

Article II of the Liability Convention imposes absolute liability on a State '*for damage caused by its space object on the surface of the Earth*.' Article I defines 'damage' as meaning 'loss of life, personal injury or other impairment of health'. It is asserted that this definition not only includes physical injury, but also encompasses harm to mental as well as social well-being. Such an interpretation is based on the World Health Organization's definition[13] of 'health' as '[A] *state of complete physical, mental and social well-being and not merely the absence of disease or infirmity*'. Pursuant to this reasoning, a violation of an individual's privacy right can potentially be construed as a personal injury under the Liability Convention. Since the injury was caused by a satellite and it occurred on Earth, the Liability Convention imposes absolute liability for the injury.

The Resolution 41/65 on the Principles of Remote Sensing, adopted on 3 December 1986 [9], specifies the regime for observing land by recognizing the right of the State observed to access primary data and data processed in its territory. Principle XII further states that '*the observed State shall also have access to the analyzed information available in the territory under its jurisdiction which is in the possession of any State participating* in rem*ote sensing activities without discrimination and under the same conditions, with due regard for the needs and interests of developing countries'*. This access is not free; it is provided under 'reasonable price conditions'.

A growing number of national laws[14] now complement this international mechanism to foster the development of a space imagery market, while ensuring the protection of the interests of the States concerned. This is particularly the case of US or Canadian legislation, or the French law of 3 June 2008, the latter setting up a flexible system of reporting to the General Secretariat of Defense and National Security of any primary data operator. The US law is characterized by more stringent provisions, subjecting operators to a system of prior authorization and organizing a shutter control for the benefit of the Department of Commerce, which can thus impose on licensed operators limitations on the collection and the

---

[13] Preamble to the Constitution of the World Health Organization, reprinted in Final Acts of the International Health Conference, U.N. Doc. E/155, at 11 (1946).

[14] Please refer to our electronic data-base of national space legislations (www.spacelegaltech.com), also accessible from SIRIUS website (www.chaire-sirius.eu). Adde: LRapp, *Space Law Making*, The Space Review, 2 July 2018, Comments.

distribution of data. This *shutter control* is not exclusive of legal mechanisms more respectful of the requirements of the operation of commercial enterprises. For instance, a licensed operator can also negotiate with the United States Government a contract for the supply of data collected on a commercial basis.

More specifically, the contracts concluded by operators of satellite systems constitute an autonomous layer of legal obligations, setting the conditions for the provision of observational data and, especially, intellectual property (IP) rights.

None of these laws or contracts, however, deal with questions of privacy [17], nor even consider risks that could result from significant progress made in the resolution of satellite imagery. This role, for now, is fulfilled by the General Data Protection Regulation[15] which caters the definition of personal data.

3 BIG DATA ANALYTICS

Among others, two main trends are currently developing in the satellite imagery industry: (1) the increasing availability of very high-resolution satellite imagery; and (2) processing-intensive image analysis tasks. Orbital Insight, SpaceKnow, Descartes Labs, Exogenesis, Remote Sensing Metrics, OmniEarth, DataKind are examples of analytic support companies offering actionable insights or intelligence.

The significance of space big data is a direct outcome of the value, both scientific and commercial, that this data is able to produce [10], through a multi-modal pipeline of advanced data analysis methods called big data analytics [15] comprising content analytics crawlers (mining unstructured content), machine learning (ML) algorithms for image analysis, natural language processing tools (NLP), and data mining techniques (DM). Some of the aspects of big data analytics are briefly mentioned to foresee its (potential) implications for privacy and data protection [7]:

(a) The use of large numbers of ML algorithms deployed against image data to analyse and process high-resolution satellite imagery and for having precise visual representations, in order to find automatic correlations and inferences from datasets. New information solutions such as the ones offered by new entrants like HawkEye 360 to help monitor global activity across the air by using radiofrequency technology should accelerate this trend. Algorithms are hence 'trained' to analyse the imagery automatically, based on a small

[15] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 Apr. 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, CELEX:32016R0679.

subset of images that has been manually classified beforehand. Once relevant correlations are identified (originally unforeseen), a new ML algorithm can be created and deployed to specific cases in the future. The new algorithms are autonomous, self-learning, and self-managing. Moreover, Target Matching Recognition algorithms for satellite images are improving robustness and accuracy, tackling image-matching errors and reducing matching recognition time [1].

(b) The tendency, among operators, to collect/analyse *all* the data that is available;
(c) The repurposing of data for purposes other than those for which it was originally collected, as analytics can mine data for new insights and find correlations between apparently disparate datasets; and
(d) The use of new types of data automatically generated and coming from IOT devices, as sensors.

As a consequence, there is a growing use of face-based technology systems in commercial settings which companies as Planet, Orbital Insight, Capella Space are taking into consideration to diversify their respective service offering. At the same time, the technology often involves the collection and use of special categories of data, requiring careful assessment of identifiability and privacy issues raised. There are various levels[16] and distinctions of facial scanning technology. They stem from:

(a) facial detection systems, e.g., counting customers in line, in stores, in amusement parks, etc., which, when properly designed, neither create nor implicate any personally identifiable information and no privacy concerns;
(b) facial characterization, in which characterization technologies can inform businesses whether individuals are smiling or frowning, male or female, and old or young. It does not raise identifiability concerns, while there is a possibility of discrimination based on gender, race, and/or other characteristics;
(c) unique persistent identifiers, where it is potentially identifiable if linked to other data and entails privacy concerns (detailed profiling, tracking); and finally,
(d) full-scale facial identification programs matching a person's image to a database in order to identify the individual to someone, who otherwise wouldn't be possible to recognize them, e.g. verification 1:1 and verification 1:many, raising identifiability and privacy concerns.

[16] https://fpf.org/wp-content/uploads/2018/09/FPF_FaceRecognitionPoster_R5.pdf.

A series of other applications and payloads can also be installed on LEO smallsats, allowing the gathering and processing of personal data and seriously interfering with, and potentially violating citizens' rights to privacy and data protection. Examples of such technology include high-power zoom, facial recognition, movement detection of people and objects (or 'patterns of life'), license plate recognition, thermal sensors, radar, see-through imaging, Wi-Fi sensors, GPS systems, systems to read IP addresses and track RFID devices, and systems to intercept electronic communications.

Big data analytics endow entities with fine-grained data to extract value, trends, and patterns to create competitive-edge, thereby enabling customizing and leverage products and services – predictive capabilities from its combination create the potential for well informed decisions and accustomed services. Increasingly higher-resolution satellite images linked with this significant challenge for data analysis pushes the limits on bandwidth, storage, computing power, and the statistical expertise of most organizations.

Technology companies are developing new tools to perform image processing on distributed computing platforms. An example of such a tool is Google Earth Engine, which provides access to both the imagery and the platform to analyse it. Other competing platforms are likely to emerge. These methods also increase risks of hampering privacy and data protection [5]. The potential imposition of risks in the space industry are analysed in the next section.

4 DATA PROTECTION, PRIVACY AND ETHICAL RISKS OF SATELLITE IMAGERY

The territorial scope of the GDPR on space data[17] (Article 3(1)) states that any entity that directly or indirectly collects or processes data of EU residents is subject to the GDPR, even if taken from a satellite under the jurisdiction and control of a non-EU country.[18] In light of the law's broad scope, it is relevant to assess what is personal data in big space data, the nature and extent of data collection and processing that takes place in the remote sensing industry, and finally, the potential risks that these images hold.

---

[17] Within the satellite industry, the GDPR will have an impact on Direct-to-Home (DTH) broadcasting, satellite telecommunications services and geolocation services.

[18] It is argued that a transfer of European personal data from a satellite under the jurisdiction and control of the US to a data controller or processor located in the US is a cross boarder transfer GDPR-based, even if none of the ground stations that receive the data are physically located in Europe [3].However, this data cross boarders get complicated. Stefoudi mentions '*satellite signals,* e.g. *are transmitted within fractions of seconds among multiple satellites in-orbit, ground stations, databases, and all sorts of electronic devices. The variety of data available and the speed, in which they are transmitted, along with the numerous ways of processing, create a situation where the data subject, the data analyst and the final product are hard to distinguish and locate'* (Art. 29 Working Party Opinions – henceforth WP-, 243/2016).

4.1. What is personal data in big space data?

A large proportion of big data is not personal, namely, weather information, satellite imaging, and operational machine data. But some space big data may include elements that link directly to a person, and hence, could be considered personal data.

The GDPR regulates the use of multiple data formats – including images – which help to identify, either directly or indirectly, any person. Personal data is therein broadly defined, and means:

'*any information relating to an identified or identifiable natural person ("data subject")',* (Article 4 (1)).

The Article 29 Working Party[19] (WP29) 'Opinion on the Concept of Personal Data' (WP136) decomposes the definition of personal data in three main building blocks:

(a) *any information*
(b) *relating to*
(c) *identified or identifiable natural persons*.

We will analyse these three constituents separately, following the cognition of the WP136 to argue that it is possible to identify natural persons from space imagery.

(a) '*Any information*' means that the concept of personal data includes any sort of information. From the point of view of the *nature* of the information, it covers 'objective' information, subjective information, opinions or assessments. For information to be personal data, it is not necessary that the object or a person depicted in a picture is true or proven; this definition also holds for incorrect information. From the point of view of the *content* of the information, the concept of personal data includes data providing any sort of information. Considering the *format* in which that information is contained, the concept of personal data includes information available in whatever form, as images, insofar as they may represent information regarding an individual.
(b) '*Relating to*' building block helps to find out which are the relations/links that are important. In general terms, information can be considered to 'relate' to an individual when it is *about* that individual, e.g. the image of a person captured or filmed is considered to be about that person. In some situations, the information conveyed by images concerns objects rather

[19] The opinions of the WP29 are not formally binding, but possess 'persuasive authority' on this domain [26].

than individuals. As the Article 29 Working Party Opinion elaborates, '*those objects usually belong to someone or may maintain some sort of physical or geographical vicinity with individuals or with other objects. It is then only indirectly that it can be considered that the information relates to those individuals or those objects'*

(c) '*Identified or Identifiable natural persons*'. In general terms, a natural person can be considered as '*identified*' when, within a group of persons, he is 'distinguished', known, or 'singled-out' from all other members of a group. A person is '*identifiable'* when, although the person has not been identified yet, it is possible to do it.

In the following section, we shall discuss an issue related to the third element – the possibility of identifying individuals through very high-resolution images (VHR) satellite imagery.

4.1.1 *Potential Identification*

Direct identification is normally achieved through direct or unique identifiers which have a particularly privileged and close relationship with the particular individual.

Article 4(1) articulates how identification is enabled through identifiers:

'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'.

Hence, the phrase 'direct and unique identifiers' covers data types which can be easily referenced and associated with an individual, including descriptors such as a name, an identification number or username, location data, phone numbers, online identifiers, or an image, in combination with additional information, if the image is not unique.

Persons may be associated with online identifiers '*provided by their devices, applications, tools and protocols, such as internet protocol (IP) addresses, cookie identifiers or other identifiers such as radio frequency identification tags. This may leave traces which, in particular when combined with unique identifiers and other information received by the servers, may be used to create profiles of the natural persons and identify them*', as Recital 30 expounds.

An individual is *indirectly identifiable* by *combinations* of indirect (and therefore non-unique) identifiers that allow an individual to be singled out. They are less obvious information types which can be related to, or 'linked' to an individual, such as, for instance, physical, physiological, genetic, mental, economic, cultural or social elements of a person.

It is not currently possible to directly identify an individual´s face using today's satellites. The resolution does not suffice to depict optical characteristics of a person's features. Satellite imaging consists of coarse resolutions that do not typically allow for recognition of individual's faces, and they tend to image structures and features that are themselves publicly viewable.

Arguably, a person – as a whole – can be depicted on these pictures, as for the resolution might allow for the indirect identification of a person, considering, for example, the person's height, body type, and clothing, which configure physical, physiological economic, cultural or social elements of that person, as article 4(1) *in fine* refers to.

Likewise, objects and places (location data) linked to a person could also enable identification of a person via VHR, such as the person's home, cars, boats, and other property.

WP136 and Recital 26 provide for a twofold standard to ascertain whether a person is identifiable. Pursuant to this standard, a person is identifiable when:

(a) '*all the means of identification are 'reasonably likely' to be used to identify an individual*
(b) *either by the controller or any other person'*.

A mere hypothetical possibility to single out the individual is likely not enough to consider the person as identifiable. To assess such possibility, it is needed to take into account *objective factors* such as:

(a) The cost and time required for identification in light of new technology, security developments, or changes to the public availability of certain records;
(b) Available tools for identification;
(c) Risk of organizational dysfunctions, e.g. breaches of confidentiality duties, technical failures;
(d) The purpose pursued by the data controller in the data processing (e.g. satellite surveillance of a specific inhabited area, in which identification is argued to only happen in a small percentage of the material collected, entails the purpose of processing of location data);
(e) State of the art in technology at the time of the processing, and the possibilities for development during the period for which the data will be processed.

Regarding the latter factor, the WP136 perceives that identification may not be possible today, with all the means likely reasonably to be used today, as the case of current very-high resolution satellite images. However, it they are intended to be kept for five years, the controller should consider the possibility of identification

that may occur also in the fifth year of their lifetime, and which may make them personal data at that moment.

The capacity of identification is increasing at a rapid pace as technology develops. The growing number of throughput satellites with increasing reliance on satellite technology, coupled with new analytical technologies, extends the possibility for its processing[20] and identification.

For instance, if the footage taken through VHR imaging only shows the top of a person's head and one cannot identify that person without using sophisticated means, it is not personal data. However, if the same photograph were taken in the backyard of a house, with additional imaging analytical algorithms that may enable an identification of the house and/or the owner, that footage would be considered personal data. Thus, personal data is very much context-dependent.

In fact, this scenario escalates with the advances of 'ultra-high' definition images[21] being published online, from commercial satellite companies, and the consequential application of big data analytic tools. It might be possible to identify *indirectly* an individual (and to depict individual households, etc.), when high-resolution images are combined with other spatial and non-spatial datasets.

Thus, while the footage of people may be restricted to 'the tops of people's heads', once these images are contextualized by particular landmarks or other information, they may become identifiable. This other information can include 'demographically identifiable information' (DII) or 'community identifiable information' (CII), which may contain personal information, or other datasets regarding transportation, administration, demographics categories, survey data available online, or other imaged information (geo-tagged or otherwise identifiable by location, and crowdsourced geographic information [34]) (WP 7/2003; 3/2013; 6/2013).

While apparently innocuous, not privacy-affecting or *anonymised* sources [8], such combination of datasets may enhance a jigsaw of indirect correlation of identification of grounded-based objects and individuals.

'*The combination of publicly available data pools with high resolution image data, coupled with the integration and analysis capabilities of modern Geographic Information Systems (GIS) disclosing geographic keys such as longitude and latitude, can result in a technological invasion of personal privacy'*[1].

---

[20] Data processing includes accessing, collecting, recording, organizing, structuring, storing, adapting, altering, retrieving, consulting, using, disclosing by transmission, disseminating (or otherwise making available) personal data. The use of satellites for any of these operations will consist in processing of personal data. As a result, collecting satellite data – where personal data may be *inferred* – e.g. images of people, licensed plates, transmitting location data of an individual – is likely to constitute 'processing' for the purposes of the GDPR.

[21] https://www.offthegridnews.com/privacy/googles-newest-high-res-satellites-can-monitor-your-every-move-in-real-time/.

Moreover, even when this data has been aggregated and *pseudonymized* to remove explicit identifiers and tags, machine learning algorithms applied to very large datasets renders it technically possible to reidentify a person (WP 05/2014). Thereby privacy-relevant new facts, set through identification, intrinsically abides to legal requirements, as identification not only means the possibility of retrieving a person's name and/or address, but also includes potential identifiability by singling out, *linkability* and inference (WP 05/2014, [6]). As data collected by the ubiquitous computing is, presumably, personal data (WP136), the processing of non-sensitive data can lead, through data mining, to data that reveals personal information, thus, blurring the conventional categories of data.

One can argue, at this point, which and how many sources of data are needed for an ideal combination of data in order to achieve that high level of accuracy required by the relevant applications (such as to enable profiling and identification).

4.2. The risks of ultra-high resolution of satellite imagery

The marrying of visual data with refined image algorithms, high-speed image refresh re-scan rates, and down-to-the-pixel detail can create different risks. 'Harmless' intelligence could be combined to create personal data and hence responsibilities and burdens follow. The danger is that the use of this personal big data of up-to-date high resolution pictures broadcasted online can lead to several privacy, data protection and ethical risks. We acknowledge that these risks motivate the integration of privacy within remote sensing principles.

Even if VHR satellite missions are focused on objects, or on a piece of property – rather than people – they may capture information 'relating to individuals' accidentally, inadvertently or in unusual circumstances. Such possibilities of capturing data 'relating to' data subjects, turns this data into personal data, regardless of its nature or content.

The hyper-connected world of ubiquitous computing provides fertile ground for such accidental impact, for unpredictability of unintended outcomes is considered one of the characteristic features of the advanced data analytics [26].

In addition, even if multiple individuals are targeted (and not one single person), such generalization considers them as members of a group – either 'ad-hoc groups' (whose link is based on a third party interest), 'ascriptive groups' (incidentally developed characteristics), and 'collective groups' (explicit shared traits) [29].This strand raises the question whether privacy should be protected at an individual level (each data subject) or in a group level [29]. It should be kept in mind here that the GDPR protects data regarding individuals (and legal persons), but not groups. Floridi [32] conversely argues that people are not targeted as

individuals, but instead as a member of a specific group and privacy protection is owed to the group:

[I]t is the group, not its members, that is correctly identified as the right-holder.

Several scenarios can be anticipated: scanning the landscape, inspection thereof, capturing images of buildings, cars, real estate showcasing, stock image production, production of footage for publicity purposes, and the like. Those familiar with the area and/or familiar with the individuals who may be in the vicinity may be able to identify them. These privacy risks are explored in the following section.

4.2[a] *Privacy Risks*

The right to privacy protects the secrecy of personal information, as well as the different facets related to the private sphere of each individual against external intrusions. Personal privacy must be safeguarded in order to protect a person's right of self-determination with regard to one's body, sexual orientation, relations with others, construction of one's own identity, etc., regardless of whether data is collected. The collection of high-resolution satellite imagery in a public space may interfere with privacy, in breach of Article 8 of the European Convention on Human Rights (ECHR), in the following circumstances: (1) when satellites monitor and record data in a systematic and permanent way, regardless of whether the monitoring is covert or overt; (2) when high-resolution satellite operators distribute images of someone previously collected; or (3) when these operators do not record images, but monitor a public space through 'sophisticated' means.

Privacy risks that flow from the use of such satellite technology are various in numbers:

(a) *Transparency and (in)visibility*. This risk applies when individuals on the ground may not know VHR satellites are in operation, and if they do, may be unsure about who is operating them and the purpose for which it is being used, causing discomfort somehow.
(b) *Function creep*: This risk occurs when the purposes of VHR usage expand, either to additional operations or to additional activities within the originally envisaged operation. This risk also arises when such imagery is disseminated in the internet and there is a risk for it to be reused widely.
(c) *Privacy of location and space* [23][24]: this risk hampers the right of individuals to move in their own home (yards and gardens) and/or other public or semi-public places without being identified, tracked or monitored by satellite images or video [28].

(d) *Privacy of association*: This refers to the freedom of people to associate with others [28]. It is related also with the fact that footage might indicate, for example, the number of adults living in a household (based on the number of vehicles) or the clues as to their relationships.

4.2[b] *Data Protection Risks*

In addition to the privacy risks described above, the following additional risks to data protection could also be taken in to account:

(a) *Lack of transparency*: Transparency of data collection requires that the data controller notify the data subject of the personal information collected, the purpose of that collection, and use of the data, as well as details of the VHR operators to enable the data subject to exercise their rights of access, correction, and erasure.
(b) *Data minimization and proportionality*: considering the collection capacity of high-res earth imagery, in substance, space technology entails the tendency of extensive collection, aggregation and algorithmic analysis of all the available data for various reasons, which hampers the data minimization principle, Article 5 (1)(c). In addition, irrelevant data is also being collected and archived, undermining the storage limitation principle, Article 5 (1)(e).
(c) *Purpose limitation and repurposing of data*. As data analytics can mine data for new insights and find correlations between apparently disparate datasets; hence, automatic capture of big data can be mostly reused for secondary unauthorized purposes, profiling, undermining the purpose specification principle, which convenes that the purpose for which the data is collected must be specified and lawful, Article 5 (1)(b).
(d) *Accuracy*. Results drawn from data analysis may not be representative or accurate if its sources are not accurate. Machine learning itself may contain hidden bias, which leads to inaccurate predictions and even profiles about individuals. Hence, high-resolution images need to be validated (on the ground) to ensure the trustworthiness of a given interpretation and avoid interpreting an image incorrectly. Indeed, '*at best, satellite images are interpretations of conditions on Earth – a 'snapshot' derived from algorithms that calculate how the raw data are defined and visualized*'.[22] For example, one recently developed algorithm is designed to identify artillery

[22] Melinda Laituri, 2018,https://theconversation.com/satellite-imagery-is-revolutionizing-the-world-but-should-we-always-trust-what-we-see-95201.

craters on satellite images – but the algorithm also identifies locations that look like craters but are not.

4.2[c] *Ethical Issues*

A final category of risk that arises from the use of high-resolution satellite imagery contains that risks that are of an ethical nature:

(a) *Discrimination*. While profiling is used as 'pattern recognition, comparable to categorization, generalization and stereotyping' [30], and is therefore used to gain information to generalize it for multiple individuals, VHR imagery combined with analysis technologies *can* lead to discriminatory profiling [5]. Also, this refers to the fact that VHR usage (and the potential privacy and data protection impacts) may be more prevalent in relation to certain populations or areas which are less likely to be able to effectively voice or act upon those concerns (e.g., marginalized populations or areas).With the use of ML and data mining, individuals might be clustered according to generic behaviours, preferences and other characteristics, even without being identified [31]. Profiling ultimately involves creating derived or inferred data, occasionally leading to incorrect and biased decisions (discriminatory, erroneous and unjustified, regarding for instance, their behaviour, health, creditworthiness, recruitment, insurance risk, etc.) [25].

(b) *Public dissatisfaction*: This refers to the possibility that people could become disillusioned with VHR imagery use based on the possibility that they are compromising privacy and data protection rights or that they are feeling 'over-run' by such technology.

(c) *Chilling effect*: This refers to situations where individuals might be unsure about whether they are being observed, even if no VHR satellites are operating, and they hence attempt to adjust their behaviour accordingly [28].

(d) *Imbalance*. In certain situations, space technologies might produce situations of imbalance, where data subjects are not aware of the fundamental elements of data processing and related consequences, being unable to negotiate their information, which leads to a side consequence of enhanced information asymmetry. Even exercising the right to be forgotten seems hard to apply. Photos captured for use in Street View may contain sensitive information about people who are unaware they are being observed and photographed [20].

5 DEMOCRATIZATION: PROMISES OF LOW-COST, HIGH-VOLUME SATELLITE DATA

Open access and open dissemination policy are frames of the 'Space Strategy for Europe' (COM(2016)705 final) [20], and hence, access to satellite imagery are being largely democratized. Public access to imagery holds the promise to democratize this data stream, ensuring that civil society and public advocacy groups have the opportunity to analyse and use information from satellites alongside governments and corporations. Sectors as diverse as agriculture, deforestation, crisis mapping, and human rights protection can benefit from satellite data analysis.

Tools such as Google Earth Engine, Earth Explorer (EE) (https://www.usgs.gov/earthexplorer-0) Earth Explorer, Google Earth Outreach, etc., are enabling access. Skybox, Planet Labs, UrtheCast and other start-ups promise of low-cost, high-volume satellite data. Competitors such as GeoEye, DigitalGlobe, RapidEye, and Spot foster access to space imagery.

Orbital Insight, SpaceKnow, Descartes Labs, Exogenesis, Remote Sensing Metrics, OmniEarth, DataKind are examples of analytic support companies offering insights or intelligence that enhance data analysis.

As they compete for market shares, the costs of high-resolution imagery are likely to decrease. These companies have shown interest in collaborating with NGOs and academia by providing imagery for free or at reduced cost [2].

6 MITIGATION RISK APPROACHES

International space law (comprised primarily of the OST together with its follow-on treaties) currently does not address privacy concerns and, consequently, contractual provisions and policies from remote sensing satellite operators do not address these issues either [17]. Satellite imagery policies need to be formulated focusing on image distribution, and on the level of access given to private corporations. Such policies need to be related to each technology, to common use-cases, benefits, concerns, and risks. National legislation could be the dedicated *locus* for regulating privacy regarding space imagery.

Imagery access, analysis, and dissemination must be consistent with the 'Common European Data Space' (SWD(2018)125 final) [22] and in furtherance of the GDPR, as it happened with the use of drones [4].

Given the rapidly changing technology and the big space data context, it is advisable that privacy issues be considered at every stage of a dataset´s life cycle, and not only to the point of selling, which means that satellite operators capturing images need to account for the data protection by design principle (Article 25(1)). This principle refers that:

> *the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organisational measures, such as pseudonymisation, which are designed* ***to*** *implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects.*

Release of datasets of images that might raise potential privacy issues might call for a special regime of licensing [27] that restrict their use to certain contexts (e.g. non-commercial), or that prohibit activities aimed at re-identification. Nevertheless, such licensing terms would depend on compliance by its users and on the data provider's legal action when breaches occur.

Information and transparency protocols, both on the missions and the operators, should be devised and implemented, as well as codes of conduct (by industry groups of remote sensing satellite operators) with recommended practices for big space applications, or guiding on the different categories of data that require special care.

Data controllers can proactively carry out a data protection and privacy impact assessment processes,[23] notably where there are risks for data protection and privacy, respectively, according to typical VHR scenarios, e.g. this requires defining the purpose of the use; choosing the right tools; using the most privacy friendly approaches, or privacy-aware analytics methods; and ensuring the security of the data collected. These processes require that before using a privacy-limiting device, means must be in place to limit the impact as far as possible.

A dialogue with manufacturers could be envisioned to pre-emptively implement privacy by design and by default measures and embed data protection requirements in data space applications[24] to ensure compliance from the outset.

Remote sensing companies can set up mechanisms to automatically process images by blurring faces, filtering out or obscure identifiable features on, house holding, whenever identification scenarios occur due to forthcoming image improvement.

The Remote Sensing Principles [9] contain no specific restrictions on what may be observed, give no veto rights to a sensed state or entity, put no operational conditions on the sensing, and, in general, do not provide any useful guidance on privacy. The focus of the Principles is on the interests and rights of states to sense

---

[23] https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/data-protection-impact-assessments/.

[24] Instantiations of space data applications are, e.g. inferred traffic patterns in large cities that can be used to better design future cities; maritime transportation activity and logistics can be tracked; foot traffic patterns at retail locations can be analysed to determine consumer behaviour; and farmers can better understand what factors influence the growth of crops, https://www.satellitetoday.com/innovation/2019/03/08/turning-space-data-into-smart-insights/.

and be sensed, not on the rights of individuals [3].Perhaps the future holds an update of these principles which would expand the scope of the Principles to encompass risks to an individual's privacy and data protection.

7 CONCLUSIONS

High-resolution remote sensing data is a key element of the value chain of the space sector. The opportunities provided by increasingly higher-resolution satellite images pose a significant challenge for big data analytics.

One cannot always predict what the forthcoming usages of VHR images will be, given the myriad tools and technologies available. As satellite images become more ubiquitous – providing mankind with god-like views from above – reflection on how they are created and the purpose for their use is timely.

Privacy and data protection rights are not absolute, and they are balanced against other competing public interests, such as transparency and the 'right to know' or the 'access to information' rights mandated in many countries.

There is a blurred line between the sharing of high-resolution images for the welfare of public safety and human advancement, and the pillars of privacy and data protection. A fair compromise will need to be considered on the further tracking of the changing human footprint across the globe.

The GDPR is intended to be broad enough to capture new forms of technology that may affect privacy and data protection. Satellite-based imaging companies (and those who use and process images from such satellites) need to be aware of the scope and jurisdiction of the regulation, as it is very likely to be applicable to their data operations.

The processes and lifecycles around data collection and management of space data need to be examined since fines for non-compliance are significant – up to EUR 10 million, or 2% of the worldwide annual revenue of the prior financial year, whichever is higher, shall be imposed for infringements.

REFERENCES

[1] S. Chun & V. Atluri, *Protecting Privacy from Continuous High-Resolution Satellite Surveillance*, in *Data and Application Security,* IFIP 73 (Thuraisingham B., et al. eds, Springer 2002).

[2] P. Baylis, G. Kroll & T. Madon, *Micro-satellite Data: Measuring Impact from Space*, Goldilocks Deep Dive (2014).

[3] B. Cohen, *Remote Sensing and the New European General Data Protection, Proceedings of the International Institute of Space Law 2017*, 415 Eleven Int'l Pub. (2017).

[4] Directorate General for Internal Policies Policy Department C: Citizens' Rights and Constitutional Affairs, *Privacy and Data Protection Implications of the Civil Use of Drones* (2015).
[5] ICO, Big Data, Artificial Intelligence, Machine Learning and Data Protection, UK, 2017
[6] P. Leonard, *Customer Data Analytics: Privacy Settings for Big Data Business*. 4(1) Int'l Data Privacy L., Oxford 53–68 (2014).
[7] A. Mantelero & G. Vaciago, *The 'Dark Side' of Big Data: Private and Public Interaction in Social Surveillance*, 14 Computer L. Rev. Int'l 161–69 (2013).
[8] P. Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57(6) UCLA L. Rev., Los Angeles 1701–77 (2010).
[9] Principles Relating to Remote Sensing of the Earth from Outer Space. G.A. Res. 41/65, U.N. Doc. A/RES/41/65, 1986.
[10] Proceedings of the 2017 conference on Big Data from Space, EU Publications, 2017
[11] R. Purdy, 'Ruling on sharper satellite images poses a privacy problem we can no longer ignore' (2014), https://theconversation.com/ruling-on-sharper-satellite-images-poses-a-privacy-problem-we-can-no-longer-ignore-28133
[12] D. Stefoudi, 'Space Big Data, Small Earth. Laws: Overcoming the Regulatory Barriers to the Use of Space Big Data Applications', Proc. of the 2017 conference on Big Data from Space (BiDS'17), EU Publications, 2017.
[13] D. Stefoudi, *Space Big Data: Big Data Troubles in the Final Frontier* (2017), https://leidenlawblog.nl/articles/space-big-data-big-data-troubles-in-the-final-frontier
[14] Swiss Re Report 'New space, new dimensions, new challenges: how satellite constellations impact space risk' (2018).
[15] K. Waterman, P. Bruening, *Big Data Analytics: Risks and Responsibilities,* 4(2) Int'l Data Privacy L., 89–95 (2014).
[16] F. Von Der Dunk, *Europe and the 'Resolution Revolution': 'European' Legal Approaches to Privacy and their Relevance for Space Remote Sensing Activities*, Space & Telecomm. L. Program Fac. Publications 810 (2009).
[17] F. Von der Dunk, *Outer Space Law Principles and Privacy*, in *Evidence from Earth Observation Satellites: Emerging Legal Issues*, 243–58 (Denise Leung & Ray Purdy eds, Leiden: Brill 2013).
[18] European Space Policy Institute, *Current Legal Issues for Satellite Earth Observation* 38 (2010).
[19] Y. Chen, W. Wei Xu et al., *Target Matching Recognition for Satellite Images Based on the Improved FREAK Algorithm*, 2016 Mathematical Prob. in Eng'g 2016.
[20] President's Council of Advisors on Science and Technology, Big Data and Privacy: A Technological Perspective, Report to the President (May 2014)

[21] Space Strategy for Europe, 2016 (COM(2016) 705 final), Communication
[22] Towards a common European data space (SWD(2018) 125 final) Communication
[23] H. Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life*, 70–72 (Stanford, CA: Stan. L. Books 2010).
[24] D. Solove, *Understanding Privacy*, 24–29 (Cambridge, MA: Harvard University Press 2008).
[25] L. Edwards, *Privacy, Security and Data Protection in Smart Cities: A Critical EU Law Perspective,* 2 Eur. Data Prot. L. Rev., Berlin 28–58 (2016).
[26] N. Purtova, *The Law of Everything. Broad Concept of Personal Data and Future of EU Data Protection Law*, in *Law, Innovation and Technology* (Routledge 2018).
[27] F. Borgesius, L. Gray et al., *Open Data, Privacy, and Fair Information Principles: Towards a Balancing Framework*, 30(3) Berkeley Tech. L.J. 2073–30, 2097 (2015).
[28] R. L. Finn, D. Wright et al., *Seven Types of Privacy*, in *European Data Protection: Coming of Age* 16 (S. Gutwirth, R. Leenes, P. de Hert, Y. Poullet eds, Springer, Dordrecht 2013).
[29] B. Mittelstadt, *From Individual to Group Privacy in Big Data Analytics*, in 30 Philos. Technol. 475 (2017).
[30] M. Hildebrandt & S. Gutwirth. *Profiling the European Citizen* (Springer 2008).
[31] B. van der Sloot, *Privacy in the Post-NSA Era: Time for a Fundamental Revision?,* 5(1) J. Intell. Prop., Info. Tech. & E-Commerce L. 2014.
[32] L. Floridi, *Open Data, Data Protection, and Group Privacy* 1–3 (*Springer* 2014).
[33] Future of Privacy Forum, *Location Data: GPS, Wi-Fi, and Spatial Analytics*, https://fpf.org/wp-content/uploads/2018/12/DDF-2-Materials.pdf
[34] P. Mooney, Olteanu-Raimond et al., *Considerations of Privacy, Ethics and Legal Issues in Volunteered Geographic Information*, in *Mapping and the Citizen Sensor* 119–35 (Giles Foody, Steffen Fritz, Linda See eds, London: Ubiquity Press 2017).
[35] C. Santos, D. Miramont & L. Rapp, *High Resolution Satellite Imagery and Potential Identification of Individuals*, Proc. of the 2019 conference on Big Data from Space (P. Soille, S. Loekken, & S. Albani eds, BiDS'2019), EUR 29660 EN, Publications Office of the European Union, Luxembourg, 237–40 (2019).