Promoting Responsible Research and Innovation in Data-intensive Technologies: a Societal Impact Methodology

Gemma Galdon, Leticia Duboc and Iris Huis in 't Veld

Eticas Research & Consulting

Barcelona, Spain

Abstract

The development of security and surveillance technologies produces externalities, both positive and negative. But how can these be measured and how can impacts on different stakeholders—such as citizens—and their perspectives be taken into account?

Anticipating and mitigating negative externalities in data-intensive systems can be challenging. On one hand, these often become apparent only when technologies have been widely used for a long time. On the other, accounting for such externalities requires a difficult balance between the values, needs and expectations of society and the technical and financial restrictions of technology developers.

Drawing on tools developed elsewhere and combining them with these specific challenges, this paper proposes a four-part societal impact assessment (SIA) methodology for the appraisal of data-intensive technologies that is sensitive both to the technological and economic concerns of engineers and decision-makers and the societal values and the perspectives of citizens.

1. Introduction

Science and technology has been shaping our society for centuries, and vice versa. Historically, an implicit social contract has been established, which underpins the trust relationship between science, technology and society. Under this social contract, science and technology have been expected to produce reliable knowledge and tools, and to be transparent in the communication of their practices with society. In exchange, society has been offering freedom, social licence, as well as funding to invent, innovate and pursue scientific endeavors. This arrangement sets out the expectations held by each other, which includes that existing norms, laws, and standards of conduct are adhered to and that appropriate sanctions can be taken if these are not met.

With time, it has become clear that technology can have unanticipated externalities, both good and bad. Society is becoming increasingly aware that some of these can have profound

consequences, such as the harmful effects of pesticides on the environment (Carson, 1962) or, more recently, the profound consequences of complex financial products on the economy and the welfare state.

Regulation has been modern society's main response to ensure that new technology is in line with societal norms and values (e.g., the protection of health and a clean environment). Through regulation, society enforces specific certain technology can be marked (e.g., in pharmaceuticals and industrial chemicals innovation). However, society has also realised that regulations do not always exist when new technologies or solutions emerge, that when they do they are not always respected, and that the mechanisms to reverse those negative externalities once the social contract has been breached are few and far between.

Therefore, responding to societal problems through regulation has its limits. In particular, regulating too soon means acting when the evidence of undesirable impacts is poor, the case for control is less evident, and the risk of lost benefits in the future is greater. This approach could lead to an innovation plateau. The obvious alternative is the opposite: regulating too late, only when there is sufficient evidence. In this case, the risk is technological lock-in – by the time we decide to regulate something, it is so embedded in society that the cost of doing so is too great and there are too many vested interests involved. Furthermore, on many occasions, the legality of an innovation is not the main issue. Perfectly legal products and services often fail to achieve its full potential or to be adopted at all when society or its target audience resists them because they feel that their values have not been taken into account in the conceptualization of these innovations.

Embedding societal values in innovation and putting in place safeguards for the consequences of scientific and technological developments is thus not trivial. Technical innovation suffers from an information problem – the difficulty to assess impact before it happens – and a power problem – the inability to control the technology once it has been developed and deployed (Collingridge, 1982). Moreover, the increasing intrusiveness and ubiquity of some technological innovations (e.g. biometrics, location technologies, and ambient assisted living technologies) represent a paradigm shift with unprecedented social, ethical and legal concerns. The body and the daily lives of people are becoming the new frontier of technological development and, on several occasions, people may be even unaware that they are interacting with technology. The need to develop assessment methodologies and responsible solutions tailored to these new frontiers is paramount.

Responsibility in innovation is about the governance of unintended consequences as much as it is about the governance of intent. The question thus is, how to conduct research and innovate under these circumstances? Is there any alternative to reacting too early or too late? How do we

embed in our innovation society's moral responsibility towards those living now, those yet to be born, and those beyond our own species?

Responsible Research and Innovation (RRI), although not a solution to all these questions, is a remarkable effort. It recognizes that a responsible innovation is not only about the final product or service being offered, but also about the process that develops these solutions. Embracing RRI means shifting the focus from rules to values. These values need to be derived from a deep understanding of what is and what is not socially desirable, and what ends are legitimate and should be promoted. Values, however, are flexible. They need to be assessed often, and in relation to the aims and means of each new development. Therefore, RRI cannot be a bureaucratic hurdle, or a fixed methodology or tick-box exercise.

Responsible Research and Innovation (RRI) compiles a series of principles that should guide the relationship between the scientific community and society at large and foster public engagement. It is described by the European Union as:

> *Responsible Research and Innovation means that societal actors work together during the whole research and innovation process in order to better align both the process and its outcomes, with the values, needs and expectations of European society (European Union, 2012).*

Alongside the general principles of RRI, we have been working on a Societal Impact Assessment (SIA) framework that responds to this ever-adjusting nature of the relationship between science, technology and society. The framework is deployed during the research and innovation process, and concerns both this process and the final products and services.

It important to note that our framework, which we call the Eticas Framework, is aimed at data-intensive systems. As such, it encompasses not only an analysis of the potential impact of such systems in society, but also practical technical and financial considerations that will support the tasks of engineers and decision makers. Security and surveillance technologies, the focus of this book, are often implemented as data-intensive systems. Therefore they lend themselves to analysis following the framework we suggest.

2. Related Work

Responsible research and innovation has its roots in technology assessment (Grunwald, 2011) – a form of policy research that studies short- and long-term consequences of the application of technology (Banta, 2009). However, early technology assessment held a deterministic account of technology where the people followed the dynamics of the technology to examine how it affected social systems and then provided insights in recurring patterns. The deterministic conception of technology changed with the rise of the social constructivist paradigm,

emphasising that technology can be ‘shaped’ according to social desires and values (Pinch & Bijker, 1987). RRI draws on the knowledge, procedures and methods of technology assessment, extending it to the ethical issues of science, technology and society and governance (Grunwald, 2011).

Approaches for studying the impact of technology differ in rationales and methods. With the rise of RRI and the constructivist paradigm, emphasis is placed on impact assessment to broaden the design practice itself. A remarkable attempt on influencing design practice with impact assessment is Constructive Technology Assessment (CTA) (Schot & Rip, 1997). CTA has three particular analytical achievements (Guston & Sarewitz, 2002): socio-technical mapping, early and controlled experimentation and promoting dialog between innovators and the public. With its focus on co-design and consultation, CTA may fail to capture the complexities of data-intensive technologies which people perceive as obscure ‘black boxes’. The individual and social impact of data-intensive technologies on privacy, for instance, may only become negative in the future, when circumstances change, and this can be difficult to capture using co-design methods. CTA’s methods of socio-technical mapping, public involvement and its attempt to be involved in design practice are contributions to technology assessment we build on. However, the outcomes of CTA are generally not clear in how to proceed because there is a lack of specific action guiding principles for what to do with the social implications identified. In addressing data-intensive systems, and being oriented towards engineering practice, it is of great importance to provide clear recommendations which can be understood by a technically trained audience.

Another well-known type of technology impact assessment is a Privacy Impact Assessment (PIA), which is ‘a systematic process for evaluating the potential effects on privacy of a project, initiative or proposed system or scheme’ (Clarke, 2009, p.1). This kind of assessment is an appropriate tool for identifying and assessing privacy risks of a data-intensive system. However, PIAs do not necessarily take into account that the social implications of such technologies can be broader than just privacy, it is a method that does not question the social desirability of the object of study, and it is an approach that does not always operationalise its findings into specific technical recommendations.

The method we suggest is an attempt to combine useful methods found in other technology impact assessments, making the most of them, with the aim to systematically address all the issues that can contribute to a thorough assessment of a technology’s desirability, acceptability, ethics and legality, as well as the options that exist to mitigate the identified risks. Most importantly, this framework translates this societal impact into practical considerations to be taken into account by engineers and decision makers.

3. Lessons Learned from Data-Intensive Technologies

Responsible Research and Innovation (RRI) is the response to a relative failure – the failure to build effective bridges between technology and society, and provide citizens with safeguards when faced with the consequences of some technological and scientific developments. Below, we briefly review some of these instances of failure.

3.1 Full-body scanners

Full-body scanners were introduced in airports after a terrorist managed to board a flight to Chicago from Amsterdam airport in late 2009. Even though investigations pointed to an intelligent or resource problem (the terrorist had been reported by his own father), the EU's reaction was technological. There seemed to be no doubt about their desirability. However, the public reaction showed that citizens did not share the policy-makers' opinion, and months of campaigning, trials and inconclusive reports led to their widely implementation at the EU being cancelled. Today, they can only be deployed provided that there is no mandatory use for passengers, that they do not take pictures (only representations of the body), and that these are not stored. Such requirements could have been anticipated, if the technology had been guided by proper technology assessments and public scrutiny.

3.2 Electronic Health Records

Electronic patient records hold out the promise of numerous benefits – less paperwork, more time for doctors to see patients, better diagnosis thanks to the availability of the relevant data from different sources, etc. However, efforts to create such records at the national level have more often led to failure than success. In Britain, the government embarked on the largest investment in health information technology in the world – only to shelve it in 2011, with costs to the tune of 10 billion GBP for the British taxpayer. France abandoned its first endeavour to develop Dossiers Médicaux Personnels. In 2010, Germany shelved an electronic card based system, which was three years behind schedule and had cost 1.7 billion euros. While all these cases are slightly different, over-reliance on the possibilities of the technology, inability to address data concerns early on in the process (who owns the data? Who is responsible for mistakes and their consequences?) and a disregard for the difficulties of implementation (trust, training, acceptability) played a significant part in these failures.

3.3 Smart Meters

The use of ‘smart meters’ has been heralded as a dramatic potential improvement in the provision and billing of utilities such as water, gas and electricity. However, smart meter technology’s data gathering requirements and potential for surveillance and remote sensing showcase the unpredictable nature of technological development. When smart meters started being installed in the Netherlands in 2009, the Dutch senate put a stop to their implementation (AlAbdulkarim, Lukszo & Fens, 2012). This followed the Senate’s receipt of a scholarly report by a consumer rights group calling the legality of smart meters. The report also stated that the remote collection of each household’s data in 15 minute intervals put the right to privacy in doubt. Had academic or civil society groups been consulted in the design of this technology, it might not have provoked such public discontent.

3. Eticas' Societal Impact Assessment (SIA)

Externalities produced by technology have to be mapped, and the impacts on different stakeholders and their perspectives must be taken into account. On the basis of our work accompanying the development of technologies in fields as diverse as border and identity management, crisis response, policing applications and forensics, we propose a method for data-intensive systems that is sensitive as much to the technological and economic concerns of engineers and decision-makers as to the societal values and the perspectives of citizens. More precisely, the framework is based on four pillars that combine technological policy and sociological perspectives - Law and Ethics, Desirability, Acceptability and Data Management.

Pillars are organized in three overlapping "levels", as illustrated in Figure 1.

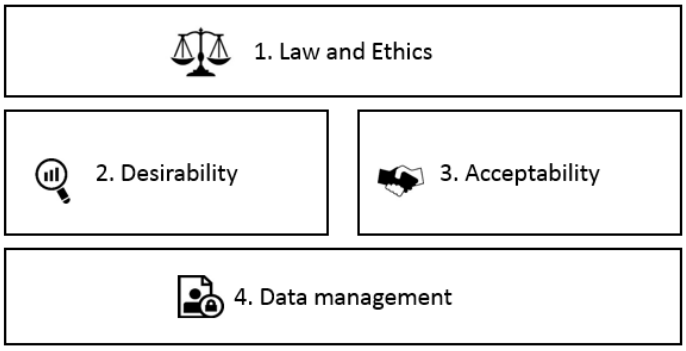

Figure 1: Eticas’ Societal Impact Assessment Framework

We organize them in levels in the sense there is a suggested order of analysis. However, the pillars can also be seen as "entry points" to the framework that can be deployed in any order or even individually. In practice, the impact assessment will be an *iterative process*. As the analyst

gains knowledge of the domain, the system and the societal issues involved, he or she is likely to go back and forward between the pillars, as well as identify issues that will affect more than one pillar. An overview of the framework is given below:

1. The SIA starts with an a priori legal and ethical assessment: what are the basic rights and values that an innovation should take into account in every situation? In this pillar, *Law* covers an analysis of national and international regulations as well as the so called 'soft laws' (e.g. recommendations, opinions and strategies). The later is important because many data-intensive systems apply to contexts and involve phenomena where there is no settled legislation.

Ethics relates to the values and moral standards guiding a project. These include both societal and individual values, as well as the notion of a social contract between the state and citizens, which may be reinforced or threatened by a technological innovation. As these stage, these are launched as hypothesis to guide the assessment in an attempt both to grasp the underlying values that guided a specific piece of existing legislation and to forecast all potential ethical issues that may be raised by the new data project.

It is worth noting that some of the key values at stake are included in the legal framework, as values are coded as law to ensure compliance of organizations and individuals with society's norms and values. Legislation, however, can lack the specificity required to be immediately transpositioned into state-of-the-art innovations, or not yet take into account new, emerging values and concerns, thus being, in itself, an insufficient foundation for data-intensive projects.

Once the legal and ethical foundation has been established, the framework suggests a more pragmatic analysis, which looks into the system's desirability and acceptability.

2. *Desirability*, the second pillar analyzed under this framework refers to the very need of a technological solution. An assessment of desirability is required because the *raison d'entre* of a new technology is often forgotten, assuming that the incorporation of technological advances in the society is invariably good. This is however not always the case: any technological solution must be proportionate to the problem it aims to address.

Desirability therefore can be achieved through a clear definition of the problem and the solution, a careful planning of its implementation (with indicators, maintenance needs, etc.) and a cost-benefit analysis of the system. The costs to be considered are not only economic, but also societal. While this pillar will not always quantify costs, it is a key decision-making support for designers.

3. *Acceptability*, the third pillar, accounts for the crucial issue of how citizens perceive, consent to and incorporate a given technology. The importance of acceptability has been exemplified in Section 2: often, perfectly legal systems have to be withdrawn because they have not been accepted by the society due to varied reasons, which may range from health to

distrust, privacy or cultural concerns. Acceptability therefore requires a public debate with an informed user base and broader public, as well as ensuring that choice, consent and control are accounted for.

It is worth noting that desirability and acceptability 'operationalize' law and ethics, in the sense that they translate the values examined under law and ethics to practical system's concepts, such as problem definition, alternative solutions, co-design, and security.

4. Finally, many of the concerns highlighted on the previous pillars can be implemented by a socially sustainable *data management* policy. In other words, data management is where often, in practice, both the problems and the solutions reside. It includes the legal framework of privacy and data protection, as well as broader considerations relating to individual control and consent, methods of anonymization, and how privacy issues can be designed into technologies and projects.

Data management, therefore, is where social, legal and ethical issues are identified, but also where the technical implementation of the other pillars takes place.

To avoid leaving engineers and decision-makers lost in abstract concepts, the framework further breaks down each pillar into a set of concerns (captured as keywords), techniques and questions. This gives the analyst guidance and allows he or she to start asking the right questions about the issues that need to be taken into account when designing and deploying data-intensive technologies or initiatives. The SIA is therefore not just a way to identify problems from a theoretical perspective. It actively involves a variety of stakeholders and provides action-guiding for the future design of a technology and the process to develop it.

3.1 Law and Ethics

The first pillar of the Eticas framework assesses the legal and ethical implications of a proposed innovation. These are discussed below:

3.1.1 Law

Compliance with regulation is not enough, but it an essential part of any responsible research and innovation. Below, we highlight some basic regulation that should apply to every research and innovation project.

3.1.1.1 Fundamental Rights

Fundamental rights are basic rights and freedoms to be respected regardless of nationality, sex, national or ethnic origin, race, religion, language, or other status. These rights are protected by documents such as the Universal Declaration of Human Rights (1948) and the Charter of

Fundamental Rights of the European Union (2000). Among the rights to be protected, are: the prohibition of discrimination and equality before the law; the right to respect for private and family life, and for the protection of personal data; the right to liberty, freedom of movement and security; the right to due process of law; and the right to freedom of expression, to peaceful assembly and freedom of association.

Any responsible research and innovation should make sure that these fundamental rights are respected, both by their final products and services, as during their development process. While this might be taken for granted, history has shown us that ignoring broad rights can have a profound impact. Take, for example, the principle of due process, which includes a set of rights such as the right to be presumed innocent, the right to be informed of the accusations, to defend oneself, to appeal, etc. The use of and reliance on personal data and surveillance measures in criminal proceedings have significant impacts on the principle of due process and on the rights of the accused in criminal proceedings (IRISS, Deliverable 1.1, p.77). Take, for example, the effect of some data-intensive technologies on the presumption of innocence. The UK DNA database registers over four million DNA profiles belonging not only to people convicted of crimes, but also to suspects not even charged yet. Once in the database, people whose DNA samples have been collected can become suspects in a criminal investigation. These individuals are not aware and have no control over how their data will be used, and if a match is found, the burden of proof would fall on them, not the prosecution, contravening European Convention principles. The presumption of innocence can also be affected by society's reliance on technology, which can make some people more easily persuaded and less critical in the face of technological 'evidence', which can lead to false positives.

3.1.1.2 Privacy and Data Protection

There is a considerable number of legal references regarding privacy and data protection. Privacy is a fundamental right, closely related to the concept of *personal data*, which is any information which can be used to identify an individual (e.g. names, telephone numbers, ID numbers, IP addresses, as well as medical, educational, financial, and employment information). Some personal data is *sensitive*, which means that it can be used against an individual in a discriminatory way (e.g. racial or ethnic origin, political opinions, religious beliefs).

While the concept of privacy is difficult to define, perhaps the most succinct and familiar definition is 'the right to be left alone' (Warren and Brandeis, 1890), which was later extended to 'the right to be left alone to live one's life with the minimum degree of interference' (Nordic Conference on the Right to Privacy, 1967).

Data protection refers to the maintenance of 'administrative, technical or physical measures to guard against unauthorised access to personal data' (David Lyon, 2007). While it is less of an

all-encompassing concept than ‘privacy’, data protection tends to guarantee a data subject's rights to control its personal data, which is fundamental to ensure privacy.

Currently, the main legal document regarding personal data protection in the EU is the Directive 1995/46/EC on the Protection of Individuals with regard to the Processing of Personal Data and on the Free Movement of such Data, which is the common reference for the national regulations of the EU Member States. The Data Protection Directive lays down enforceable principles of good practices regarding personal data, among which are fairness, finality, data minimization, data quality, conservation, confidentiality, security and notification to the supervisory authority.

However, the Directive 1995/46/EC will be replaced by the European Data Protection Regulation (GDPR), directly applicable to all EU Member States without a need for national transposition. The new regulation extends the 1995 directive in a number of ways, including its application to data controllers and processors of EU citizen data regardless of their location, forcing breaches notifications to be made in 72 hours of discovery, granting more power to data protection authorities, and only allowing the process of personal data with explicit consent (unless certain conditions exist). In particular, it obliges data controllers to conduct privacy impact assessments when privacy breach risks are high. This is an excellent opportunity to design methods and create better techniques to achieve this goal. At the time of writing, the new regulation has been approved and is expected to come into force in 2018.

3.1.2 Ethics

While adequately responding to societal challenges requires adherence to the law, legal compliance alone will not guarantee that society will endorse a technology. Therefore a wider range of ethical concerns have to be taken into account.

Ethics is fundamentally about values. In the words of Laudon and Starbuck (1996), ‘ethics are moral standards that help guide behavior, actions, and choices’. It refers to well-based standards of right and wrong that prescribe what humans ought to do. Furthermore, it is attributed to the study and development of personal and community ethics standards, in terms of behavior, feelings, as well as social habits and norms.

In the context of RRI, ethical concerns can be divided in two categories: the social contract between science, technology and society (discussed in Section 1) and the values that underlie this social contract.

However, what we have witnessed in recent years is the break-up of this social contract by a number of data-intensive systems. History collects continuous cases of privacy failures, security breaches and malfunctioning of the mechanisms for accountability and redress. Furthermore, there is an overall lack of transparency with respect to what technological systems are doing, as

well as a common absence of protection of citizens when faced with the unintended consequences of innovation, especially when personal data is involved.

These problems undermine the values that should uphold the relationship between science, technology and society. An awareness of these values will help guide decision-making in the innovation process, as discussed below.

3.2.1.1 Values

One of the most fundamental values is *liberty*. As captured by the law, people have the right to be free. That includes freedom of speech, of assembly, association and expression, of conscience, of personal property, and from arbitrary arrest, among others. Data-intensive technologies can potentially infringe upon different types of freedoms. Take, for example, the freedom of assembly, association and expression. This freedom lies on the assumption that citizens can engage in activities and discussions away from the power of the state and its gaze. However, the proliferation of data-capturing systems, especially when operated by public authorities, can affect people's sense of freedom: information about a person's activities and connections can be used against them. As another example, consider the freedom of movement. Data-intensive systems provide the means to identify, monitor, categorise and manage those who move, and create 'soft checkpoints' away from critical areas such as border crossings (Razac, 2009). These soft checkpoints can not only limit the freedom of movement in areas where it should prevail, but can also create vast databases and blacklists which could further impact human rights.

Liberty is also closely related to *autonomy*, the capacity to make informed decisions without external coercion. Autonomy is an essential part of self-worth and it is also related to the control one has of its personal data. When data-intensive systems collect data without consent, subjecting it to analyses whose purpose are unknown to the data subject, and keeping it for later uses, it is undermining people's control over their own data and their autonomy.

If freedom and autonomy are not respected, this could in turn have an effect on an individual's state of *self-worth* and *dignity*. The former refers to one's self perception in terms of self-respect. The latter, closely related, can be seen as the state of being worthy of esteem or respect (Dilon, 1995). Some types of data-intensive systems, such as social networks, can expose people's identity and images. This exposure can lead to all sorts of consequences, including stigmatization and the intrusion on private lives, potentially undermining a person's sense of self-worth. Furthermore, dignity and self-worth can also be affected when one feels their personal space has been entered.

Justice is another value related to the social contract that concerns fairness. In that sense, technology should ensure equality of treatment and avoid worsening inequalities by deepening the digital divide, for instance.

Equality of treatment states that similar situations must be treated equitably and prohibits discrimination and discriminatory treatment. In relation to technology, this means that systems should not systematically disadvantage certain groups of people. This is a danger especially when systems make automatic decisions, as they rely heavily on categorization and automated profiling practices, reinforcing mechanisms of social differentiation (Monahan, 2003). Automated sorting by categories of personal data can (re)produce marginalising effects and negative discrimination (Monahan, 2003). Possibly the most widely discussed cases relate to credit rating software, which can use a plethora of information gathered from several sources (e.g., social media, data brokers, and online trails) to automatically decide whether someone can access a loan. Decisions such as these can have a real impact upon the individual being profiled, affecting their confidence and ability to succeed in work and life. They can be even more damaging when causing a digital divide in the society. That is, when they result in the pre-emptive and disproportionate criminalization of whole sectors of society – such as migrants, students, the poor or the unemployed - damaging society's inclusive character (Woolgar, 2002).

Last but not least, a key value in the relation between science, technology and society is *democracy*. Democracy concerns collective decision-making. Technology developers can enhance democracy by taking the public opinion into account, rather than forcing their ideas and technologies upon society. Especially when public money is used to fund research and technology, the public deserve a say in what technologies are being developed.

3.2.1.2 Other Concerns

Besides values, ethics also encompasses some key concerns that can undermine the social contract between science, technology and society.

A basic principle to take into consideration is the *precautionary principle*, which provides a framework for decisions under uncertainty, aiming to anticipate and minimize potentially serious or irreversible risks. It calls for precautionary action: in the absence of full scientific certainty or consensus over the potential beneficial or harmful effects of any action or plan, the burden of proof must fall on those taking the action.

Value transfer is another important consideration. It happens when there is a mismatch with respect to the costs and benefits of a technological solution. For example, when costs borne by public entities are only enjoyed by private parties, or when investments by large constituencies can only be enjoyed by small sets of people.

Finally, an important and often overlooked issue is *function creep*. This phenomenon takes place when a given technology ends up being used for a completely different purpose. Take, for example, the European Dactyloscopie (EURODAC), a central repository that contains the biometric data of asylum seekers. While EURODAC was originally established to enhance the common asylum policy, later access was offered to police and other law enforcement agencies (Finn and Wright, 2011, p. 131).

Overall, ethics builds on the need for innovation to respect, reflect and adhere to a broader social consensus about which world we want to live in, how to achieve it, and what role technology should play. While these issues are crucially important, technology developers and engineers are often not acquainted with them, and may never be. Therefore, the need to operationalise these principles, values and concerns emerges as a crucial challenge of the information, data-driven society.

Below we present sample keywords, methods and questions that can help to assess legal and ethical issues.

<table>
<tr><td colspan="2">Operationalizing Law and Ethics</td></tr>
<tr><td>Example Keywords

Social contract, liberty, autonomy, self-worth, dignity, justice, equality of treatment, democracy, precautionary principle, value transfer, function creep.</td><td>Methods

Desk research, qualitative methods, press review.</td></tr>
<tr><td colspan="2">Sample Questions

- Are there earlier cases of failure or concern that can be taken into account?

- What are the legal regulations the technology should comply with?

- What procedures are in place to verify the on-going legality of the technology?

- What is the social contract the technology concerns? E.g: a surveillance technology might concern the relation of the police and citizens, where the police is supposed to protect citizens from harm.

- What is the potential for 'dual use'?

- What are the risks of ‘backdoor’ access in this technology?</td></tr>
</table>

- What steps have been taken to identify key societal values and how are these being translated into concrete requirements?
  - Can the technology coerce certain actions upon a user?
  - What are the short, medium and long-term impact/s of the technology?
- Initial assessment of the life-cycle of the data in the technology
  - What type of data analytics will be carried out by the technology?
  - Can this result in unequal treatment of the people interacting with the technology?

3.2 Desirability

Desirability considers the very need of a solution and whether it is proportionate to the problem at hand.

Proportionality is a balancing test between competing principles, which finds its rationale in the need to put costs and benefits in relation and avoid wrongdoing or unfairness (IRISS, 2012). While it is difficult to provide objective guidelines to assess proportionality, finding indicators to determine whether a project is proportionate is desirable and achievable, and should be an important part of any planning or evaluation exercise.

Therefore, the first step to assess desirability involves *defining the problem* to be addressed by the technology. The 'problem' is usually the discrepancy that exists between an existing situation and a desired situation. Stating clearly what this difference is contributes to the coherence of any new project being considered. While problem definition sounds like an obvious step, it is often ignored or taken for granted in many new initiatives.

Secondly, the context needs to be laid-out; that is, developing an understanding of the norms, perceptions and expectations in each setting. Norms, perceptions and expectations do not travel well, and so each 'recipe' to solve a problem will have to be tailored to every specific situation. An appreciation of this aspect also serves to avoid simplistic or deterministic approximations of the relationship between science, technology and society.

Once the problem and the context have been defined, attention is turned into the *proposed solution*. Solutions should have both clear and measurable *objectives*, against which the state-of-the-art can be compared either quantitatively or qualitatively, depending on the problem being tackled, and a specific and detailed *implementation* plan. Finally, *alternative* ways of solving the problem must be explored and defined. The do minimum alternative should always be taken

into account (what are the costs and benefits of not doing anything and how these compare to other alternatives).

In an ideal scenario, a complete assessment of *costs and benefits* should be carried out for all alternatives considered. Such analysis should consider not only the financial but also societal costs. In line with the iterative nature of this framework, it is important to note that while the cost-benefit analysis is performed under desirability, it will refer to values discussed under ethics and social costs uncovered under acceptability (e.g. in a public debate).

The sample keywords, methods and questions concerning desirability are listed below.

<table>
<tr><td colspan="2">Operationalizing Desirability</td></tr>
<tr><td>Example Keywords

Proportionality, problem definition, solution definition, project objectives, implementation, alternative solutions, costs and benefits.</td><td>Methods

Desk research, expert interviews, stakeholder interviews, focus groups, business canvas model, performance scoring, cost-benefit analysis (CBA), technical forecasting, social forecasting, training, stakeholders mapping, surveys, quasi-experiments, simulations.</td></tr>
<tr><td colspan="2">Sample Questions

- What is the problem the technology is trying to solve?

- How does the technology provide a solution to this problem?

- How can the improvement brought by the technology be measured?

- What is the range of possible solutions for the problem at hand?

- Is solving the problem with a technological solution the only option?

- Can solutions also be found in policy?

- Is it better to do nothing than to proceed with the identified solution?

- Is implementation included in the plan?

- Will first-line operators be properly trained?

- Have maintenance costs been taken into account?

- What are the financial and social costs of the technology on users and non-users?</td></tr>
</table>

3.3 Acceptability

The concept of acceptability encompasses all principles related to the promotion of people's awareness, information and control over the systems.

Here, the promotion and existence of a *public debate* is an essential element to assess whether the system will be acceptable from a societal perspective. In order for a public debate to exist, the relevant stakeholders need to be identified and have access to adequate *information* – who is promoting a specific new process, what information there is about its potential and risks, how is implementation monitored and what mechanisms are there for redress. Additionally, promoting public debate requires sensibility to the language and cultural variations of different stakeholders: information must be given in a way that they can understand, to ensure that no social group is discriminated against.

The concerns that are likely to come up during the public debate and that have a direct impact into acceptability include security, transparency and accountability. A public understanding of *security* requires that expected trade-offs are made explicit. What is to be secured, from whom, why, for how long, and to what extent are relevant questions that should find spaces where they can be asked, and system promoters should be able to answer them. *Transparency* relates to a clear statement of the system purpose, internal processing, costs and benefits, risks and their mitigation strategies. Finally, any new technology requires *accountability* from its developers and decision makers; that is, they must answer to and be liable by their innovation. This emphasis on security, transparency and accountability, in its turn, will contribute to the necessary building of *trust* between the stakeholders.

Promoting an honest public debate is also a precondition for *co-design,* which actively involves all stakeholders in the design process and helps ensuring that its outcome meets people's needs. Co-design will also make it easier to account for the *digital literacy* – different user groups are likely to have different abilities to locate, organize, understand and analyze information in a system – and yet all should feel comfortable interacting with the new technology.

Good planning and co-design are also a good basis for the creation of *responsible governance* mechanisms: accountability, enforcement, redress and oversight. Simplistically, responsible governance ensures that the new research or technological development respects the rights and interests of its stakeholders; that is, it 'says what it does and does what it says'.

Acceptability, however, goes beyond raising people's awareness and providing them with adequate information. It also requires giving people choice and control over the new technology, and making sure that they consent to it. Therefore, we see choice, control and consent as enablers of acceptability. In today's dynamic environment, *choice* evokes the need to

collectively take responsibility and direct technological developments in a direction that is desirable from a societal point of view. Choice is what provides citizens and users the possibility to influence and steer the course of things 'to the kind of world we are building' (Leenes, 2011). Choice is also intimately related to *control*. People's acceptability of a technology will increase if they feel they understand it and have some control over it – either in the decision-making process (being able to express their preferences or concerns) or at the point of delivery (being able to effectively turn on or off some functionalities).

Finally, *consent* concerns the process of getting permission from the user/subject to perform activities that may have an impact on his or her fundamental rights. In order to be valid, consent should be explicit, full and free, and lack of consent should not prevent the citizen from accessing goods or services – consent can and should be granular, and provide the subject with information and alternatives. Consent guarantees *transparency* and enables the subject to become aware of its own position, to adapt its behavior to a given situation, and even to refuse participation in specific processes or practices. Therefore, mechanisms to seek out the consent of those involved in an innovation process or product needs to exist and be clear and known by all those involved, willingly or unwillingly. Similarly, it must be clear that the fact of agreeing to something does not make the practice acceptable or desirable in itself, but it does create the procedures of redress.

Again, consistently with the iterative nature of this framework, choice, control and consent will be revisited also in the context of data management.

Some keyworks, methods and questions related to acceptability are shown in the box below.

Operationalizing Acceptability	
Example Keywords Public debate, co-design, digital literacy, responsible governance, security, transparency, accountability, choice, control and consent.	*Methods* Surveys, focus groups, stakeholder interviews, social forecasting, action research, simulations, participatory stakeholder mappings, quasi-experiments, simulations, co-design workshops.
Sample Questions - Who can be potentially affected by this technology, both directly (e.g. end users) and	

indirectly (e.g. society and non-users)?

- Is there a public debate on similar technologies?

- How is consent gathered?

- Has the public been consulted about plans for the technology?

  - Through what channels and to what extent?

  - Has the language been adapted to the different communities?

  - How has their input been integrated?

- What assessment has been made in terms of the end users digital literacy?

  - How did the results shape your planning?

- How is control ensured?

3.4 Data Management

Data management offers technical solutions to many of the concerns raised on the other pillars, and is closely related to the protection of privacy and personal data.

The way in which data is collected, managed, processed and stored in research and innovation processes is often ´black box´, meaning that citizens have no information about these procedures nor their rights in such processes. Therefore, non-compliance and lack of protection is the norm rather than the exception when it comes to privacy and data protection. Only when policies and technological solutions are put in place, the black box of data management can be opened up.

In this framework, we suggest that the best way to approach this step is by drawing the life-cycle of the data involved, mapping its different stages, the vulnerabilities to look out for and the principles and precautions to be enforced.

In order to map the data life-cycle, its main stages need to be defined and identified (see Fig. 2):

1. *Data collection*: the data is gathered under the condition that it will only be used for a legitimate purpose; that only the data needed for that purpose will be collected (data minimization); that the people to which the data concerns gave consent to it; that the data is of good quality; that there are mechanisms in place for the people to opt-in and opt-out; and finally that the data is properly anonymized when possible.

2. *Data storage*: when stored, data should be kept under certain conditions, such as it should be secure; who has access to the data has to be carefully considered; it should be established who is accountable for the data; and the process should be frequently audited.

3. *Data Analysis*: there are also principles for when the data is being analysed. Data should be anonymised when possible; the necessity of certain analysis techniques (e.g. profiling and sorting) should be considered, as well as the potential harm of the data produced by the analysis and possible mis-identification of an individual by an automated algorithm. Finally, accountability and auditing mechanisms should be put in place to make sure that all analyses performed are consistent with the system's purpose.

4. *Data Sharing*: there are different possibilities of data sharing, within the data controller, with data processors, with third parties and as open data. Whenever data is shared, the data life-cycle starts again, as it will be stored, analyzed, shared and eventually deleted. Each type of data sharing requires specific precautions to be taken, including anonymization and the adoption of security, accountability and auditing mechanisms.

It is also worth noting that data may be shared under emergency protocols. This can imply the suspension of certain rights and freedoms, depending on the specific threats. Nevertheless, this should not result in a 'legal vacuum' or non-compliance. Respect to non-derogable fundamental rights must be ensured, and the basic mechanisms of data management should be still in place.

5. *Data Deletion*: when data is no longer needed for the purpose of which it was lawful collected, it should be deleted (ideally, the deletion date should be defined at the moment of collection). Mechanisms must be put in place for requesting data deletion. Additionally, secure deletion mechanisms should ensure that deleted data cannot be retrieved and that there are not copies available (e.g. temporary storage and back-ups).

Finally, there are also a few precautions that concern the whole process: a data protection officer (DPO) should be designated to ensure the responsible management of data; personnel should be trained; and the entire information processing system must be placed under a legal footing, which includes the notification to relevant authorities.

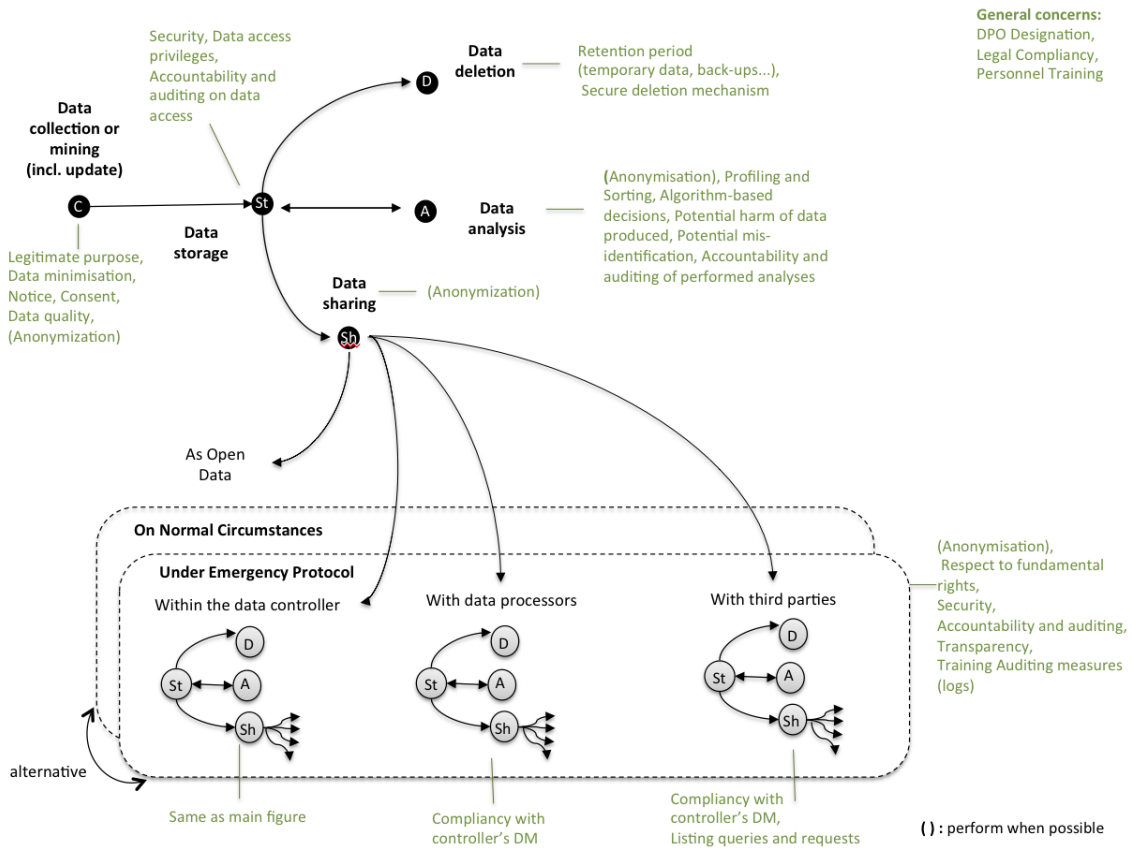

Figure 2: Eticas Data Management Life-cycle

The box below exemplifies some keywords, methods and questions regarding data management.

Operationalizing Data Management	
Example Keywords Purpose, data minimization, anonymization, notice, consent, data quality, security, accountability, auditing, transparency, mis-identification, harm of produced data, DPO designation, personnel training, retention period, secure deletion	*Methods* Privacy By Design (PbD), Privacy Enhancing technologies (anonymizers and pseudo anonymizers, encryption, filters and blockers, track erasers), authentication and authorization, secure printing, secure deletion, algorithm auditing.
Sample Questions - Does the data you collect include personal data?	

- What steps can be taken to minimize the amount of personal data required?

- For what purpose personal data is collected?

- Is this purpose legitimate?

- For how long this data is needed to achieve your purpose?

- Are you favouring privacy protection as part of your default settings?

- What measures are implemented to ensure that personal data is anonymized, when full information is not a necessity?

- Is data automatically and securely deleted when it is no longer needed?

- What mechanisms are in place to store and process data securely?

- Are your measures for the security of data stored adequate for the amount and nature of data?

- How is hacking or unauthorized access prevented, including from within?

- What measures are taken to prevent the accidental disclosure of personal details?

- Have you informed the relevant national data protection authorities (DPA) of your plans?

4. Conclusion

The externalities of data-intensive systems are often hard to predict, as it may take the widespread adoption of a technology and a long time for them to become apparent. Yet, a growing number of cases have emerged over the last decades, making society increasingly aware of this problem but also providing cues as to prevent and minimise negative externalities.

Responsible research and innovation (RRI) comes as a proposal from the scientific community to ease society's concerns about such externalities. RRI brings together societal issues to better align the project's process and outcomes with the values, needs and expectations of society.

This paper proposes a Societal Impact Assessment (SIA) framework for data-intensive systems that follows the general principles of RRI and stems from the authors' experience working on such technologies in different fields and domains. The framework combines an analysis of the societal impact with the practical technical and financial concerns of technology developer, using four perspectives: Law and Ethics, Desirability, Acceptability and Data Management.

Our framework, which we refer to as the Eticas framework, has been applied to data-intensive systems from a variety of domains, including security and surveillance systems. We have found this framework to be malleable; that is, it can be adapted to different systems and contexts, as

well as to the resources of the organizations performing the assessment. Even though an order of assessment is suggested, the analyst can choose any entry point that he or she feels comfortable with. The assessment of one pillar will inevitable raise issues regarding other pillars, easing the navigation to other types of concerns.

The framework also accounts for the fact that often technology developers have a different understanding of the concepts being discussed in a societal impact assessment (e.g. dignity, autonomy, trust). It therefore seeks to translate these abstracts concepts into simple keywords with clear definitions and concrete questions that can be asked to or by technology developers. Analysts can also pick and chose among the suggested tools and techniques, accordingly to their skills, time and budget. Finally, the framework provides practical data management recommendations to tackle many of the issues identified, and contribute to furthering privacy-enhancing solutions and methods.

However, adopters of the framework may also face a few challenges. The framework's iterative nature means that analysts may find themselves navigating from one pillar to another. Even though that might seem frustrating at times, this format allows for a gently, yet deep dive into the potential externalities of any system. Another challenge worth highlighting concerns the globalized nature of many data-intensive systems. For software-based systems that know no geographical boundaries, the legal assessment can be quite challenging, due to differences in national regulations and language issues. Yet, it is a step that should not be overlooked. Finally, the framework is not a sure recipe for avoiding negative externalities. Its success depends on a genuine commitment from all stakeholders. This is particularly the case with technology designers, which should adopt a mind-shift from technology inventors to solution providers; that is, people that develop technologies that are proportionate to real problems, while considering the values, needs and expectations of the communities beyond their user base.

In terms of future work, even though this framework has been used in a variety of projects, we see it as a 'live framework'; that is, as an ever-evolving tool to accommodate technological advances and social changes. As such, we are continuously exploring existing tools and techniques and creating new ones to be used within the framework. In particular, we continue to work on improving ways to find out and assess peoples' true concerns and behaviours when confronted with the choice between short-term benefits and potential negative externalities in the far future, as well as their attitudes towards individual versus societal gains. We are currently designing experiments to gain better insight into people's current and future values and expectations when it comes to data-intensive technologies and the changes they bring about.

5. References

Alabdulkarim, L. O., Lukszo, Z., & Fens, T. W. (2012, June). Acceptance of privacy-sensitive technologies: smart metering case in The Netherlands. In*CESUN 2012: 3rd International Engineering Systems Symposium, Delft University of Technology, The Netherlands, 18-20 June 2012*.

Assembly, U. G. (1948). Universal declaration of human rights. *UN General Assembly*.

Banta, D. (2009). What is technology assessment?. *International Journal of Technology Assessment inHealth Care*, *25*(S1),7-9.

Carson, R. (1962). *The Silent Spring*. Boston: Houghton Mifflin.

Clarke, R. (2009). Privacy impact assessment: Its origins and development.*Computer law & security review*, *25*(2), 123-135.

Collingridge, D. (1982). *The social control of technology*. London: Pinter

Dillon, R. S. (1995). Dignity, character, and self-respect. Psychology Press.

European Convention (2000) Charter of Fundamental Rights of the European Union, OJ C364/01. Available at: http://www.europarl.europa.eu/charter/pdf/text_en.pdf

European Union (2012). *Responsible Research and Innovation: Europe's ability to respond to societal challenges*. Pg. 3. Available at https://ec.europa.eu/research/swafs/pdf/pub_public_engagement/responsible-research-and-innovation-leaflet_en.pdf.

Finn, R. and Wright, D. (2011) *Deliverable 2: Privacy, data protection and ethical issues in new and emerging technologies: Five case studies.* PRESCIENT EU-FP7 Project.

Gibbons, M. (1999). Science's new social contract with society. *Nature*, *402*, C81-C84.

Grunwald, A. (2011). Responsible innovation: bringing together technology assessment, applied ethics, and STS research. *Enterprise and Work Innovation Studies*, *31*, 10-9.

Guston, D. H., & Sarewitz, D. (2002). Real-time technology assessment.*Technology in society*, *24*(1), 93-109.

IRISS (2012) Deliverable 1.1. Surveillance, fighting crime and violence. Pg. 77. Available at www.irissproject.eu.

Leenes, R. (2011). Framing techno-regulation: An exploration of state and non-state regulation by technology. *Legisprudence*, *5*(2), 143-169.

Lyon, D. (2007). *Surveillance studies: An overview*. Polity.

Monahan, T. (2002). Surveillance and inequality. *Surveillance & Society*, *5*(3).

Nordic Conference On The Right To Privacy, (1967). Bulletin of the International Commission of Jurists.

Pinch, T. J., & Bijker, W. E. (1987). The social construction of facts and artifacts: Or how the sociology of. *The social construction of technological systems: New directions in the sociology and history of technology*, *17*.

Razac, O. (2009). Le bracelet électronique et la virtualisation de l'enfermement.*Cultures et Sociétés*, (10), 52-57.

Schot, J., & Rip, A. (1997). The past and future of constructive technology assessment. *Technological forecasting and social change*, *54*(2), 251-268.

Woolgar, S. (2002). *Virtual society?: Technology, cyberbole, reality*. Oxford University Press on Demand.

Warren, S. D., & Brandeis, L. D. (1890). The right to privacy. *Harvard law review*, 193-220.