SPARTA

D2.5

Concrete ELSA solutions for the SPARTA programs: A process for assessing and improving the Societal Readiness Level of Cybersecurity Research

Project number	830892
Project acronym	SPARTA
Project title	Strategic programs for advanced research and technology in Europe
Start date of the project	1st February, 2019
Duration	41 months
Programme	H2020-SU-ICT-2018-2020

Deliverable type	Report
Deliverable reference number	SU-ICT-03-830892 / D2.5 / V1.0
Work package contributing to the deliverable	WP2
Due date	June 2022 – M41
Actual submission date	21st September, 2022

Responsible organisation	Fraunhofer
Editor	Michael Friedewald
Dissemination level	PU
Revision	V1.0

Abstract	This deliverable outlines the concept of Responsible Research and Innovation, proposes an operationalisation to increase societal readiness of cybersecurity research and reflects the conditions for mainstreaming RRI in cybersecurity.
Keywords	Responsible Research and Innovation, Stage gate process, Societal Readiness Level, Reflection tool, ELSA governance

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 830892.

Editor

Michael Friedewald (Fraunhofer)

Contributors (ordered according to beneficiary numbers)

Michael Friedewald, Greta Runge, Ralf Lindner (Fraunhofer)

Reviewers (ordered according to beneficiary numbers)

Aleksandra Pawlicka (ITTI)

George Kokkinis, Giannos Kotrogiannos (KEMEA)

Disclaimer

Executive Summary

In this deliverable, we suggest a process to consider ethical, legal and societal aspects in cybersecurity research during the development process. The aim is to identify possible problematic aspects as early as possible and to find a solution. This is done in an iterative and participatory process in which the interests of the involved stakeholders are weighed against each other and an adequate (technical or organizational) solution is decided upon. In this way, cybersecurity solutions should be created that are acceptable to all parties involved and thus have a high "societal readiness level".

In the Chapter 2, we introduce the theoretical background of this process, the concept of “Responsible Research and Innovation” (RRI) that was adapted by the European Commission a few years ago and implemented in the research framework programs. We discuss procedural and content related elements of RRI and how they can contribute to the goal of “better technology in a better society”.

Chapter 3 is showing that it is not enough to create a stand-alone reflection and evaluation process involving only the researchers. In order to actually develop sustainable solutions that have a high level of social readiness, it is necessary to spread the ideas of RRI to other areas of a company or an authority. Only through “deep institutionalization” can RRI friction losses and conflicting goals be avoided and additional benefits realized.

In chapter 4, the elements presented before are brought together in a stage-based reflection process, which was originally generic in nature but has been adapted here to the specifics of cybersecurity. This process foresees interventions at four points during the development process, before research and development work begins, after the system architecture is complete, at the end of implementation, and finally before commissioning. At each of these times, all relevant involved and affected parties are consulted on potential (negative) impacts and possible countermeasures are decided (depending on the outcome it may also be necessary to repeat one of the earlier development phases). To support the consultation, a list of criteria and questions covering relevant procedural and content-related aspects is provided to avoid blind spots.

Finally, Chapter 5 discusses some practical experiences in dealing with ethical, legal, and societal issues from the SPARTA project and draws conclusions about what is needed to successfully address such issues in future cybersecurity activities and structures.

Table of Content

List of Figures

List of Tables

Chapter 1 Introduction

Digital products and services are now ubiquitous in all areas of business and society. As a result, cybersecurity has become one of the most important techno-social issues of our time. All stakeholders are required to do their part to ensure cybersecurity. Strategies have been developed and institutions established to protect assets important to the state, the economy and society, and to defend against possible attacks on them. However, the various actors involved and affected have different interests, (technical) capabilities and power to enforce them (see Figure 1).

Simplified, there are decision-makers in politics and industry who determine which values and whose interests are particularly important and which means are appropriate to protect them. Solution designers (researchers and engineers) decide how these prioritized goals can be achieved in a concrete implementation. In our context, such solutions are primarily information technology solutions, but they are embedded in a complex environment that also includes regulatory, organizational, etc. designs. Finally, there is the group of those who operate cybersecurity solutions on a regular basis (security authorities, companies, private individuals) and those who are merely affected in some way by the use of cybersecurity (e.g., by the activities of the security authorities).

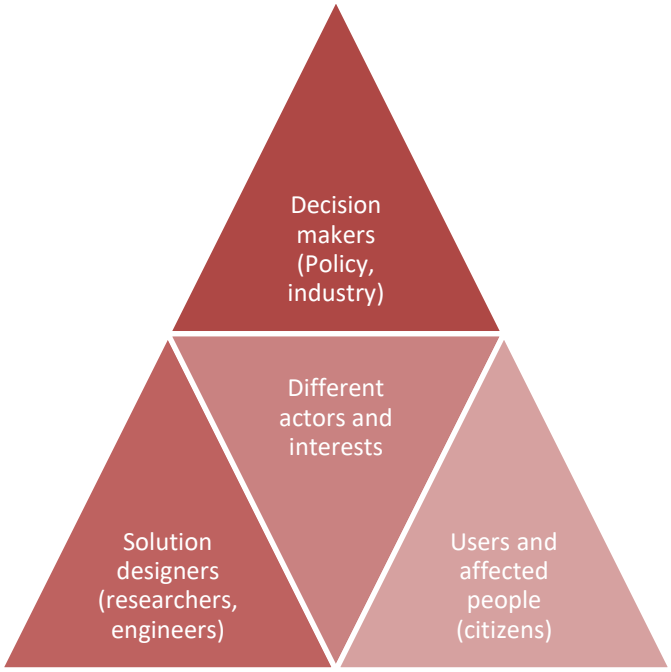

Figure 1: Actors in the cybersecurity ecosystem

So, although cybersecurity is a task for society as a whole, the influence of the different groups is very different. As a result, the impacts of cybersecurity measures often do not affect those who make decisions on goals and system design. And even among those affected, there are those who, due to their (technical) knowledge, financial possibilities, etc., are able to protect themselves against undesirable effects better than others.

It has been a declared political goal for many years that technical systems should not only be functional, i.e., achieve a predefined goal. They should also comply with legal requirements (e.g., data or consumer protection) and not violate the rights of users. Cybersecurity should be designed in a human-centric way [22], “as if ordinary citizens mattered” [63]. In general, the goal should be to achieve "better technology in (a better) society" as Schot and Rip [60, p. 263] phrased it. To this end,

it is necessary to work closely with researchers and developers in the development process, to identify societal issues and potential controversies at an early stage and to find solutions to them together as early as possible. In practice, this task is not easy to implement. Technical developments are subject to their own logic, mostly an economic profit motive. Without external pressure (such as legal requirements), internal organizational support and a certain (intrinsic) motivation of those involved [66], consideration of ethical, legal and societal issues is usually not realistic. Researchers have long been developing, testing and modifying concepts and procedures on how to proceed in order to implement "better technology in (a better) society" with the involvement of all stakeholders in a way that all interests are fairly taken into account.

The SPARTA "responsibility activities" precisely pursue these goals, but with the awareness that cybersecurity is a very complex subject and (in terms of technical procedures) incomprehensible for most laypersons. Moreover, (cyber) security technology is considered by many - especially governmental – actors to be in need of secrecy. The challenge was thus to define a process that involves stakeholders actively, does justice to the complexity of cybersecurity and leads to technical solutions that do not raise ethical, legal and societal problems and can thus be considered "societally ready". In essence, this attempt is based on the approach of "responsible research and innovation" (RRI), which has been propagated and promoted by the European Commission in its Framework Programmes for Research and Innovation for several years [59; 78]. The background and approach of RRI are explained in more detail in the following Chapter 2.

Researchers are increasingly concerned with aligning their work with societal needs and preventing it from having predictable negative impacts, particularly in fast-moving and ethically sensitive areas such as cybersecurity. In a recent paper, Bernstein et al. [6] describe a tool to help researchers and research funders incorporate responsible research and innovation values into their work to promote society-centered research. The Societal Readiness Thinking Tool (SRT) is designed to help funders and researchers incorporate considerations of aspects of open and responsible R&I practices from the outset of the development process and include them in proposals. Against this background, in Chapter 3, this deliverable will address what principles are necessary for the responsible design and development of effective society-centric cybersecurity solutions (software) and what contribution the SRT could make to integrate social and ethical dimensions of responsibility into the actions of actors in the cybersecurity software development process.

SPARTA was fortunate to be able to build upon such preparatory work from other EU funded projects that have gradually developed a set of RRI heuristics, procedures and tools in recent years. Many of these instruments are very general. Such generic tools and guidelines are good if you want to deal with a variety of technical developments in a unified way, but most of the recommendations for decision makers derived from them are so general that it remains unclear for many organizations how to implement them [see e.g. 30]. We decided to adapt the assessment (or reflection) tool developed in the NewHoRRizon project[1], which has been tested in all areas of EU research funding, including "civil security for society". Chapter 4 explains how this tool was adapted to the specifics of cybersecurity, including a modified set of assessment criteria that emerged from our own preliminary work in the SPARTA project [42; 76], validated in exchange with the other SPARTA partners.

Chapter 5 finally presents and discusses practical experiences from the governance of ELS in the SPARTA programs and derives recommendations for future cybersecurity research and its organization. Some of the emerging ELSA issues were eventually taken up by partners involved in the “responsibility activity” and addressed through independent research [documented in 75].

[1] https://newhorrizon.eu/

Chapter 2 Background: Responsible research and societal readiness in cybersecurity

2.1 Theoretical background and origin

Responsible Research and Innovation (RRI) is an eclectic approach, building upon several earlier concepts, partially integrating and developing them further. It makes use of concepts and disciplinary contributions from science and technology studies (STS), ethics of science and technology, ELSA/ELSI research (ethical, legal, social aspects/implications), sustainable technology development, value sensitive design, responsible development, participatory and transdisciplinary research, research integrity, responsible metrics, etc. [11]. RRI became a cross-cutting theme in the EU's research framework program, but it was also implemented as a mandatory element of technology-oriented projects in certain work programs.

The concept of RRI is seen as a measure for addressing major societal challenges in a range of areas, from transport and climate to social participation and cybersecurity [2; 14]. Despite the broad interest in RRI across different societal sectors such as politics, business, and academia, there is, on the one hand, academic consensus on what the concept means in terms of what can be understood by responsible research and development. Von Schomberg [59, 72] understands RRI to be

> *"…a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view on the (ethical) acceptability, sustainability and societal desirability of the innovation process and its marketable products (in order to allow a proper embedding of scientific and technological advances in our society)".*

The European Commission also conceptualizes RRI as a process in which "all societal stakeholders (researchers, citizens, policy makers, businesses, civil society organizations, etc.) work together throughout the R&D process to better align R&D outputs with the values, needs, and expectations of European society" [48]. From a conceptual point of view, the RRI concept combines the following six key elements: public engagement, gender equality, science education, open access, ethics, and governance [26]. From a operational perspective, RRI can be understood rather as a process to integrate the key elements into the R&D process rather than as an outcome (see also [13]). The key elements are to be understood as levers that ensure that the R&D results are in line with the values, needs and expectations of European society.

Stilgoe et al. [70] believe that the central challenge of responsible innovation (RI) is to address societal challenges - the main challenge is that these challenges are unpredictable. In light of this unpredictability, the researchers conceptualized a framework for prospective RRI, consisting of four dimensions: *Anticipation, Inclusion, Reflexivity, and Responsiveness*. According to Stilgoe et al., the dimensions originate from a set of questions that have proven important in public debates about emerging areas of science and technology [70, 1570]. It often remains unclear which tools and instruments can be used to achieve an outcome in the relational dimensions (operational aspects). In other words: How the shift from concept and theory to practice can succeed (cf. [68]). Against this background, the question of suitable tools for RRI must be addressed with regard to a concrete problem within a social field [7; 13]. This should not only consider the specific integration of societal concerns via the six key elements of RRI, but also the relational dimensions along the entire development process. It is elementary to include the involvement of various actors from research, business, politics, education, and civil society at the system and project level in research and innovation. The development and introduction of innovations and technologies and the associated market penetration is a social process with various social impacts. In the context of Cybersecurity, RRI creates an "institutional ambiguity" [9]. The scientists argue that every security research project/instrument raises ethical, societal, and legal questions, but at the same time RRI helps determine the goals, scope, and outcomes of security research and policy [9]. According to this, RRI would consist of a "dialectical process" as "values, principles and legal provisions (are attempting to) define the range of research and innovation" [9].

Public acceptance of technology and innovation is essential for overcoming complex challenges such as the protection of IT-Infrastructures and cyber security applications, which is the subject of this deliverable. In the context of RRI, the focus on acceptability is crucial for the development of society-centered technology, processes and products. In this context, the question also arises as to what extent the acceptance of innovation and technology in society depends on cultural factors and thus differs from region to region. Factors influencing technology development are dealt with in the scientific discussion under the "readiness" concepts. "Technological Readiness (TR)" was originally developed as a cross-organizational method to advance technology projects in the space and weapons industries to avoid reverse salient or bottlenecks [34]. In this context, TRL provided a framework for assessing the state of development of new technologies based on a systematic analysis of aspects such as technical and economic costs or potential risks associated with a technology [32]. The focus of TRL is not on ethical considerations regarding the social acceptability of technology [37]. To fill this lack, Bernstein et al. [6] argue that RRI addresses existing gaps in TRL and thus can also be opened in the direction of "Societal Readiness (SR)". The concept can be broadly understood as society's readiness to accept and use (new) technologies [5, 8].

To address this socially oriented readiness in the R&D process, a *Social Readiness Thinking Tool* was developed by the "NewHoRRizon" project (2017-2022). The Project aimed “at further integrating Responsible Research and Innovation (RRI) in the research and innovation systems on national and international levels which broadly addresses the integration of RRI into all phases of research and innovation project development and implementation". The *Societal Readiness Thinking Tool* offers practical guidance on how to mature the societal readiness of research projects. In the form of a social lab, Bernstein et al. developed and tested the tool with various stakeholders who wanted to pay more attention to promoting social dimensions in the R&D process [6]. The tool complements existing Technology Assessment (TA) approaches to the effort to integrating RRI, sustainability, and design thinking into the R&D lifecycle. Design thinking is an approach "that uses the designer’s sensibility and methods to match people’s needs with what is technologically feasible and what a viable business strategy can convert into customer value and market opportunity" [10, p. 2]. By adopting a staged, process-oriented approach [18], the Societal Readiness Level (SRL) Thinking Tool provides practical guidance to researchers on how to reflect on and expand the level of accountability in their research and innovation projects (see Chapter 4). Given the high systematic reflexivity of social concerns, the tool can also provide valuable indications in the field of cybersecurity applications as to whether software in the cybersecurity sector is being developed under socially ethical and responsibility-centered criteria.

2.2 Procedural dimensions for reflection and benchmarking

SPARTA has demonstrated that digital transformation requires solid security and that security is the enabler for successful digital transformation. In doing so, the project always focused on concrete, application-oriented challenges and offered different actors the opportunity to collaborate in order to foster the development of transformative capabilities and the formation of world-leading centers of excellence. Against this backdrop, we took on the question of *how the procedural RRI dimensions of anticipation, reflexivity, inclusiveness, and responsiveness can be meaningfully translated into research and innovation practices related to cybersecurity*. This section first presents the dimensions at the theoretical level in order to be able to draw references back to the application area of cybersecurity.

2.2.1 Anticipation

On a conceptual level, RRI reflects anticipatory governance [40]. Understanding governance as a basic framework for activities, anticipation can be understood as an approach to mitigate intended and unintended consequences and applications of R&D in different societal contexts [54, 754]. According to Stilgoe et al., anticipation involves "systematic thinking aimed at increasing resilience, while revealing new opportunities for innovation and the shaping of agendas for socially-robust risk

research" [70, p. 1570]. At the operational level, anticipation combines several techniques and approaches. These include: Foresight, Technology Assessment, Horizon Scanning, Scenario Development, Vision Assessment, and socio-literacy techniques. Furthermore, factors can be cited that influence the implementation of anticipation in RRI processes. One factor is the confrontation with existing and differing worlds of imagination. In the context of technology assessment, it is about participation instead of prediction, in horizon scanning primarily about plausibility. Particularly, the acting actors in the R&D process are part of the dimensional view of Anticipation. Anticipation concerns the ability of the actors involved to reflect on the potential positive and or negative impacts of research and innovations and to take them into account in the development process [39, 2331]. In this context, multidisciplinary collaboration, as well as the integration of social scientists and ethicists in research settings, is considered an approach to Anticipation in the RRI context (embedded social scientists, sometimes referred to in the literature as “embedded humanist” [62] or “engagement agents” [73]).

Transferring the theoretical approaches of the RRI dimension “anticipation” to R&D activities in cybersecurity, three main features can be identified. First, the *Implications* that must be anticipated from the threat, which could range from impacts on citizen privacy to the failure of critical infrastructures [46]. This is also the basis for the *Strategic Character* that must be considered in responsible development and design. Important methods here are foresight processes or the early involvement of various (societal) stakeholders in *Multidisciplinary Cooperation* settings [3].

2.2.2 Reflexivity

According to Stilgoe et al., *reflexivity* challenges scientists: “In public, to blur the boundary between their role responsibilities and wider, moral responsibilities. It therefore demands openness and leadership within cultures of science and innovation” [70, 1571]. Stovall names “professional self-awareness” as a synonym for reflexivity and further as “a sort of master virtue that fosters the reflective deliberation necessary for a professional to pursue their work in an aspirational frame of mind" [69, p. 256; 71, p. 110]. In addition, reflexivity is also described as the ability of actors to respond towards the needs from the public, in the R&D process and to deal with the complexity of broad participation [39]. At the application level, promoting the ability to be reflexive regarding one's own ideas on dominating ways of thinking about the relationship between science and innovation is central according to RRI condition [72]. In this light, Stilgoe et al. [70] define several techniques and approaches as well as implementation factors for the RRI process. Deliberative techniques and approaches including various approaches focusing on ethical aspects, codes of conduct but also citizen-appointed panels and deliberative techniques that could be used, for instance, in dialogue formats or focus groups [64].

Regarding the ability of innovation systems to be *reflexive*, Lindner et al. define that "reflexivity or reflective capacities should be major criteria of the quality of innovation systems. The ability of innovation systems to collectively reflect about orientation and objectives, and to bring together actors to deliberate goals, to prioritize and to specify" [45, p. 6]. R&D processes related to cybersecurity are also embedded in a systemic level, so that "culture, infrastructure, standards, norms or policies define the 'rules of the game'" [45, p. 18]. Which, in the context of the reflexivity dimension, also underscores the *complex nature* of RRI in R&D processes. Zedler [77] sees a need to catch up on this strategic level, especially in the area of cybersecurity, which entails an organizational and technical level. On the one hand, these fields of action can be concretized in terms of political and legal framework conditions (societal level). On the other hand, reflective capabilities must be developed in terms of communication, coordination, expertise, specialist personnel and technical equipment (organizational level) [77].

2.2.3 Inclusion

Inclusion is linked to the need to involve and engage societal stakeholders in RRI decision-making processes [39, p. 2331]. According to Stilgoe et al. [70], RRI processes that aim at inclusiveness are forcing to deal with issues of power. According to Goodin and Dryzek [28], Inclusion is based on deliberative processes; these are embedded in sociopolitical systems and are necessary to achieve legitimacy [24, p. 202; 28, p. 232]. In this context, Brand & Blok [8] note potential tensions between

transparency and competitive efforts and suggest that the aspect of Inclusion should be reflected at the level of the outcome rather than in terms of the RRI process [8, p. 19]. At the application level, van Mierlo et al. define that "participatory, tailor-made techniques for public dialogues are used to include the public, NGOs and other stakeholders that are usually absent from science, development and innovation, with the aim to open up the innovation process" [49, p. 361]. Also on the level of Inclusion Stilgoe et al. [70] define various techniques and approaches to integrate RRI into application-related contexts. In this regard, deliberative tools such as deliberative mapping / polling stand for the ability to mitigate power imbalances or to question social and ethical aspects associated with new knowledge and technology. In addition, aspects of *user-centered design*, as well as the role of open innovation in integrating citizens into the R&D process also plays an important role in this dimension. In particular, user-centric design and security is considered as a major challenge in the field of information- and cybersecurity systems. Following Poel et al., inclusion addresses the dialogue that a company has with stakeholders and end users, so at the operational level, companies rely on user-centered design techniques and user groups to learn the exact needs and concerns of users and stakeholders" [21, p. 12]. At the institutional level, Stilgoe et al. mention regulations, standards and mechanisms for transparency (open access) as tools for RRI integration. Aspects that could have an effect on institutional implementation include the institutional structure or dominant political discourses [70, p. 1573].

The thematic complexity of cybersecurity requires the active *involvement of a range of stakeholders* [51]. Important Stakeholders who should be engaged include the public and private service providers who can provide input about critical infrastructure and software development. The complexity of the topic and the diversity of stakeholders raises the question of how *participation* should be structured and which formats (at the system and organizational level) open up the discussion space for strategic interactions between the various stakeholders [31].

2.2.4 Responsiveness

The fourth dimension of the RRI framework, *Responsiveness,* covers the capability of stakeholders in the R&D process to respond to needs and changes [39, p. 2331]. *Responsiveness* is conceived as a condition for responsible R&D processes: “responding to new knowledge as this emerges and to emerging perspectives, views and norms. For responsible innovation to be responsive, it must be situated in a political economy of science governance that considers both products and purposes” [70, p. 1572]. In this light, Responsiveness is closely intertwined with the other RRI dimensions and can reinforce their consideration. At the governance level, Kooiman and Jentoft argue that the responsiveness of an innovation depends on “Metagovernance” and the values, norms, and principles that that support innovation action – on the process and system level [44, p. 819]. The focus here is on the institutional conditions that are decisive for the design of future-oriented activities on the one hand and for the evaluation of governance activities that have already taken place on the other [44, p. 822]. Approaches and enablers for RRI in application in this dimension include Niche Management, Value Sensitive Design, or Stage-gates. Stilgoe et al. [70] also define institutional cultures, management, openness and transparency, and technological standards as factors that can have an influence on the implementation of responsiveness.

An important aspect of responsiveness is the *adaptability* of the RRI process in response to the *outcome of the other dimensions*. All activities in the context of cybersecurity should be based on principles (see Chapter 4.1) and integrated into governance (system and organization level). Moving this to the application level in cybersecurity, “adaptive security” can be an approach that analyzes behaviors and events to protect against threats. In this way, potential risks from applications can be assessed, continuously evaluated and proportionate measures can be taken automatically [33; 43].

Dimension	Indicative techniques and approaches	Factors affecting implementation
Anticipation	Foresight	Engaging with existing imaginaries
	Technology assessment	Participation rather than prediction
	Horizon scanning	Plausibility
	Scenarios	Investment in scenario-building
	Vision assessment	Scientific autonomy and reluctance to anticipate
	Socio-literary techniques	
Reflexivity	Multidisciplinary collaboration and training	Rethinking moral division of labour
	Embedded social scientists and ethicists in laboratories	Enlarging or redefining role responsibilities
	Ethical technology assessment	Reflexive capacity among scientists and within institutions
	Codes of conduct	Connections made between research practice and governance
	Moratoriums	
Inclusion	Consensus conferences	Questionable legitimacy of deliberative exercises
	Citizens' juries and panels	Need for clarity about, purposes of and motivation for dialogue
	Focus groups	Deliberation on framing assumptions
	Science shops	Ability to consider power imbalances
	Deliberative mapping	Ability to interrogate the social and ethical stakes associated with new science and technology
	Deliberative polling	Quality of dialogue as a learning exercise
	Lay membership of expert bodies	
	User-centred design	
	Open innovation	
Responsiveness	Constitution of grand challenges and thematic research programmes	Strategic policies and technology 'roadmaps'
	Regulation	Science-policy culture
	Standards	Institutional structure
	Open access and other mechanisms of transparency	Prevailing policy discourses
	Niche management[a]	Institutional cultures
	Value-sensitive design	Institutional leadership
	Moratoriums	Openness and transparency
	Stage-gates[b]	Intellectual property regimes
	Alternative intellectual property regimes	Technological standards

[a] Schot and Geels (2008).
[b] See below and Macnaghten and Owen (2011) for an example of this.

Table 1: Four dimensions of responsible innovation

Source: Stilgoe, Owen [70, p. 1573], p. 1573.

2.3 Content dimensions for reflection and benchmarking

As already mentioned, the RRI approach or the early consideration of “ethical, legal and societal aspects” (ELSA) aims to identify and adequately take into account the interests of all stakeholders involved in the development and use of cybersecurity solutions as well as the sensitivities of those affected. This is to avoid conflicts and possible non-acceptance. But which principles or values are relevant in this context and how can acceptable compromises be found for all sides? In an earlier deliverable [42], we have explained which values are at stake, which fundamental trade-offs exist in pursuing their guarantee and how a balance can be achieved. In the following paragraphs, we will briefly summarize this.

Following Stahl, we assume that normative principles and legal rules are the basis of RRI, as they define whether a certain type of research and innovation is desirable or acceptable [67, p. 711]. Moreover, normative values can provide concrete guidance for both research design and the assessment of the impact of security policy [9].

But which values need to be considered when developing cybersecurity solutions? According to Christen et al. [16], there are generally a large number of potential values[2], but only a limited number of them are relevant for cybersecurity. In particular, these are the "European values" - as enshrined in the Charter of Fundamental Rights of the European Union. It is demanded -- at least from the policy makers -- that future technology should be oriented towards these values in order to counter the libertarian US-American model and the authoritarian Chinese model [25].

[2] These are, inter alia, accountability, democracy, dignity, equality, fairness, freedom, justice, non-discrimination, pluralism, privacy, respect for human rights, rule of law, security, solidarity, tolerance.

There are suggestions [20; 23] that in the field of cybersecurity, the still large number of "European values" should be grouped into a small number of clusters that stand for the areas where potentially the greatest societal impact can be observed. In D2.1, we followed van de Poel and suggested four clusters, namely security, privacy, fairness and accountability. In the context of operationalising an RRI-based assessment and design process, which is described in more detail in the next chapter, it has been shown that a slightly modified set of clusters better covers the specifics of cybersecurity. These are trustworthiness, safety and human-centricity.

In [42], we have argued for a virtue ethics approach to the question of how to resolve ethical problems or value conflicts, as other approaches to decision-making (such as utilitarianism and deontology) are not flexible enough for practical problems occurring in cybersecurity [20]. For this reason, we have opted for a deliberative approach in which all groups involved in or affected by a specific technical development are involved in the R&D process[3]. In this approach, possible conflicts are identified in several iterations, possible solutions are discussed, and then design decisions are made that are acceptable to all as far as possible (see Chapter 3).

2.3.1 Trustworthiness

Trust is an important prerequisite for a technical solution to be accepted and used and to avoid possible conflicts. This cluster also includes aspects of transparency, explainability, and, especially, accountability. It also includes the question of how we are dealing with uncertainty and with “dangerous” (i.e., not trustworthy) actors. According to Hunt [35], trustworthiness can only arise when there is a "readiness or preparedness to give an explanation or justification to relevant others (stakeholders) for one’s judgments, intentions, acts and omissions when appropriately called upon to do so." This also includes the "readiness to have one’s actions judged by others and, where appropriate, accept responsibility for errors, misjudgements and negligence and recognition for competence, conscientiousness, excellence and wisdom." This corresponds with all the principles propagated by RRI. In the context of cybersecurity, this means for instance the ability of the controller to demonstrate that their activities and measures are compliant with the law: In the GDPR, the basic assumption is that a controller can be trusted by default, but that they must always be able to prove that they deserve this trust [42, sect. 2.1].

2.3.2 Safety and reliability

One possible point of conflict is the question of whether the cybersecurity solutions can cause disadvantages or damage to the affected persons. Therefore, reliability is a new value cluster that we did not foresee before. It is gaining importance as cybersecurity solutions in many cases have the character of an infrastructure. Therefore, availability, correct and predictable functioning, robustness/resilience and controllability are prerequisite for smooth operation and acceptance by stakeholders. Related to the solutions being an infrastructure or even public service, there are also links to values such as transparency and accountability [42, sect. 2.1].

2.3.3 Human-Centricity

As mentioned in the introduction, the landscape of cybersecurity development is characterized by a power imbalance between those who actively make decisions about the goals and implementation of cybersecurity technology, and those who are only passively affected by it. Therefore, it is of utmost importance to take a human-centered approach, i.e., to ask how individuals or groups in society have their rights impacted or even violated by the functionality of cybersecurity technology and the practices the technology is used, for instance, by companies and cybersecurity agencies. Consequently, questions of privacy, fairness, accessibility, inclusiveness and non-discrimination are important in this context, but also broader issues of (social) sustainability [42, sect. 2.1].

The lack of people-centeredness has repeatedly led to technical solutions being controversial, not being used as expected or even failing. Particularly in the case of very complex systems with an

[3] A similar approach has been taken by the IEEE Systems and Software Engineering Standards Committee in the development of the 7000-2021 standard for addressing ethical concerns during system design [38; 65].

infrastructure character, this has led to the insight that it is not only the level of technical maturity (or readiness) that is important in the research and development process. Rather, in a holistic view, the degree of maturity from the perspective of other actor groups is also important. These are first and foremost the affected citizens, but in the same way it can also be asked how suitable a technology is for organizations such as companies that are to adapt it [12]. This is the point of departure for the concept of assessing the societal readiness that is in the center of the reflection process that we will present in the next section.

Chapter 3 Supporting institutionalisation of responsible R&D in cybersecurity with ELSA and RRI

Institutionalizing the conditions and objectives of RRI is the *conditio sine qua non* for the successful development of cyber security research to reach higher levels of social responsibility. In view of the importance of institutionalization processes as a prerequisite for change, both academic research and more practice-oriented projects have devoted considerable resources in generating a better understanding of the conditions and mechanisms supporting and hindering institutionalization. It is essential to place cybersecurity at the service of society, to improve knowledge about the risks and opportunities of cyber and information security systems and to emphasize the strategic importance of security and privacy for society as a whole. Against this backdrop, it is elementary to operate R&D processes which consider the procedural RRI dimensions of anticipation, reflexivity, inclusion, and responsiveness and translate them into practices. At the same time, it is equally essential to embed the assessment and improvement of societal readiness in the emerging or changing organizational structures of cybersecurity in Europe. Only in this way can we address with sufficient vigor the issues of "(ethical) acceptability, sustainability and societal desirability of the innovation process and its marketable products" [59, p. 72], which are important for societal resilience.

Randles et al. [56; 57; 58] developed the theoretical approach of deep institutionalization to better analyze processes of institutional change in organizations - their conditions, barriers and drivers. This section briefly outlines the concept of deep institutionalization to provide guidance for advancing the organizational conditions of cybersecurity R&D that are conducive to the ambitions of responsible and society-centered technology design. Deep Institutionalization - in contrast to shallow institutionalization and window dressing - means the embedding of values (normative orientations) in the design, practice, and identity of organizations. This can be summarized as a model of descriptive theory of organizations and described in terms of four main characteristics [58]:

> *"(a) its long-haul, long-term and resilient nature; (b) its transformative dynamic: affecting institutionalization which creatively destroys prior orientations; (c) its inter-dependent systemic nature, comprising integrated and mutually supporting infrastructures of technologies, social norms and routines, governance tools as well as economic and ideological logics; (d) is a methodological point: that we cannot truly evaluate the effectiveness of transformation towards particular normative goals ex-ante"* (p. 61)

Against this background, it becomes clear that the adoption of RRI is not only dependent on the internal logics and institutional context of the organization, but lies in a thorough analysis of the external environment. For a profound institutionalization, Randles defines three supportive orientations ("Multi-level alignments") [56, p. 38]:

1. *"Institutional context & external conditional factors by which trends, pressures, challenges and opportunities (including role of the State and other forms of 'shadow hierarchy') are meant;*
2. *intra-organizational translation and 'getting along' with multiple institutional logics, within different units and functions of large/complex forms of 'organizational pluralism;*
3. *Institutional entrepreneurship and /or forms and expressions of leadership as well as intermediation at different levels of the organization is defined as the orientation" (p. 40)*

In view of these complex, multi-faceted conditions of institutionalization processes, they propose a concept composed of four elements or axes that should be addressed if deep institutionalization is to be reached:

- Evolution of dominant narratives: Instead of replacing previously existing understandings of responsibility in an organization, the deep institutionalization concept assumes a process of sedimentation of new understandings over the existing ones. In effect, the at first dominant organizational narrative co-exists and overlaps with the new responsibility understandings and narratives, often entailing tensions and contradictions [19; 56]. For actors attempting to institutionalize

an understanding of RRI in an organization, the processes of sedimentation and the resulting hybrids of old and new responsibility values needs to be taken into account.

- Maturation process: The maturation process covers different levels of maturity, from emergence over maturity to resilience. While “emergence” refers to responsibility understandings “in the making”, “maturity” already denotes a general acceptance of RRI in an organization. This can be reflected, for instance, by RRI principles being embedded in organizational routines, practices and incentive structures. “Resilience”, finally, refers to a stage at which the newly institutionalized responsibility understanding is resilient even in the face of irritations and challenges [19; 56]. This perspective on the different phases of an institutionalization process is important to keep in mind when embarking on a change process in an organization. Depending on the respective maturity level, different measures need to be taken.
- Systemic consolidation: This axes refers to situations in which an organization moves from mostly ad-hoc, isolated and largely independent experiments and expressions of new responsibility understandings to broadly shared and routinised norms, practices and rules within the organization [19; 56]. Again, this perspective can be useful for institutional entrepreneurs seeking the uptake of RRI in an organization and who might have to start the process by supporting coalitions of the willing, then actively connecting different parts of the organization moving forward, and eventually turning to activities supporting the scaling-up and broadening out of RRI.
- Vertical multi-level alignment: The last element of the concept refers to the relationship of an organization’s responsibility understanding to its external environment. As organizations constantly interact with their environments, the coherence between the responsibility values of the organization holds and the context within which it operates can cause tensions [19; 56]. Depending on the ability and power of an organization to either exert influence on the environment and/or resist external influences that are in contrast to the sought responsibility values, different organizational strategies are required.

Thus, the *Deep Institutionalization* concept can provide guidance for the integration of RRI in organizations within the cybersecurity context. Conditions of the macro and meso levels of interaction arise for this (see Figure 2):

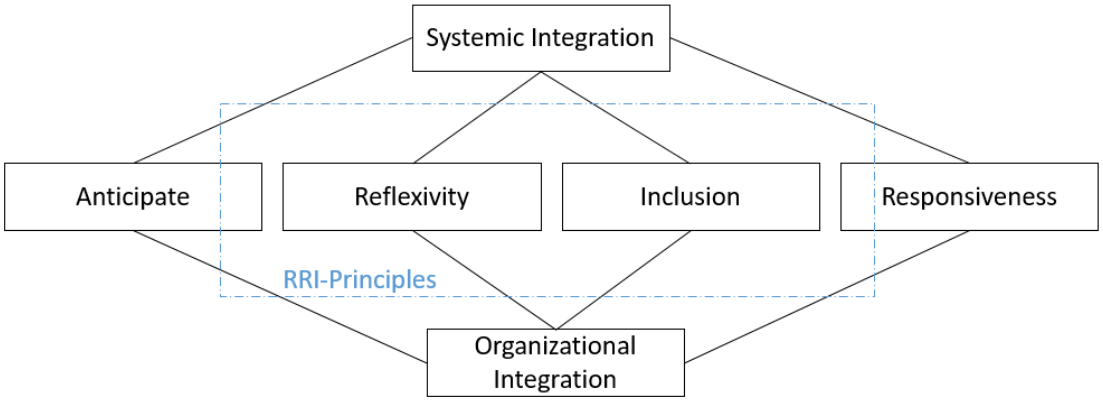

Figure 2: Context and organizational preconditions for Institutionalization of RRI

Source: Own illustration.

These conditions can also act as guardrails in the context of the far-reaching institutionalization of RRI (and ELSA) with a view to cybersecurity and be imposed as reflexive conditions on R&D projects. If the normative goals formulated with the preconditions are taken into account from the outset in the project or process context, on the one hand, the potential for aligning innovative action with societal goals without disregarding aspects of value creation increases (macro). On the other hand, this could also promote the institutionalization of ethical aspects in the organizational context (meso). In view of the upcoming closure of the SPARTA project, the mentioned characteristics could be a forward-looking answer to the procedural question of how to not only talk but drive a "Deep Institutionalization" of RRI aspects in cybersecurity projects. This brings up the question of specific RRI principles in the cybersecurity context, which will be explained in the next chapter.

Chapter 4 A gate-based process for the systematic assessment and improvement of societal readiness

4.1 Overview of the tool

This section outlines the general rationale of the tool, how it aims to support reflexivity, who should use the tool and how, and also show the limitations of the tool.

While the ambitions of RRI are well documented and have found ample representation both at the policy level (e.g., RRI as a cross-cutting theme in Horizon 2020, RRI principles implemented in research councils) as well as in the academic discourse, rather practice-oriented approaches and tools supporting the effective integration of responsibility principles in research and innovation processes remain scarce. Tools offering support to research and innovation practitioners who aim at higher levels of responsibility are particularly important in the context of technology development due to the often-cited control (or Collingridge) dilemma [17; 27]. Due to the far-reaching information deficits about the future implications of a technology on the one hand, and the high degree of diversity and openness of most contemporary technology development processes on the other, the establishment of detailed frameworks, rules and standards aiming to control and regulate technology more often than not misses the problem at hand. While such regulatory approaches are undoubtedly essential building blocks of any responsible technology governance framework, particularly the early stages of technology development require approaches that are more flexible. To address this challenge, the core of the RRI conceptual thinking is directed at actively supporting reflexivity, anticipation and responsiveness at the level of the actual research and innovation processes (cf. Chapter 3 of this del.). With the aim of turning these abstract principles into a useful tool providing practical orientation, the Societal Readiness Thinking Tool was developed. In essence, the Thinking Tool has the purpose to support researchers to reflect on the societal implications of their research and innovation activities, thereby increasing the integration of societal values and ethical considerations and reducing harmful effects of new technologies and applications.

The Thinking Tool was developed by the EU-funded project NewHoRRIzon (“Excellence in science and innovation for Europe by adopting the concept of Responsible Research and Innovation”). Initially inspired by the well-established Technology Readiness Level (TRL) concept [47], NewHoRRIzon set out to develop a practical tool to ensure higher levels of "societal readiness" of research and innovation projects. However, early in the conceptual development process of the Tool, it became clear that both the TRL's linear notion of technology development and the definition of distinct "levels" were inadequate to fulfil the goal of providing practical guidance to actors wanting to better integrate broader responsibility considerations in their practices [15]. Particularly the challenges of providing practical and concrete process support to researchers and innovators with regard the manifold uncertainties of societal and ethical implications of their work, led the developers of the Tool to emphasise "thinking" and "reflection" [6]. Such an approach was seen as significantly more realistic and context-sensitive compared to the attempt of developing a one-size-fits-all framework based on predefined criteria and operationalised, for example, by means of a checklist.

Thus, to fulfil the requirements of the Tool as a practitioner-oriented resource supporting the integration of societal and ethical perspectives, the general design of the Tool is based on the following key features [52]:

1. Adoption of a stage-gating approach, covering the main phases of research and innovation projects,
2. supporting reflexivity and thinking across all phases of a project lifecycle and all RRI dimensions by means of thought-provoking questions, addressing all RRI dimensions,
3. provision of potentially useful methods and resources to be applied by the users of the Tool.

At the final stages of the development process, the Tool was turned into an online resource that can be used free of charge.

4.1.1 Stage-gate approach

A main structuring feature of the Thinking Tool is the stage-gate approach. Inspired by a model widely used in commercial product development [18], the Tool distinguishes between four distinct stages or phases of the research and innovation process. Each of the phases is separated by a "decision gate" at which the users of the Tool are provided with questions intended to support reflection about the "societal readiness" of their specific project at this stage. The four stages of the Tool are (1) Research design and problem formulation, (2) Implementation, data collection and testing, (3) Data analysis and evaluation, (4) Launching and dissemination (see Figure 3)

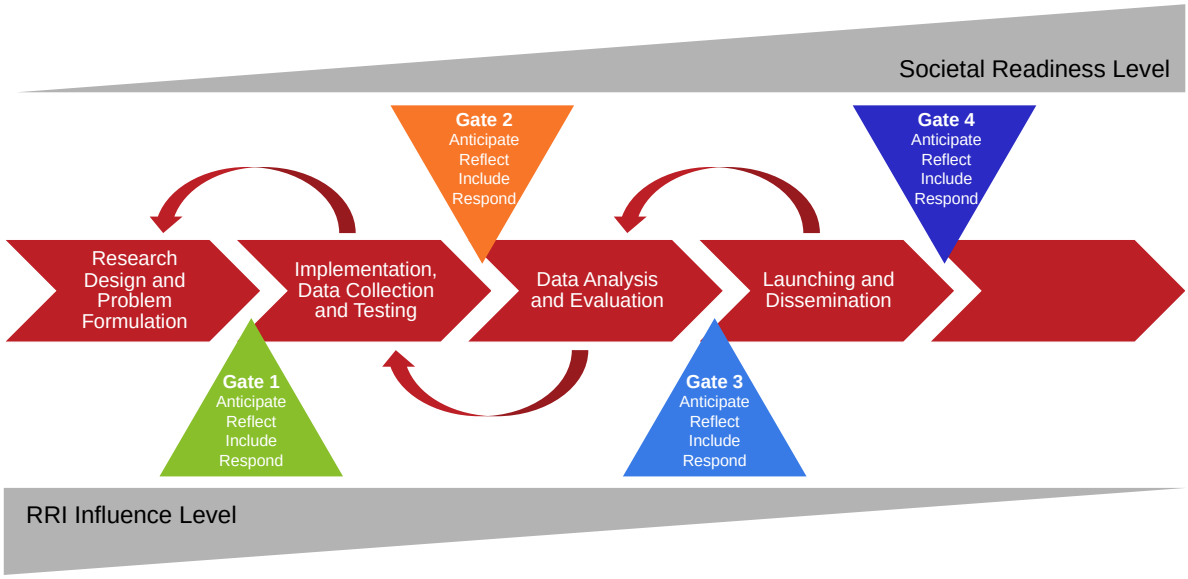

Figure 3: Elements of a RRI based reflection process.

The users of the Tool are encouraged to consider the different questions posed at each of the stages before passing the gate and proceeding to the next phase. While this stage-gate approach implies a rather linear understanding of research and innovation processes, the developers of the Tool emphasise that "In reality, these phases are not always consecutive" [52, p. 11]. Thus, in the actual application of the Tool, it is recommended to iterate between the four phases, depending on the respective needs, necessary improvements and responsibility ambitions of the users and the project at hand.

4.1.2 Reflexivity enhancing and guiding questions

The core of the Tool is based on responsibility-related questions intended to spark critical thinking about the full spectrum of RRI dimensions and perspectives. To this end, the questions were developed by drawing on the results of a review of key RRI literature and further developing them for the purpose of the Tool [52]. The questions are structured along the four stage gates and cover the four RRI process dimensions “anticipate”, “reflect”, “include” and “respond”. Moreover, the guiding questions systematically refer to the so-called RRI keys (public engagement, open access, science education, gender and ethics) established by the European Commission in an attempt to operationalise RRI. As a result, the Tool is currently composed of 89 questions. With regard to the present deliverable, the questionnaire served as a reflection framework for the responsible development of cybersecurity applications, taking into account multiple implications along the RRI key elements

4.1.3 Methods and resources provided

Another key feature of the Tool is the provision of methods and resources intended to support research projects to actually reach their responsibility goals. Examples of these methods include stakeholder mapping, value-based design, ethical impact assessment or public engagement. In total, the

Tool currently offers 28 methods and resources. Each of the methods and resources is briefly explained and, where possible, linked to appropriate and more detailed online resources. Also, depending on the type of question posed to the user, the Tool provides the relevant methods and resources [6].

4.2 A Thinking Tool to support societal readiness in cybersecurity R&D

The "Societal Readiness (SR) Thinking Tool" is designed to align research activities more closely with societal values, needs and expectations. For this purpose, the tool combines the process-related dimensions of RRI according to Stilgoe et al. [70], as well as the key operational elements for RRI defined by the European Commission [26]: *public engagement, gender equality, science education, open access and ethics*. The tool can also prove very useful in the context of developing software for cybersecurity applications by improving the integration of RRI, sustainability, and design thinking into research and innovation cycles. In the following, the operationalization of the tool for application to cybersecurity software is presented. For this purpose, the responsibility-related questions, which are intended to stimulate critical thinking across the spectrum of RRI dimensions and perspectives, have been specified in relation to the use case. First, the key RRI elements (public engagement, gender equity, science education, open access, and governance) are outlined against the background of cybersecurity applications. Then, the stage-gate approach of the SRT will be applied to research and innovation processes in the field of cybersecurity, followed by an overview of the model in its complexity.

Public Engagement and Science Education

Public engagement is understood as the involvement of the population in strengthening research and innovation processes in the field of cybersecurity. For our application, science education is understood as a didactic approach within the framework of public engagement, so that it can contribute to a "democratization of technology developments by conveying knowledge to the public and thus creating awareness" [61, p. 7].

Open Access

Open Access includes free online access to data and publications of publicly funded research in the field of cybersecurity. Open Access addresses the general public with the purpose of "breaking down boundaries between research and society and making it more porous" [36, p. 26].

Diversity

In the context of cybersecurity innovation processes, we use the term diversity (instead of gender equality). The reason for this is the current discussion about the lack of diversity in the development of cybersecurity applications (cybersecurity profession[4]) , but also the homogeneous data basis for such applications (data diversity), which may lead to biases.

Ethics

The key factor "ethics" targets a shared framework of values in the European community to be able to address societal challenges adequately. According to the European Commission [26], ethics in the context of cybersecurity innovation processes should not be seen as an inhibitor to development processes, but as a way to ensure high-quality outcomes.

[4] https://www.forbes.com/sites/forbestechcouncil/2021/04/23/how-to-address-the-lack-of-diversity-in-cybersecurity/?sh=6252ecae2255 or https://www.ukcybersecuritycouncil.org.uk/diversity-inclusivity/

4.3 Operationalization of Questions in context of Cybersecurity

As presented in Chapter 4.1, Bernstein et al. operationalize the SR Thinking Tool "through generic reflective questions intended to aid identification and accounting for key societal dimensions of innovation at different stages of a project" [6, p. 5]. Based on this approach, the questions within the gates were adapted to the specifics of cybersecurity innovation projects. The operationalization of the tool's questions was formulated on the theoretical background of the RRI dimensions and the specifics of the RRI key elements (Table 2). In addition, the ethical principles identified in section 2.3 were integrated for the operationalization of the questionnaire in the context of cybersecurity.

<table>
<tr><th colspan="5">RRI-Framework</th></tr>
<tr><th rowspan="2">RRI Key Factors</th><th colspan="4">RRI Dimensions</th></tr>
<tr><th>Anticipation</th><th>Reflexivity</th><th>Inclusion</th><th>Responsiveness</th></tr>
<tr><td>Public Engagement
Knowledge transfer and awareness building</td><td rowspan="4">Implications

Strategic character

Multidisciplinary cooperation</td><td rowspan="4">Reflective ability

Critical handling of knowledge

Dealing with complexity</td><td rowspan="4">Involvement of stakeholders

Participation</td><td rowspan="4">Adaptability

Monitoring the outcome</td></tr>
<tr><td>Open Access
Access to knowledge</td></tr>
<tr><td>Diversity
Dealing with diversity</td></tr>
<tr><td>Ethics
Quality and value orientation</td></tr>
<tr><td>Values / Ethical principles
for developing cybersecurity solutions</td><td colspan="4">Trustworthiness Safety and reliability Human-centricity</td></tr>
</table>

Table 2: SRT-Framework to match key factors, RRI-Dimensions and ethical values in cybersecurity R&D

Source: Own illustration.

4.3.1 Public engagement and science education

Public engagement is the involvement of the general public in strengthening research and innovation processes in the field of cybersecurity. This seems to be particularly relevant in the field of cybersecurity, as the topic is complex, but at the same time affects many stakeholders and society as a whole. In order to face societal challenges from cyber-attacks right now, science education, communication and knowledge transfer seem to be essential to build up the cybersecurity industry in a targeted and robust way, while considering societal concerns. In addition, the problem-centered questions of the RRI key factors public engagement and science education of the gates aim at aspects of knowledge transfer, how this can be shaped and how it can succeed with the involvement of different stakeholders. A multidimensional perspective is always adopted by exploring options for different stakeholders and their participation in the development process.

4.3.1.1 Gate 1: Research design and problem formulation

Anticipate:

- How will you ensure that you maintain good relations with your stakeholders?
- At which phases in the project will stakeholder involvement have the most crucial impact, and why?
- Who will be the primary users/ beneficiaries/ stakeholder of the project, and could this change?
- Who bears the social costs?
- Who might be excluded from the benefits of the project, and how will you address this?

- How will different stakeholders benefit from your project?
- What are the disadvantages for stakeholders/What are the disadvantages of stakeholder engagement?
- Will any potentially relevant stakeholder be missed by the selected method for identifying requirements?
- How might the project benefit from involving stakeholders in identifying methods for software development process (Requirements, Design and Implementation)?
- Will the project contribute new knowledge of relevance for science education, and how?
- Could your project benefit from involving citizens in data collection and analysis, and how?

Reflect

- Have you considered alternative definitions of and approaches to the problem at stake?
- Who are the relevant stakeholders of your project?
- Have relevant stakeholders been involved in defining the research problem?
- Can RRI perspectives be integrated into the training and supervision of project staff, and how?
- What would it take to better accommodate citizens interested in contributing to your work, and how?
- How do you plan to communicate the uncertainty of your research?

Include

- Which actions will be taken to involve all potentially relevant stakeholders including researchers, representatives from industry, policy-makers and civil society actors in the project?
- Which stakeholders will take part in the project's education and training activities, and why?
- Will your education and communication activities be tailored to specific stakeholder groups, and which ones?

Respond

- Is it possible to change problem formulation or project design in response to changing stakeholder viewpoints or unforeseen ethical issues arising throughout the project?

4.3.1.2 Gate 2: Requirements and software design, implementation and verification

Anticipate

- Will any potentially relevant stakeholder be missed by the selected method for identifying requirements?
- How might the project benefit from involving stakeholders in identifying methods for software development process (Requirements, Design and Implementation)?
- Will the project contribute new methods and techniques of relevance for other researchers and practitioners?

Reflect

- Who have been involved in designing methods for the requirement, design and Implementation of the software? How has the nature and purpose of the project been communicated to external stakeholders?
- Did the requirements gathering and design process give rise to new consideration about potentially relevant stakeholders, and which ones?
- Will it be possible for interested citizens to contribute to the collection of data, and how?
- How can you ensure that interested stakeholders understand the purpose and approaches of the project?

Include

- How will you ensure that all stakeholders feel empowered to voice their opinion?
- How will you ensure that all relevant stakeholders have the information they need to engage in a meaningful dialogue about proper procedures for data collection and testing?

- With whom will you share potential documentations of software development process?
- Which stakeholders are taking part in your education activities, and why?
- If your project contributes new methods and techniques of relevance for other researchers and practitioners, how do you plan to support the education of these groups?

Respond

- Is it possible to change procedures for requirements gathering, design and implementation process in response to ethical issues or stakeholder viewpoints in this phase?

4.3.1.3 Gate 3: Data analysis and evaluation

Anticipate

- Which stakeholders will benefit from your result and which will not?
- Will the project contribute new analytical and evaluative methods of relevance for other researchers and practitioners, and how do you plan to support this?
- What do people not participating in the project (teachers, students museums, Civil society organizations) need to know about the data analysis and evaluation of project results to learn about/engage with the outcomes of your group?

Reflect

- Who have been involved in data analysis and evaluation, and why? Did the data-analysis and evaluation give rise to new considerations about potentially relevant stakeholders, and which?

Include

- How will you ensure that all stakeholders have the information they need to engage in a meaningful dialogue about data analysis and evaluation?
- With whom will you share the documentation of your analysis and evaluation?

Respond

- Is it possible to change procedures for data analysis and evaluation of project results in response to ethical issues or stakeholder viewpoints in this phase?

4.3.1.4 Gate 4: Launching and dissemination

Anticipate

- How can your stakeholder engagement experiences inform future engagement activities in your research area?
- How may your results contribute to the public interest in an understanding of science?
- How may the results of this project be used in the education of future generations of researchers and engineers?

Reflect

- Does your dissemination plan address the stakeholders of the project?
- Which other communication channels than peer-reviewed journals will you use to communicate your work?

Include

- Is your dissemination plan tailored to the needs and characteristics of specific stakeholder groups?
- Will the results of your project be communicated for science education purposes in other languages than English?

Respond

- Is it possible to change your launching and dissemination activities in response to needs and concerns of societal actors?

4.3.2 Open Source

The role of Open Source in the field of responsible Cybersecurity applications is characterized by aspects of transparent software development and accessibility of applications by various stakeholders. In addition, the ability to integrate different perspectives into the development process seems to be crucial for the society-centered design of cybersecurity software. The questions posed by Gates under the key element of open source are aimed further at aspects of participation, so that a holistic view of aspects of social inclusion is made possible at an early stage in the development process of cyber security software.

4.3.2.1 Gate 1: Research design and problem formulation

Anticipate

- Which aspects of the project do you plan to make open access?
- What can you do to ensure that all project partners comply with your open-access strategy?

Reflect

- What are the potential barriers to making your data, coding and publications open access and how could these barriers be addressed?

Include

- What can be done to make proceedings and the final results of your project easily accessible and intelligible to a diverse set of stakeholders?

4.3.2.2 Gate 2: Requirements and software design, implementation and verification

Anticipate

- How may the selected methods for requirements gathering, design and implementation best be documented to ensure transparency and allow for replication and knowledge transfer?

Reflect

- How do you plan to document your methods for the software development process in an intelligible and transparent way?
- What are the potential barriers to making documentations of requirements gathering, design and implementation publicly accessible (e.g. intellectual property rights, competing interests)?

Include

- How will you ensure that all stakeholders feel empowered to voice their opinion?
- How will you ensure that all relevant stakeholders have the information they need to engage in a meaningful dialogue about proper procedures for data collection and testing?
- With whom will you share potential documentations of data collection and testing?

Respond

- Is it possible to change procedures for requirements gathering, design and implementation in response to ethical issues or stakeholder viewpoints in this phase?

4.3.2.3 Gate 3: Data analysis and evaluation

Anticipate

- How may the requirements gathering, design and implementation best be documented to ensure transparency and allow for replication and knowledge transfer?

Reflect

- Did you document your data analysis / evaluation in an intelligible and transparent way, and how?
- What are the potential barriers to making code-scripts and documentation of the full analysis publicly accessible (e.g. intellectual property rights, competing interests, confidentiality etc.)?

Include

- With whom will you share the documentation of your analysis and evaluation?

4.3.2.4 Gate 4: Launching and dissemination

Anticipate

- Who will be responsible for maintenance and storage of the open-access information after the project ends, and for how long? Could the data collected as part of this project be useful for other research purposes, and which?

Reflect

- Is the open access information accompanied by clear and transparent documentation of data editing, statistical procedures and analytical decisions made through-out the project?
- Is the information made open access accompanied by clear specifications on data structure and variable descriptions to allow for replications or new research purposes?

Include

- Is licensed software required to benefit from your open access information?

4.3.3 Diversity

Closely related to the consideration of social participation of diverse social groups and perspectives in the process of developing cybersecurity applications is the one of diversity. The questions in the four gates under the key factor combine aspects of the gender gap to be bridged in cybersecurity workforce. There is also a focus on addressing *diversity* in the design process of applications. On a societal level, the questions also serve to reflect the impact of gender equality through appropriate cybersecurity applications for society.

4.3.3.1 Gate 1: Research design and problem formulation

Anticipate

- Could the outcomes of this project benefit from incorporating a gender dimension into research content, and how?

Reflect

- What are the barriers to gender balance among researchers and leaders in this project and how can these be addressed?
- What are the possible gender and sex dimensions of the problem at stake?

Include

- What can be done to ensure gender balance among researchers and leaders in this project?
- What can be done to ensure gender diversity among research subjects?

4.3.3.2 Gate 2: Requirements, software/system design, implementation & verification

Anticipate

- Will the selected methods for requirements gathering, design and implementation allow for nuanced analysis of possible gender and sex-related differences and similarities?

Reflect

- Have gender and sex related issues been taken into consideration in the selected methods for data collection and testing, and how?
- What is the composition of the team members involved in the design process?

Include

- How do you plan to identify participants that do not identify as men or women (e.g. non-binary or gender fluid subject) in the data collection?

4.3.3.2.1 Gate 3: Data analysis and evaluation

Anticipate

- How may your system solutions impact gender norms and gender relations in society?

Reflect

- Has your data analysis paid attention to possible gender and sex related differences and similarities, and how?

Include

- Have you analyzed possible interactions between gender and sex and other sociodemographic variables such as class, ethnicity, race, nationality and age, and how?

4.3.3.3 Gate 4: Launching and dissemination

Anticipate

- What impact do expect your project will have on gender equality?

Reflect

- What can be done to help support the future career of both men and women junior scholars in the project?

Include

- How will you communicate your results in a way that does not reinforce gender stereotypes?

4.3.4 Ethics

In the context of ethical implications of cybersecurity applications, the question arises not only of the exact subject matter of issues and ethical concerns in different time perspectives, but also of appropriate measures that can be taken early in the technology development process. In this context, the question-based reflection framework for researchers and developers ranges from the consideration of moral aspects of data protection, to the preoccupation with informational autonomy and the privacy of individuals and society. The use of the ethics questionnaire can thus sharpen the focus on relating the importance of IT security for consumers and individuals in general, also to the trust and responsibility of concrete products.

4.3.4.1 Gate 1: Research design and problem formulation

Anticipate

- What ethical issues/ ethical concerns in online privacy and data security/ value conflicts could your project potentially give rise to?
- Can you imagine possible scenarios where the outcomes of the project may be misrepresented or misconstrued in the public discussion?
- To what extent will you be able to predict the long-term societal outcomes of the project?

Reflect

- Which actions should be taken to ensure research integrity and compliance with professional guidelines in the project?
- Does your project involve any risks of negative impacts or value conflicts, and which?

Include

- Who will be involved in identifying the ethical issues and possible solutions to these issues in your project, and how?

- Which actions will be taken to ensure diverse perspectives on the potential ethical issues arising in your project?

4.3.4.2 Gate 2: Requirements and software design, implementation and verification

Anticipate

- Can you think of potential scenarios of misuse of the research outcomes?
- What are the foreseeable changes in the principles/society standards for the transmission of information through the software?

Reflect

- Is the planned research methodology ethically acceptable, including aspects related value-sensitive software development?
- Does your software development process require informed consent from the participants?
- Does your project involve any risks of breach of confidentiality and what might they be?

Include

- Who have been involved in identifying the ethics-related issues to be considered in the system design?
- Have certain groups of stakeholders been excluded from the system design due to ethical concerns, and how may this limit your analysis?
- How was the involvement of non-experts as well as expert knowledge in the software development process ensured?

4.3.4.3 Gate 3: Data analysis and evaluation

Anticipate

- Can you think about benefits of your application/system beyond the original scope of your work (for example societal)?
- Can you imagine possible scenarios of misuse?
- Could your findings be misinterpreted, and how?

Reflect

- What ethics-related issues are involved in your data analysis?
- What types of sensitivity analysis have been used to test the robustness of your methods and results?

Include

- Did your analysis devote attention to possible variations across subgroups of participants, and how?

4.3.4.4 Gate 4: Launching and dissemination

Reflect

- How will you brief the participating research subjects about the project results?
- What can be done to ensure that your results are not misrepresented or misinterpreted in the public debate?

Include

- Do you plan to involve possible stakeholders in discussions about the ethical implications of your project results?

Chapter 5 Experiences and conclusions

Finally, the question arises: What are the prerequisites for ethical, legal and social aspects to be taken into account in the desired way in cybersecurity research and development? This is the subject of this chapter.

5.1 Experiences

Though it was not possible to implement large-scale collaborative tests for validation together with the SPARTA programs as originally foreseen it was possible to collect experiences from various sources.

Firstly, staff members of the responsibility activity were also involved in SPARTA's technical programs at the same time and in this way had the opportunity to learn first-hand how the programs dealt with upcoming ELS aspects – if any. These and other interactions that took place between the technical programs and the legal and ethics experts clearly show that a tool for systematic assessment and improvement of social readiness, as presented in chapter 4, promises clear advantages:

1) *Knowledge*: Most researchers in the SPARTA programs were not aware of the legal and ethical issues because they did not have an overview of all the relevant regulations on national and European level. This could be counteracted by early and comprehensive legal and ethical counselling, which is documented in several deliverables by the SPARTA programs.[5]
2) *Anticipation*: Researcher in most of the programs initially had difficulties in seeing links between the solutions they developed and possible legal and ethical challenges. Developers usually assume that their technical solutions, developed with the best intentions, will work and be used as they imagine. They often overlook the fact that the real world is more complex than the laboratory solution and that there is misuse and abuse. For example, in the CAPE Program, it was possible to identify together with the developers what incidents might occur in their platooning scenario that could raise issues of trust and liability.
3) *Reflexivity*: Reflecting the consequences of research and development activities is a particularly difficult matter, as usually a large number of stakeholders are affected and no simple causal chains exist. Furthermore, risks and benefits are usually unequally distributed among stakeholders, so that a - non-trivial - balancing has to take place. This already starts with the seemingly simple question of which problem actually needs to be solved, but which is very important for the character of the technical solution. In the technical environment, for example, it is often assumed that there is a technical solution and the problem is formulated (or translated) accordingly [74]. With the help of a tool that can be used to systematically examine and assess possible impacts, learning effects could be observed, leading to improved architectures and implementations.
4) *Inclusion*: We also observed that developers often have a very clear, but often too simple, idea of who users and stakeholders are and how they interact with technology and that the (perceived) benefits are the same for everyone. It was even expressed that on occasions when an exchange with users and stakeholders actually took place (see below), interesting new aspects emerged that had not been considered in the development so far.

Secondly, the T-SHARK program had implemented an elaborated governance process that also addressed – albeit marginally – ELS issues [41, section 3.2.2.1]. To do this, they established an "arbitrage group" of stakeholders (though not directly involving citizens as affected persons) and used a stage-gate procedure and tested the concept of Societal Readiness Levels. Finally, at the end of the project, interviews with SPARTA program and task leaders captured the experiences in dealing with ELS issues that arose during the project duration.

[5] E.g. T-SHARK [53], CAPE [50], HAII-T [4] SAFAIR [1; 55]

As mentioned at the outset, the successful integration of ELSA in cybersecurity research depends on close, intensive and trusting cooperation of the ELSA teams with management, developers and other internal as well as external stakeholders. In the following sections, recommendations are made on how this can be achieved.

5.2 ELSA Governance

Most importantly, addressing ELSA should not be seen as a marginal or additional activity. This has a number of implications:

- ELSA must be considered an *integral part* of the overall development plan from the outset and taken as seriously as functional requirements during implementation. A dedicated work package gives ELSA adequate visibility within the governance of the overall project, but should still be closely linked to the project management. In this way, the handling of ELSA is not perceived as a mere service - this leads to a stronger identification with the mission of the overall project.
- This also means that this task must not be delegated to an advisory board that meets only a few times during a project, that has only a few budgeted working days or is even voluntary. Instead the position of a “Responsibility Leader” or “Ethics Manager” should be established [41, sect. 4.1.1].
- Integration should take place not only among research actors in academia or industry but also among research funders (as with the nascent ECCC/NCCs).
- In order to actually influence the actual design decisions, the concept of the *embedded social scientist* who are continuously involved in the development process and thus much closer to the decision-making locus, has proven its value. These researchers are no longer just distanced critical observers, but become more influential *insider agents* who identify problematic aspects of the developed technology at an early stage and work out possible solutions together with the developers.
- It is fundamental to have SSH and legal experts involved who can analyze developments from the *political and social spheres* in terms of their relevance to specific topics.

5.3 Successful implementation of the ELSA reflection process

Even though we were not able to collect any new empirical results, it became clear in principle that a suitable procedure, such as a stage-gate process, is useful for systematically integrating the consideration of ELSA into the development process of organizations. However, there are a number of aspects that need to be considered:

- The right degree of specificity: It is naïve to believe that there is a small number of guiding rules that are universally valid and implementable at the same time – a critique recently also formulated about the numerous "ethics guidelines" for AI [29]. As cybersecurity technology is extremely complex, tools such as scenarios need to elaborate concrete application-specific questions and to illustrate alternative solutions with their functions, impacts, etc.
- Engagement and participation: The assessment should be carried out by a deliberation group involving developers and selected other stakeholders. Its success depends on all relevant stakeholders being involved and able to provide their input in a low-threshold and effective manner:
    - Participation in this group should include representatives of the main stakeholder group (affected people/civil society, end users, industry, science, policy)
    - The group needs to be big enough so that all interests are represented and small enough to be workable – 15 people is a reasonable size.

- The members of the deliberation group should have an intrinsic motivation to participate. Therefore, the group should not be composed along "political" lines. This also helps to ensure that there is not constant changing within the group.
- The group should have a broad mandate, i.e., the scope for decision-making should not be restricted from the outset.
- The deliberation group must be able to continuously monitor development progress in order to make timely decisions; however, the number of consultations must be limited to avoid fatigue.
- For issues that are not covered by the deliberation group's expertise, a corresponding external input is necessary. It makes sense to make use of the "embedded social/legal scientist" to identify topics and prepare such inputs.

- Putting ELSA into action: Ensure that the monitoring group does not only make general recommendations but is also involved in specific design decisions.

5.4 Ensuring conducive conditions for ELSA and RRI

All the above require framework conditions and organisational settings that support the effective and sustainable integration of ELSA ambitions in cyber security research and development. Without serious attempts to deeply institutionalise the values of responsible cyber security R&D and ELSA practices in the relevant research performing and research funding organisations, many ELSA efforts run the risk of being reduced to one-off, unsustainable initiatives or decaying to various forms of window dressing and "responsibility washing". Thus, in order to support processes of ELSA and RRI institutionalisation, organisations should:

- Take effective measure to ensure that the values and ambitions of responsible cybersecurity R&D and related ELS activities are seriously integrated in the organizations' mission statements, actively welcomed by leadership, and eventually become enshrined in the organizational culture.
- Support ELSA practices by providing adequate resources to staff members, incentives, but also training opportunities.
- Be prepared that processes of institutionalization require time and effort, thus calling for staying power and perseverance.

In summary, there is still much to be done. On the one hand, there is still a need for research: the assessment framework presented in this deliverable needs to be further tested, validated and, if necessary, developed to provide a robust tool that meets the needs of researchers and developers in academia, industry and public administration.[6] This concerns in particular effective procedures for the participation of civil society, however not only CSOs but for genuine citizen participation. On the other hand, there is a need to work with the emerging institutions of cybersecurity research in Europe (ECCC, NCCs) to build structures and define processes that strengthen the consideration of ELSA.

[6] Some SPARTA partners are continuing this research in a future EU-funded project (Acronym: TRANSCEND, Topic: HORIZON-CL3-2021-SSRI-01-05, 2022-25) that aims to develop and test transdisciplinary methods for societal impact assessment and impact creation for security research technologies.

Chapter 6 Bibliography

1. Addad, B., et al., Final version of AI systems security mechanisms and tools, 2021. Deliverable 7.5. SPARTA Project (EU H2020, GA 830892).
2. Akca Prill, M., et al., Diagnosis: RRI in Societal Challenges, 2018. Deliverable 4.1. NewHoRRIzon Project (EU H2020, GA 741402). https://www.thinkingtool.eu/Deliverable_6.1_Final_April%2030_THINKING_TOOL.pdf.
3. Anagnostakis, D., *The European Union-United States cybersecurity relationship: a transatlantic functional cooperation.* Journal of Cyber Policy, 2021. **6**(2): p. 243-261.
4. Armando, A., et al., Security-by-Design Framework for the Intelligent Infrastructure, 2020. Deliverable 6.1. SPARTA Project (EU H2020, GA 830892).
5. Bellamy, R., *Social readiness of adaptation technologies.* Wiley Interdisciplinary Reviews-Climate Change, 2019. **10**(6).
6. Bernstein, M.J., et al., *The Societal Readiness Thinking Tool: A Practical Resource for Maturing the Societal Readiness of Research Projects.* Science and Engineering Ethics, 2022. **28**(1): p. 6.
7. Blok, V., *The emerging concept of responsible innovation. Three reasons why it is questionable and calls for a radical transformation of the concept of innovation*, in *Responsible innovation 2*, B.-J. Koops, et al., Editors. 2015, Springer: Cham. p. 19–35.
8. Brand, T. and V. Blok, *Responsible innovation in business: a critical reflection on deliberative engagement as a central governance mechanism.* Journal of Responsible Innovation, 2019. **6**(1): p. 4-24.
9. Briege, C., L. Mitrou, and Z. Menevidis, Policy Brief: RRI for Security, 2014. Deliverable 6.2. RESPONSIBILITY Project (FP 7, GA 321489). http://responsibility-rri.eu/wp-content/uploads/2013/08/RESPONSIBILITY-D6.2-Policy-Brief-RRI-for-Security_Final-EC-Public.pdf.
10. Brown, T., *Design Thinking*, in *Harvard Business Review*. 2008. p. 84–92.
11. Brundage, M. and D.H. Guston, *Understanding the movement(s) for responsible innovation*, in *International Handbook on Responsible Innovation: A Global Resource*, R. von Schomberg and J. Hankins, Editors. 2019, Edward Elgar: Cheltenham. p. 102-121.
12. Bruno, I., et al., *Technology readiness revisited: a proposal for extending the scope of impact assessment of European public services*, in *ICEGOV 2020: Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance, Athens, Greece*, Y. Charalabidis, M.A. Cunha, and D. Sarantis, Editors. 2020, ACM Press. p. 369-380.
13. Burget, M., E. Bardone, and M. Pedaste, *Definitions and Conceptual Dimensions of Responsible Research and Innovation: A Literature Review.* Science and engineering ethics, 2017. **23**(1): p. 1–19.
14. Bührer, S., et al., Concepts and methods to measure societal impacts – an overview, 2022. Discussion Papers Innovation Systems and Policy Analysis 74. Fraunhofer ISI: Karlsruhe. https://www.isi.fraunhofer.de/content/dam/isi/dokumente/cci/innovation-systems-policy-analysis/2022/discussionpaper_74_2022.pdf.
15. Chan, T.T. and I. Meijer, *Best practice II: Societal Readiness Thinking Tool by NewHoRRIzon*, in *Assessment of Responsible Innovation: Methods and Practices*, E. Yaghmaei and I.v.d. Poel, Editors. 2021, Routledge: London and New York. p. 218-221.
16. Christen, M., et al., *Mapping Values: Using Thesauruses to Reveal Semantic Structures of Cultural Moral Differences.* Cognitive Systems Research, 2016. **40**: p. 59-74.
17. Collingridge, D., *The social control of technology*. 1980, London: Pinter.

18. Cooper, R.G., *Winning at New Products: Creating Value Through Innovation*. 5th ed. 2017, New York: Basic Books.

19. Daimer, S., H. Berghäuser, and R. Lindner, *The Institutionalisation of a New Paradigm at Policy Level*, in *2022 European Forum for Studies of Policies for Research and Innovation (Eu-SPRI 2022)*. 2022: Utrecht. https://euspri2022.nl/wp-content/uploads/sites/556/2022/07/Daimer-et-al_2022-RRI-at-policy-level_EuSPRI-full-paper-1.pdf

20. van de Poel, I., *Core Values and Value Conflicts in Cybersecurity: beyond Privacy versus Security*, in *The Ethics of Cybersecurity*, M. Christen, B. Gordijn, and M. Loi, Editors. 2019, Springer: Cham.

21. van de Poel, I., et al., *Company Strategies for Responsible Research and Innovation (RRI): A Conceptual Model.* Sustainability, 2017. **9**(11): p. 2045.

22. Deibert, R.J., *Toward a Human-Centric Approach to Cybersecurity.* Ethics & International Affairs, 2018. **32**(4): p. 411-424.

23. Domingo-Ferrer, J. and A. Blanco-Justicia, *Ethical Value-Centric Cybersecurity: A Methodology Based on a Value Graph.* Science and Engineeering Ethics, 2020. **26**(3): p. 1267-1285.

24. Dryzek, J.S., *Foundations and Frontiers of Deliberative Governance*. **2011**, Oxford: Oxford University Press.

25. European Commission, 2030 Digital Compass: the European way for the Digital Decade, 2021. COM(2021) 118 final. Brussels.

26. European Commission, *Responsible research and innovation: Europe's ability to respond to societal challenges*. 2012: European Union Publications Office.

27. Genus, A. and A. Stirling, *Collingridge and the dilemma of control: Towards responsible and accountable innovation.* Research Policy, 2018. **47**(1): p. 61-69.

28. Goodin, R.E. and J.S. Dryzek, *Deliberative Impacts: The Macro-Political Uptake of Mini-Publics.* Politics & Society, 2016. **34**(2): p. 219-244.

29. Hagendorff, T., *A Virtue-Based Framework to Support Putting AI Ethics into Practice.* Philosophy & Technology, 2022. **35**(55).

30. Hallensleben, S., et al., From Principles to Practice: An interdisciplinary framework to operationalise AI ethics, 2020. VDE Verband der Elektrotechnik Elektronik Informationstechnik e. V.; Bertelsmann Stiftung: Frankfurt and Gütersloh.

31. He, M., L. Devine, and J. Zhuang, *Perspectives on Cybersecurity Information Sharing among Multiple Stakeholders Using a Decision-Theoretic Approach.* Risk Analysis, 2018. **38**(2): p. 215-225.

32. Héder, M., *From NASA to EU: the evolution of the TRL scale in Public Sector Innovation.* The Innovation Journal: The Public Sector Innovation Journal, 2017. **22**(2): p. article 3.

33. Hellaoui, H., A. Bouabdallah, and M. Koudil, *TAS-IoT: Trust-Based Adaptive Security in the IoT*. 2016. p. 599-602.

34. Hughes, T.P., *The Evolution of Large Technological Systems*, in *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*, W. Bijker, T.P. Hughes, and T. Pinch, Editors. 1987, MIT Press: Cambridge Mass. p. 51-82.

35. Hunt, G., *The Principle of Complementarity: Freedom of Information, Public Accountability and Whistleblowing*, in *Open Government in a theoretical and practical context*, R.A. Chapman and M. Hunt, Editors. 2016, Routledge: London and New York. p. 43-53.

36. Häußermann, J.J. and M. Heidingsfelder, *Offen, verantwortlich und verantwortlich offen.* TATuP - Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis, 2017. **26**(1-2): p. 31-36.

37. Iatridis, K. and D. Schroeder, *Responsible Research and Innovation in Industry: The Case for Corporate Responsibility Tools*. SpringerBriefs in Research and Innovation Governance. 2016, Cham, Heidelberg, New York, Dordrecht, London: Springer.

38. IEEE Computer Society, *IEEE Standard Model Process for Addressing Ethical Concerns during System Design*, in *Systems and Software Engineering Standards Committee*. 2021, Institute of Electrical and Electronics Engineers: New York.

39. Jakobsen, S.-E., A. Fløysand, and J. Overton, *Expanding the field of Responsible Research and Innovation (RRI) – from responsible research to responsible innovation.* European Planning Studies, 2019. **27**(12): p. 2329-2343.

40. Karinen, R. and D.H. Guston, *Toward anticipatory governance: The experience with Nanotechnology. Governing future technologies.* Sociology of the Sciences Yearbook, 2010. **27**: p. 217–232.

41. Kirchner, F., et al., Preparing a CCN pilot to become a sustainable CCN, 2022. Deliverable 1.5. SPARTA Project (EU H2020, GA 830892).

42. Knockaert, M., et al., Ethical, legal and societal aspects, 2020. Deliverable 2.1. SPARTA Project (EU H2020, GA 830892). https://www.sparta.eu/assets/deliverables/SPARTA-D2.1-Ethical-legal-and-societal-aspects-PU-M12.pdf.

43. Kong, J., et al., *Adaptive security for multilevel ad hoc networks.* Wireless Communications and Mobile Computing, 2002. **2**(5): p. 533-547.

44. Kooiman, J.A.N. and S. Jentoft, *Meta-Governance: Values, Norms and Principles, and the Making of Hard Choices.* Public Administration, 2009. **87**(4): p. 818-836.

45. Lindner, R., et al., Addressing directionality: Orientation failure and the systems of innovation heuristic. Towards reflexive governance, 2016. Discussion Papers Innovation Systems and Policy Analysis 52. Fraunhofer ISI: Karlsruhe. https://nbn-resolving.de/urn:nbn:de:0011-n-4087463.

46. Manjikian, M., *Cybersecurity Ethics: An Introduction*. 2018, Lodnon and New York: Routledge.

47. Mankins, J.C., *Technology readiness assessments: A retrospective.* Acta Astronautica, 2009. **65**(9-10): p. 1216-1223.

48. Meijer, I., et al., *Monitoring the Evolution and Benefits of Responsible Research and Innovation (MoRRI) – a preliminary framework for measuring RRI dimensions*, in *OECD Blue Sky Forum on Science and Innovation Indicators*. 2016: Ghent. https://digital.csic.es/bitstream/10261/161892/1/MoRRI-OECD-RRI.pdf

49. van Mierlo, B., P.J. Beers, and A.-C. Hoes, *Inclusion in responsible innovation: revisiting the desirability of opening up.* Journal of Responsible Innovation, 2020. **7**(3): p. 361-383.

50. Mirko, M., et al., Demonstrators evaluation, 2022. Deliverable 5.4. SPARTA Project (EU H2020, GA 830892).

51. Morgan, G. and B. Gordijn, *A Care-Based Stakeholder Approach to Ethics of Cybersecurity in Business*, in *The Ethics of Cybersecurity*, M. Christen, B. Gordijn, and M. Loi, Editors. 2019, Springer: Cham.

52. Nielsen, M.W., et al., Ensuring Societal Readiness. A Thinking Tool, 2017. Deliverable 6.1. NewHoRRIzon Project (EU H2020, GA 741402). https://www.thinkingtool.eu/Deliverable_6.1_Final_April%2030_THINKING_TOOL.pdf.

53. Ohm, M., et al., Cybersecurity Threat Prediction Legal Framework, 2022. Deliverable 4.5. SPARTA Project (EU H2020, GA 830892).

54. Owen, R., P. Macnaghten, and J. Stilgoe, *Responsible research and innovation: From science in society to science for society, with society.* Science and Public Policy, 2012. **39**(6): p. 751-760.

55. Pawlicki, M., et al., Preliminary description of AI systems security mechanisms and tools, 2020. Deliverable 7.2. SPARTA Project (EU H2020, GA 830892).

56. Randles, S., Deepening 'Deep Institutionalisation', 2017. Deliverable 1.2. JERRI Project (EU H2020, GA 709747). https://www.jerri-project.eu/jerri-wAssets/docs/deliverables/wp-1/JERRI_Deliverable_D1_2_Deepening-Deep-Institutionalisation.pdf.

57. Randles, S., et al., *Report of the Roundtable at S.NET Boston 2013: 'Where to Next for Responsible Innovation?'*, in *Innovation and Responsibility: Engaging with New and Emerging Technologies*, C. Coenen, et al., Editors. 2014, IOS Press: Amsterdam. p. 19-38.

58. Randles, S. and O. Laasch, *Theorising the Normative Business Model.* Organization & Environment, 2015. **29**(1): p. 53-73.

59. von Schomberg, R., ed. *Towards Responsible Research and Innovation in the Information and Communication Technologies and Security Technologies Fields*. 2011, Publications Office of the European Union: Luxembourg.

60. Schot, J. and A. Rip, *The Past and Future of Constructive Technology Assessment.* Technological Forecasting and Social Change, 1997. **54**(2/3): p. 251-268.

61. Schuch, K., et al., Verantwortungsbewusste Forschung und Innovation: Begriffsbestimmung, Herausforderungen, Handlungsempfehlungen, 2016. Positionspapier.

62. Schuurbiers, D. and E. Fisher, *Lab-scale intervention.* EMO reports, 2009. **10**(5): p. 424-427.

63. Shane, P.M., *Cybersecurity Policy as If Ordinary Citizens Mattered: The Case for Public Participation in Cyber Policy Making.* I/S: A Journal of Law and Policy for the Information Society, 2012. **8**(2): p. 433-462.

64. Slocum, N., S. Steyaert, and R. Berloznik, *Participatory Methods Toolkit: A practitioner's manual*. 2006, Brussels: King Baudouin Foundation.

65. Spiekermann, S., *From value-lists to value-based engineering with IEEE 7000 (TM).* 2021 Ieee International Symposium on Technology and Society (Istas21): Technological Stewardship & Responsible Innovation, 2021.

66. Spiekermann, S., J. Korunovska, and M. Langheinrich, *Inside the Organization: Why Privacy and Security Engineering Is a Challenge for Engineers.* Proceedings of the IEEE, 2019. **107**(3): p. 600 - 615.

67. Stahl, B.C., *Responsible Research and Innovation: The Role of Privacy in an Emerging Framework.* Science and Public Policy, 2013. **40**(6): p. 708-716.

68. Stahl, B.C., et al., *The Responsible Research and Innovation (RRI) Maturity Model: Linking Theory and Practice.* Sustainability, 2017. **9**(6).

69. Steen, M., *Slow Innovation: the need for reflexivity in Responsible Innovation (RI).* Journal of Responsible Innovation, 2021. **8**(2): p. 254-260.

70. Stilgoe, J., R. Owen, and P. Macnaghten, *Developing a framework for responsible innovation.* Research Policy, 2013. **42**(9): p. 1568-1580.

71. Stovall, P., *Professional virtue and professional self-awareness: a case study in engineering ethics.* Sci Eng Ethics, 2011. **17**(1): p. 109-32.

72. Swierstra, T. and A. Rip, *Nano-ethics as NEST-ethics: Patterns of Moral Argumentation About New and Emerging Science and Technology.* NanoEthics, 2007. **1**(1): p. 3-20.

73. te Kulve, H. and A. Rip, *Constructing productive engagement: pre-engagement tools for emerging technologies.* Science and Engineering Ethics, 2011. **17**(4): p. 699-714.

74. Valkenburg, G., *Security Technologies Versus Citizen Roles?* Science as Culture, 2016. **26**(3): p. 307-329.

75. Valutytė, R., M. Friedewald, and M. Knockaert, Guidelines for responsible cybersecurity research and innovation, 2022. Deliverable 2.6. SPARTA Project (EU H2020, GA 830892).

76. Valutytė, R., et al., Key challenges and promising solution approaches, 2021. Deliverable 2.3. SPARTA Project (EU H2020, GA 830892).

77. Zedler, D., *Zur strategischen Planung von Cyber Security in Deutschland (On Strategic Planning of Cyber Security in Germany).* Zeitschrift für Außen- und Sicherheitspolitik, 2017. **10**(1): p. 67-85.

78. Zwart, H., L. Landeweerd, and A. van Rooij, *Adapt or perish? Assessing the recent shift in the European research funding arena from 'ELSA' to 'RRI'.* Life Sciences, Society and Policy, 2014. **10**(1): p. 11.

Chapter 7 Abbreviations

Abbreviation	Translation
ECCC	European Cybersecurity Competence Center
ELSA / ELSI	Ethical, Legal and Societal Aspects / Issues
NCC	National Coordination Centre for Cybersecurity
NMB	Normative business model
R(R)I	Responsible (Research and) Innovation
R&D	Research & Development
SRL	Societal Readiness Level
TA	Technology Assessment
TRL	Technological Readiness Level