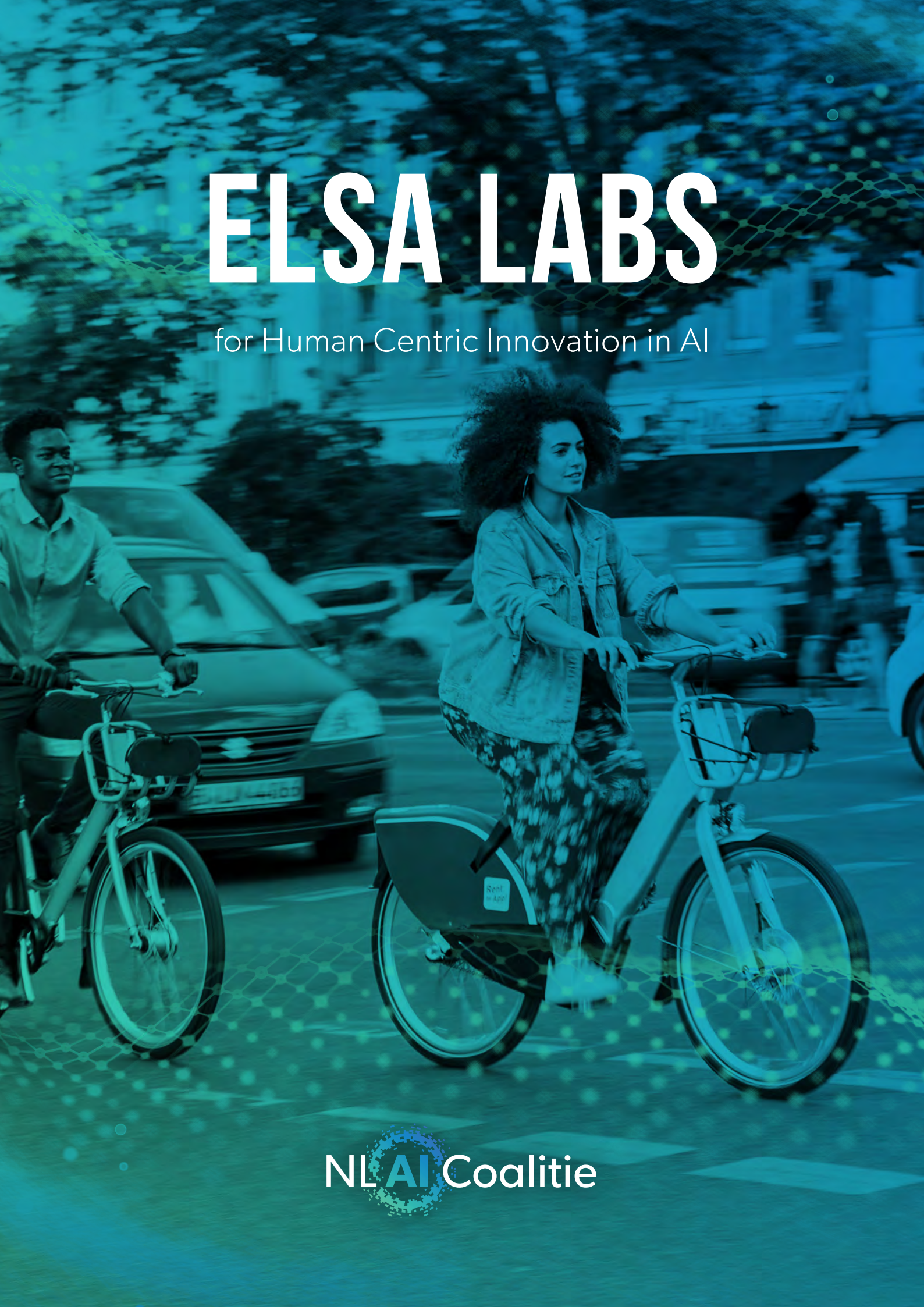
ELSA LABS
for Human Centric Innovation in AI
NL AI Coalitie

1. INTRODUCTION AND CONTEXT

The purpose of this paper is to draw from existing knowledge and experience to underpin the deployment of ELSA Labs as part of the AiNed programme by the Netherlands AI Coalition (NL AIC). Ethical Legal Societal Aspects (ELSA) Labs aim to see to the development and deployment of 'human centric AI', in line with the European focus on AI applications that respect fundamental rights and public values. The National Growth Fund aimed at stimulating long term economic growth and resilience embraced this approach and granted a substantial initial financial allocation of around 10 million euro dedicated to ELSA Labs. The Growth Fund recommends, in addition, to work with the models of Society Readiness Levels and Pathways to Impact, to include all societal and economic stakeholders from the start of AI development.[1]

As the instrument of ELSA Labs in technology development is neither new, nor particular to the Dutch AI agenda, a historical review of the notion of ELSA Labs and an analysis of the critical debates around it will help to understand critical success factors and facilitate the realisation of the agenda as emerging from the AiNed programme. Human centric AI and its different elements, including ethical guidelines for responsible AI; the legal framework including the recently published proposal for a European AI regulation[2] and the societal shaping of technology in society, constitute the starting perspectives on ELSA Labs.

With respect to the conceptual framework that may be used to set up, develop, implement and replicate ELSA Labs, two additional issues that will be elaborated in the next sections, have emerged in the history of ELSA Labs:

1. The notion of 'human centric' AI frames the goals and approaches of designers and developers of AI towards individuals and consumers, but leaves civic goals and public values somewhat unaddressed. Movements like ***AI for public good*** express similar objectives as human-centred AI, but focus on a societal context in which the human is constituted. Thus, they pursue different projects, draw different coalitions and engage in different governance. Hence, ELSA Labs for human centric AI should capture both human and public value.
2. The notions of Societal Readiness Level and the Pathways to Impact recommended by the Growth Fund, are somewhat behind the current, state-of-the-art descriptive and prescriptive articulations of technology and society, which are converging around notions of joint construction, contestation and iteration (in other words: continuous learning), rather than assuming linear growth (readiness) and impact (pathways). This togetherness is captured in concepts like mutual shaping (from science and technology studies), co-creation (from design studies), quadruple-helix (from innovation studies), actor-network-theory (from sociology) and contestable, participatory or value-inclusive design (from HCI studies). Thus, the ELSA 'Lab' way of working should be collaborative, enabling a dynamic learning process that is essential for a rapidly developing technology such as AI.

We will build on these insights and disciplines to elaborate the particular purposes, practices and partners of the ELSA Labs, in a way that produces both societally relevant and responsible AI, and cutting-edge usable knowledge about the way it should come about, be contested, governed and adjusted.

1. Expertrapporten eerste beoordelingsronde 2021 R&D&I, Expert rapporten eerste beoordelingsronde 2021 R&D&I | Rapport | Rijksoverheid.nl, p.7

2. European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Fostering a European approach to Artificial Intelligence, Brussels, 21.4.2021, COM(2021) 205 final.

2. HISTORICAL DEVELOPMENT OF ELSA LABS CONCEPT

The emergence of ELSA Labs in the human genome project

While the relation between society and technology has been a longstanding object of academic research and public debate, the notion of ELSA is relatively young and can be accredited to the researchers of the human genome project of the 1990's, who felt a need to include societal aspects and implications into their biological research. Zwart & Nelis (2009) identify this as the starting point of 'Elsification' of large techno-science programmes. These are typified by a strong interdisciplinary approach, a focus on future developments and public agenda building around emerging technologies, a participatory (inter)action methodology that involves multiple stakeholders and a close proximity to the development of technology at hand (ibid, 2009).

While this description seems relatively straightforward, a closer look at the meaning of ELSA reveals the rather different forms of society-technology interactions that have been envisioned in ELSA Labs. Variously called Ethical, Legal and Societal Aspects, Implications or Issues each of these terms implies a different understanding of the actual and desired relation between society and technology. The concept of 'implications' is the most revealing in this sense as it assumes that a unidirectional causal relation from technology to society is possible. The word 'issues' makes clear that possible tensions and conflicts are bound to emerge in the society-technology articulation. 'Aspects' in contrast seems a more neutral term, which fits the context of the National Growth Fund.

Since the first ELSA Labs in the human genome project, various ELSA activities have taken place around emerging technologies, for instance the European Defense Progam (PADR, 2019), bio- and neuro technology (Forsberg, 2015), nanotechnology (Schillmeijer, 2015), and robotics (ERF, 2017). These have not necessarily taken place in the form of 'labs' but also have a wide variety of other forms like reflections, evaluations or end user research.

Praise and criticism

These subsequent ELSA activities provide useful collections of do's and don'ts, lessons learned and risks to avoid, as well as a particular assessment about their discursive and ideological positioning. All these will be helpful to make the most of the ELSA Labs in the emerging AiNed programme.

The structural inclusion of ELSA in the design and production of emerging technologies, is thought to be a considerable improvement in comparison to past practices in which technology was said to be merely driven by engineers. While this is, evidently, part of the legitimating discourse of ELSA and a somewhat stereotypical representation of the past, ELSA does represent a much wider recognition than before of the need for society to be involved in technical developments and deployment. Not only as a boundary condition, but as a value driver for AI applications. This recognition is represented in the bid of AiNed to the Growth Fund which gives a prominent place to ELSA Labs, but even more so in the priority the Growth Fund gives to ELSA Labs in the overall allocation of funding.

The almost twenty years of experience with ELSA Labs does also present clear risks and challenges. In reflecting on the ELSA Labs in the Norwegian bio-, neuro-, and nano-technology program of the 2010's, Forsberg (2014) identifies a number of recurring disagreements and controversies about how to execute the ELSA Labs, for instance with respect to the balance between applied and fundamental research, between empirical and theoretical projects and the integration between ELS and technical disciplines. Moreover, from the perspective of the researchers, the ELSA in the overall programs were too often designed as add-ons to an already defined innovation programme and as instruments to gain societal legitimacy. Zwart & Nelis (2009) speak of the risk of ELSA Labs of becoming the 'handmaidens' of new technologies, and they quote Hedgecoe (2003) who describes ELSA as "a rhetorical device used to gain support among policy-makers and funders for particular research topics and technologies" (p. 544).

Responsible Research and Innovation

In addition to such criticism, the ELSA approach also lost its prominence due to EU Funding programmes changing their discourse from ELSA to the concept of Responsible Research and Innovation (Zwart, Landeweerd & Rooij, 2014) which prescribes that science and technology should promote "social justice, equality, solidarity and fundamental rights; a competitive social market economy; sustainable development and quality of life" (quoted in Stilgoe & Guston, 2017, pp. 855-856). Rather intensive debates between researchers and practitioners about the supposed differences between ELSA and RRI can be traced in the literature, with proponents of the latter claiming that the RRI approach is more open to social aspects and includes the knowledge about innovation from the philosophy of science.

Thus, the RRI approach is claimed to understand innovation as a multi-actor, iterative process immersed in social relations and tensions, whereas the ELSA approach would assume a more linear process of design, development and implementation. Various other authors, however, agree that there is no fundamental difference between the two and that especially the S in ELSA enables the different emphasis that RRI stands for. They claim, additionally, that the different terms mostly represent a pragmatic adaptation to changing funding language (cf. Forsberg, 2015).[3] Nevertheless, RRI has become the dominant framework to describe and prescribe the relation between society and technology as a recent overview testifies (Timmermans, 2017).

3. See for the whole debate the special issue ELSA and RRI of Life Sciences, Society and Policy, 2015.

3. HUMAN CENTRIC AI: CAPTURING HUMAN AND PUBLIC VALUES

Human centric AI: a multidisciplinary research field

Responsible development and application of data science and AI has emerged as part of several research fields, including information systems, data and computer science and AI, law, philosophy of technology, and science and technology studies. The topic has given rise to multidisciplinary research communities such as the FAccT computer science conference on ***Fairness, Accountability, and Transparency***[4] and the ***Humane AI*** project.[5] Within both communities, a range of disciplines collaborates to develop responsible AI (e.g. Dignum, 2019). Data and computer scientists aim to develop fair AI, e.g. based on the principles laid out by FACT data science: Fairness, Accuracy, Confidentiality and Transparency (Aalst, Bichler & Heinzler, 2017). Some computer scientists search to do so by decreasing bias in datasets, others develop 'human-in-the-loop machine learning', which concerns research on the interaction between humans and AI to allow for optimal development of machine learning-bases AI systems (Xin et al., 2018).

Other research fields, such as Law, Philosophy of Technology, and Science and Technology Studies focus on creating a legal and ethical frameworks for responsible AI (see textboxes on Ethical and Legal aspects), or on the mutual shaping of AI and organisations or society at large (see textbox Societal aspects). Ethical frameworks geared towards ensuring human autonomy guide AI development and application and legal frameworks are developed to enforce human rights. Social shaping of technology is concerned with mutual the impacting forces of technology on organizational and social forces and vice versa (e.g. Orlikowski, 1992); as such, technology – and especially salient regarding AI – may even be attributed agency (Latour, 2005).

4. ACM FAccT (facctconference.org).
5. Humane AI | Human-Centered Artificial Intelligence (humane-ai.eu).

Ethical

Increasing attention is given to which values (and, more broadly, ethics) should be governing data science and AI. This has led to a myriad of frameworks of ethics principles, often formulated by a cross-disciplinary research community including authors from the fields of computer science, philosophy and ethics, the social sciences, legal studies, and various interdisciplinary combinations (Floridi et al., 2018; Veenstra & Timan et al., 2021). One of the best-known being the seven principles of the European High-Level Expert Group for AI (HLEG AI, 2020):

1. human agency and oversight;
2. technical robustness and safety;
3. privacy and data governance;
4. transparency;
5. diversity, non-discrimination and fairness;
6. environmental and societal well-being;
7. accountability

However, normative as well as practical concerns emerge with this wide variety of ethics frameworks. Mittelstadt et al.'s 2016 literature overview of the ethics of algorithms found five overarching epistemics and normative concerns, such as misguided evidence (epistemic) and unfair outcomes (normative). Recently, Veenstra & Timan et al. (2021) compared fifteen well-known ethics frameworks and identified practical concerns regarding the use of these principles, including a lack of consensus on the meaning of the norms and values included in these frameworks, unclarity of how conflicting values should be handled, and a lack of a systemic approach capturing who should be responsible for what.

Therefore, ethical frameworks are often used in combination with ethical approaches to design, such as 'value sensitive design' (cf. Friedman, Kahn & Borning, 2006; Verbeek & Tijink, 2019). These approaches may be used to entangle values into the technology.

Legal

Law, too, has an important role in creating the conditions for AI innovation. The role of law in that context is three-fold: law can be an enabler of investment in innovation by creating legal certainty, incentives for (economic) investment but also the conditions for knowledge sharing and making knowledge more broadly accessible. Intellectual property law is an example hereof, with copyright law but also patent law offering protection and creating temporary monopolies on creations and inventions but also determining under which conditions others can have access to databases, inventions and knowledge.

Another important role of law is to mitigate risks and protect the right of others, including fundamental rights, such as privacy, the right to non-discrimination, autonomy and human dignity or freedom of expression. The European Commission, for example, recently proposed regulation of AI, to encourage the development and use of AI along two axes: trust and excellence (European Commission, 2021). These axes refer to the joint objectives of stimulating innovation and uptake of AI while at the same time ensuring that AI is applied responsibly. This regulation proposes a risk-based approach to AI, with applications ranging from having an unacceptable risk and therefore they should be banned, to high-risk applications that will become subject to, and other risk categories that may be subject to sectoral self-regulation. This way the AI regulation will contribute to legal certainty, but also re-enforce the European focus on human centric AI, and AI that respects fundamental rights. While the focus of the draft AI regulation is very much on taking into account the broader societal impact of AI, other pieces of regulation, like the GDPR, the European Consumer Law Aquis, or the pending rules about AI product liability focus on safeguarding the interests and rights of individual users.

Finally, law has an important role in creating the conditions for fair competition and choice, including competition for developing AI-driven products and solutions that are more responsive to the needs of users and society, and respectful of fundamental rights and public values. An important topic in this context is the 'winner takes it all' dynamics in technology markets, and the dominant role of a few (non-European) technology companies. With regulatory initiatives such as the Digital Services Act, the Digital Markets Act but also the Data Governance Act and the initiatives to facilitate data sharing and the creation of 'common European Data Spaces" law can have an important role in democratising technology adoption, competition for values but also tackling structural dependencies that can ultimately be detrimental for societies and the public values they cherish.

Societal

The emphasis on the Societal in the ELSA terminology necessitates a perspective on society and how it works. Evidently, this cannot be a neutral, uncontested perspective as many debates in sociology, political science, philosophy and a range of other disciplines tell us. A critical understanding of society will emphasize inequality and power structures and is likely to understand AI-driven innovation as a new instrument to maintain the hegemonic status quo (cf. Zubov, 2019). A pluralist view, on the other hand, understands society as an ongoing interaction between groups and institutes whose different interests will compete and collide in varying ways. In such a perspective AI does not work for one or the other societal interest, but is seen as a relatively neutral resource that can contribute to many different goals. Evidently, this is the perspective that is implicit in the AiNed framework for the ELSA Labs. An analysis of the many different groups, interests and resources engaged in technological innovation and production would be captured in the theory of social shaping, i.e. society and technology mutually shape each other; to which Actor Network Theory (ANT) offers a specific elaboration that identifies 'technology' and its affordances as an actor in the network as well.

From a critical perspective such pluralism forgets that not all actors have equal access and agency in the network. This is indeed one of the recurring critiques on ANT, that it ignores inequalities and different relations of power. However, social shaping theory and ANT alike, also assume that society is in a relatively reasonable and peaceful state and that its contending groups deploy similar kinds of knowledge, discourse and dialogue. In the current polarized climate in the Netherlands and elsewhere, this is an untenable assumption. Aside from the numerous social issues that divide groups, new technologies increasingly evokes radical and emotional opposition as the many protests against 5G, drones, wind turbines, vaccines and AI itself demonstrate.

Including 'the social' in the further design, development and deployment of AI thus necessitates the acknowledgment and accommodation of conflict and contestation, rather than working from an assumption of progressing pluralism. Recently, designers have worked towards methods that acknowledge antagonism and opposition, especially in relation to AI and algorithmic decision making. The notion they have introduced is that of 'contestability-by-design', i.e. designing systems that themselves have the capability of responding to opposition and that can make room for the influence of multiple stakeholders, in order to become fundamentally contestable (from Alfrink et al., 2020).

Capturing both human value and public value

Technical solutions for responsible AI or human-in-the-loop AI systems as well as legal and ethical approaches focus on maintaining the autonomy of individuals as end users of AI systems. While data science or computer science research often focuses on the effectiveness and usability of the AI system, in legal frameworks the human rights of citizens are central, such as privacy or non-discrimination.[6] To also include the mutual shaping of AI and social impacts, responsible AI development would benefit from taking a 'public value' perspective next to a 'human value' perspective. This may also be helpful for developers since the language of public values is often closer to home for public officials and citizens than the more abstract notion of ethics (Hayes, Poel & Steen, 2020). In this context it is important to take a broader perspective and define AI-driven systems broadly, taking into account not only possibilities for value-sensitive technology design, but also ways of implementing the technology into societal contexts and institutions, which are in themselves complex.

However, a public value perspective is not without its challenges. It assumes a societal discussion on what defines 'public good' within a great number of different public values that can be distinguished (e.g. Jørgenson & Bozeman (2007) identified 72 public values). Public value creation can therefore better be seen as a continuous process of determining what constitutes public value (Moore, 1995). In the context of the development and application of responsible AI, this will not only give rise to the question of how to perform this continuous assessment and weighing of values. It may also bring into scope societal tensions, including questions on power, democratic control, and governance of data and AI.

While the notion of ***human centric AI*** usually focuses on approaches of AI development towards individuals and end users, mainly from a consumer standpoint, the notion of ***AI for public good***, which focuses on public value is geared towards civic goals. The ELSA Lab approach would benefit from combining both human value and public value.

6. General Data Protection Regulation EUR-Lex - 32016R0679 - EN - EUR-Lex (europa.eu); the principle of non-discrimination EUR-Lex - dh0001 - EN - EUR-Lex (europa.eu).

4. CO-CREATION AND IMPACT

Co-creation and the n-tuple helix

The inclusion of public value next to human value into AI-systems, thus, brings into scope processes of co-creation. Co-creation in the widest sense includes every act of collective creativity (Sanders & Stappers, 2008). It is usually considered as a way for different actors to jointly create value (Osborne et al., 2016). Different reasons may exist for this joint value creation, for example the necessity for different actors or expertise to collaborate on complex issues, or to ensure that all different stakeholders are able to have a say in the outcome of a process (Bruns Alonso et al., 2020). In innovation processes, co-creation often takes place within a the 'n-tuple helix' setting, referring to the different institutions taking part in the co-creation process, including representatives from government, businesses, knowledge partners and societal partners, including citizens (cf. Galvao et al., 2019).

While co-creation has a longer history, its application to digital developments such as data science and AI is fairly recent. Within the field of AI development, co-creation and stakeholder engagement often seems to not have moved beyond classical usability studies such as human-computer interfacing (HCI) and A/B-testing (Gallo, 2017) with consumers as end-users, with some exceptions using ethics as a guiding force for AI development (cf. Verbeek & Tijink, 2019). The question is thus to move beyond this and use methodologies that somehow anchor public values and societal goals in the development process of AI systems. While the ethical and legal basis of responsible AI have yielded a strong basis of human rights and ethics frameworks, a question is how these shape the co-development and co-creation of technology.

'By-design' approaches guarding fundamental rights...

A prominent group of methodologies are 'by-design' strategies that may be used to connect human rights with engineering practice (Umbrello, 2019). Arguably the best know proponent of this group of methodologies is 'value-sensitive design' identifying values in the development process as well as for whom values may be added (Friedman, Kahn & Borning, 2006). They have developed elaborate frameworks, for example focusing on privacy ('privacy-by-design') (Spiekermann, 2012) and non-discrimination ('non-discrimination-by-design') (Sloot et al, 2021). These methodologies also include the aforementioned topic of 'contestability-by-design' (cf. Lyons, Velloso & Miller, 2021), which opens up for tensions and contradictions in requirements to specific AI technologies or algorithms. Transparency of AI-systems is a sine qua non for such contestation, and is also a fundamental element of the GDPR. However, in this regulation it is often arranged post hoc rather than integrated in the design as the 'by-design' approaches necessitate.

Therefore, while ethical and legal frameworks may be seen as a set of requirements that should be met, 'by-design' methodologies aim to entangle these human and public values with the technology from the very first to the very last stages of design, development, production, implementation and usage of AI, in ongoing iterative cycles. This will not happen by simply teaching data scientists about ethics or social relations, nor by undertaking ethical sessions with AI developers involving philosophers or legal experts. Other ways of 'making AI' are necessary which engage multiple stakeholders through all iterations.

… deployed through key enabling methodologies allowing for in-situ co-design

Whereas the focus of developing AI systems responsibly is often on the design and development phase of the technology, theory on social shaping of technology holds that this process continues after implementation and that it should accommodate contestation. This means that merely providing ethical and legal frameworks, nor guiding the design process based on ethics or fundamental rights may be sufficient, but, in addition, an embedded or 'in-situ' way of working is deemed necessary, involving relevant stakeholders, including citizens, and develop governance that is fitting. This is also in line with the rapidly developing nature of AI. These different elements point towards setting up Lab initiatives. Examples of Lab initiatives include AMS' Social Urban Data Labs[7] or TNO's Policy Lab[8] that attempt to bring on board all n-tuple helix partners.

However, especially involving citizens in such ways of working is a challenging undertaking (Borning et al., 2007) and more specifically for AI development (Black et al., 2020). An example where this has taken place is the citizen consensus council set up by the Danish Board of Technology.[9] Yet, as impressive, time consuming and potentially rewarding these initiatives to include citizens are, they typically suffer from self-selection bias and the need for a basic level of expertise. Self-selection bias means that only those citizens who have interest and the time and means to engage in a topic (Yang et al., 2019) while often the non-users are left out of such initiatives (Wyatt, 2003). The issue with expertise is that in order for meaningful debate to take place on a complex issue such as AI, basic knowledge is needed, which likely not all citizens have. Citizens' involvement in such a Lab way of working may, therefore, require specific attention.

7. AMS Institute - Social Urban Data Lab (ams-institute.org).
8. Policy Lab: developing data-driven policies | TNO.
9. Danish citizen technology panels (co-intelligence.org).

5. CONCLUSION: ELSA LABS FOR HUMAN CENTRIC AI

From the historical development of the notion of ELSA Labs, including ethical, legal, and societal aspects, it becomes clear that human centric AI should combine both human and public value. Furthermore, facilitating continuous learning, rather than assuming linear growth and impact, requires not entangling fundamental rights but also public values 'by-design', for example by allowing for 'contestability-by-design'. This continuous shaping of AI systems and society and citizens' involvement in technology development, such as in designing AI systems, may, therefore, require a more iterative and less linear 'post development' way of working. Therefore, ELSA Labs for human centric AI should deploy an 'in-situ' or Lab way of working together with quadruple helix partners, enabling the dynamic learning process that is essential for a rapidly developing technology such as AI. Therefore, the various principles laid out in this position paper, such as 'human centric AI', 'AI for Public Good', 'co-creation', the 'n-tuple helix', 'by-design approaches', including 'contestability' are made operational in several fundamental elements for ELSA Labs:

- In the mission that asks for articulation with sustainable development goals, and societal challenges ('public good') in addition to the development of human centric AI-systems.
- Through the requirement of using key enabling methodologies from design thinking ('co-creation' and 'by-design approaches').
- Through the requirement of working with 'quadruple helix' consortia (contestability and participatory design).

More formally, these fundamental elements have been made operational in seven requirements for ELSA Labs (NL AIC, 2021).

1. ELSA Labs address societally relevant issues which are aimed at sustainable prosperity in the broadest sense, examples may come from the 17 UN Sustainable Development Goals.
2. ELSA Labs collect validated and documented insights in a multi-stakeholder context.
3. Solutions are developed applying design thinking methods applying improvement cycles in real-life settings.
4. Insights and solutions are generated with data heavy and algorithm savvy techniques and methods.
5. In an ELSA Lab all four innovation helix dimensions assume equal responsibility for the development and coordination of the portfolio of activities carried out.
6. ELSA Labs apply a communication policy to share the insights and solutions with stakeholders and society at large.
7. ELSA Labs take a responsibility to scale-up solutions in order to impact society.

6. REFERENCES

Aalst, W.M.P. van der, Bichler, M. & Heinzl, A. (2017). Responsible Data Science. ***Business & Information Systems Engineering***, 59(5), 311–313.

AiNed programme, The Netherlands AI Coalition (NL AIC), AiNed programme | Nederlandse AI Coalitie (nlaic.com).

Alfrink, K., Turèl, T., Keller, I., Doorn, N. & Kortuem, G. (2020). Contestable City Algorithms, Contestable City Algorithms (participatoryml.github.io).

Almada, M. (2019, June). Human intervention in automated decision-making: Toward the construction of contestable systems. In: Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law, 2-11.

Black, E., Williams, J., Madaio, M.A. & Donti, P.L. (2020). A call for universities to develop requirements for community engagement in ai research. In: The Fair and Responsible AI Workshop at the 2020 CHI Conference onHuman Factors in Computing Systems.

Borning, A., Friedman, B., Davis, J. L., Gill, B. T., Kahn Jr, P. H., Kriplean, T. & Lin, P. (2009). Public participation and value advocacy in information design and sharing: Laying the foundations in advance of wide-scale public deployment. ***Information Polity***, 14(1,2), 61-74.

Bruns Alonso, M. et al. (2020, June). Sleutelmethodologieën (KEM's) voor missiegedreven innovatie Onderzoeksagenda - Onderdeel van de KIA Sleuteltechnologieën 2020-2023 Sleutelmethodologieën (KEM's) voor missiegedreven innovatie - KEM's voor missiegedreven innovatie (clicknl.nl).

Dignum, V. (2019). ***Responsible Artificial Intelligence. How to Develop and Use AI in a Responsible Way***. Cham, Switzerland: Springer.

European Commission High-Level Expert Group on Artificial Intelligence (2019). Ethics guidelines for trustworthy AI, Ethics guidelines for trustworthy AI | Shaping Europe's digital future (europa.eu).

ERF (2017). Workshop on ELSA in health care robotics. European Robotics Forum, ERF Workshop on ELSA in Healthcare Robotics (brl.ac.uk).

Forsberg, E.M. (2014). Institutionalising ELSA in the moment of breakdown? ***Life Sciences, Society and Policy***, 10(1), 1-16.

Friedman, B., Kahn jr., P.H. & Borning, A. (2006). Value Sensitive Design and Information Systems. In: Zhang, P. & Galletta, D. (eds.) ***Human-Computer Interaction and Management Information Systems: Foundations Advances in Management Information Systems, Volume 5 (Advances in Management Information Systems)***, Armonk, NY: M.E. Sharpe, 348-372.

Gallo, A. (2017). A Refresher on A/B Testing. ***Harvard Business Review***, A Refresher on A/B Testing (hbr.org).

Galvao, A., Mascarenhas, C., Marques, C., Ferreira, J. & Ratten, V. (2019). Triple helix and its evolution: a systematic literature review. ***Journal of Science and Technology Policy Management*** 10(3), 812–833.

Hayes, P., Poel, I. van de & Steen, M. (2020). Algorithms and values in justice and security. ***AI & SOCIETY***, 35(3), 533-555.

Jørgensen T.B. & Bozeman, B. (2007). Public Values: an Inventory. ***Administration and Society***, 39(3), 354-381.

Latour, B. (2015) ***Reassembling the social: an introduction to actor-network-theory***. Oxford New York: Oxford University Press.

Lyons, H., Velloso, E. & Miller, T. (2021, April). Conceptualising Contestability: Perspectives on Contesting Algorithmic Decisions. In ***Proceedings of the ACM Human-Computer Interaction*** 5(106).

Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S. & Floridi, L. (2016). The Ethics of Algorithms: Mapping the Debate. ***Big Data & Society*** 3(2), 1-21.

Moore, M.H. (1995). ***Creating Public Value: Strategic Management in Government***. Cambridge, MA, US: Harvard University Press.

NL AIC (2021, May). Achtergrondinformatie ELSA Labs, Informatie ELSA Labs mei 2021 (nlaic.com).

Orlikowski, W.J. (1992) The Duality of Technology: Rethinking the Conept of Technology in Organizations. ***Organization Science***, 3(3), 398-427.

Osborne, S., Radnor, Z. & Strokosch, K. (2016) Co-Production and the Co-Creation of Value in Public Services: A suitable case for treatment? ***Public Management Review***, 18(5), 639-653.

PADR (2019). Preparatory Action on Defence Research (PADR) Programme Guidance on How to complete your self-assessment on "ethics, legal and societal aspects (ELSA)", pa-guide-elsa-sa-padr_en.pdf (europa.eu).

Rogers, E.M. (1962). ***Diffusion of innovations***. Simon and Schuster.

Sanders, E. & Stappers, PJ (2008). Co-creation and the new landscapes of design. ***Co-design***, 4(1), 5-18.

Schillmeier, M. (2015). What ELSA/I makes big and small in Nanotechnology Research. In: ***Absence in Science, Security and Policy***. London: Palgrave Macmillan, 55-77.

Sloot, B. van der, et al. (2021). 'non-discrimination by design' Handreiking non-discriminatie by design | Rapport | Rijksoverheid.nl.

Spiekermann, S. (2012). The challenges of privacy by design. ***Communications of the ACM***, 55(7), 38-40.

Umbrello, S. (2019). Beneficial artificial intelligence coordination by means of a value sensitive design approach. ***Big Data and Cognitive Computing***, 3(1), 5.

Veenstra, A.F. van & Timan et al. (2021). AI: In search of the human dimension. Involve citizens and experiment responsibly. TNO Whitepaper, http://publications.tno.nl/publication/34638227/KTvx9Y/veenstra-2021-ai.pdf.

Verbeek, P.P. & Tijink, D. (2019). Aanpak Begeleidingsethiek. Den Haag: ECP, Aanpak begeleidingsethiek binnen de NL AIC - ECP | Platform voor de InformatieSamenleving.

Wyatt, S.M. (2003). Non-users also matter: The construction of users and non-users of the Internet. ***Now users matter: The co-construction of users and technology***, 67-79.

Xin, D., Ma, L., Liu, J., Macke, S., Song, S. & Parameswaran, A. (2018). Accelerating Human-in-the-loop Machine Learning: Challenges and Opportunities. In: ***DEEM'18: Proceedings of the Second Workshop on Data Management for End-to-End Machine Learning***, 1-4.

Yang, K. C., Varol, O., Davis, C. A., Ferrara, E., Flammini, A. & Menczer, F. (2019). Arming the public with artificial intelligence to counter social bots. ***Human Behavior and Emerging Technologies***, 1(1), 48-61.

Zubov, S. (2019). ***The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power***. London: Profile Books.

Zwart, H. & Nelis, A. (2009). What is ELSA genomics?. ***EMBO reports***, 10(6), 540-544.

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

Editors:
Anne Fleur van Veenstra, Liesbet van Zoonen, Natali Helberger

Contact:
E-mail — communicatie@nlaic.com
Website — nlaic.com

July 2021