Kenniscentrum
Data & Maatschappij

VISION PAPER

Kenniscentrum
Data & Maatschappij

Knowledge Centre
Data & Society

DECEMBER 2019

How to cite this report:

Ballon, P., Bénichou, B., Bertels, N., De Bruyne, J., Duysburgh, P., Fierens, M., Franck, G., Gils, T., Heyman, R., Laenens, W., Rossello, S., Valcke, P., Wauters, E. (2019). Vision Paper Kenniscentrum Data & Maatschappij. Kenniscentrum Data & Maatschappij, Brussel, België. [authors in alphabetical order]

www.data-en-maatschappij.ai

Introduction

This is the vision paper of the Kenniscentrum Data & Maatschappij (KDM)/ Knowledge Centre Data & Society (CDS)[1] on the social, legal and ethical aspects of artificial intelligence (AI)[2]. It outlines the role of the Centre, a proposed roadmap and themes it will work on. The vision paper was circulated among stakeholders and EWI (Flemish Department on Economy, Science and Innovation) through workshops and presentations. For each topic, we identified the challenges, discussed the envisioned steps to address the challenges and made a non-exhaustive list of existing initiatives that might be interesting to consider.

The vision paper is structured as follows:

- First, the method to arrive at this vision paper is elaborated.
- Second, the role of the Centre is clarified.
- Next, this paper is divided in two large parts. One on structural initiatives we need in Flanders and the other on themes the Centre will work on.
    - Part one is presented as a roadmap of proposed initiatives. It starts with the organisation of the Centre as a Flemish institutional panel, then its role in Digital Innovation Hubs (DIHs) is explained and lastly the Centre's possible contribution to regulatory sandboxing is proposed.
    - Next to these structural initiatives, the Centre will work on the following themes to pro-actively aid its stakeholders in part 2.
        - AI assessment tools and certification
        - Chatbots in Flanders
        - AI systems and GDPR compliance
        - Data maturity and governance

[1] Hereinafter referred to as the "Centre".

[2] For the purpose of this vision paper, we understand "AI" as defined by the High-Level Expert Group on Artificial Intelligence (AI HLEG) of the European Commission. This Expert Group defined AI as follows:

"Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems)."

More information in relation to this definition can be found here: https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top.

Preparations towards this vision paper

In preparation for this vision paper, the Centre contacted different stakeholders on three moments. Two sessions were interactive workshops hosted by the Centre, one is a presentation issued by 'Werkgroep implementatie en opleiding actieplannen AI en CS'.

- 05/11/2019: The first version was presented to a group of experts on AI on the 5th of November 2019 for input and feedback. The workshop was an ideation workshop where participants were invited to discuss what they needed with regard to the ten points raised in this document. For this workshop we invited organisations that represented our target audiences.
- 13/11/2019: The second version was presented to 'Werkgroep implementatie en opleiding actieplannen AI en CS'.
- 15/11/2019: A bottom-up workshop was organised with VVSG members where we started from concrete AI applications. Members identified ethical, legal or societal challenges for their local governments and citizens.

Table of contents

1. Role of the Centre Data & Society

The Flemish government has established the Centre Data & Society on the 22nd of March 2019. It focuses on the interplay between data, artificial intelligence and society. The Centre will enable socially responsible, ethical and legally appropriate implementations of AI in Flanders. It comprises three existing research Centres: imec-SMIT (Vrije Universiteit Brussel), imec-MICT (Ghent University) and the Centre for IT & IP Law (KU Leuven). The Flemish Department on Economy, Science and Innovation funds the initiative.

The Centre aims to enable Flemish companies, policymakers, regulators and citizens to achieve the greatest social and/or economic benefits of AI. For this it will bring together diverse stakeholder representative organisations and facilitate the creation of tools, advice and recommendations. The Centre will support umbrella organisations in the field by offering them state of the art knowledge, guidelines or teach the teacher programs, so that knowledge trickles down via their existing channels in a cooperative way.

Challenges:

- Organisations currently working on AI are unfamiliar with (the role and tasks of) the Centre.
- How can the general public be informed of AI and its societal, legal and ethical implications?
- Flemish stakeholders need a dedicated partner that focuses on social, ethical and legal aspects of data-driven applications and AI.
- A variety of international and European policy initiatives has been launched in the last few years. Regulatory developments can, therefore, be expected on those different levels, while stakeholders will expect a certain degree of uniformity or harmonisation in the results of those initiatives. What role can the Centre play in this regard?

General tasks of the Centre:

- Develop extensive expertise on legal, societal and ethical aspects of AI, thus becoming a 'regional' High Level Expert Group.
- Follow up on state-of-the-art international and national research on legal, social and ethical aspects related to AI.
- Share best practices and knowledge on AI-related topics by means of issuing short accessible and low-threshold communication tools such as 'snapshot' reports, (vision) papers or fact sheets. The output should be tailored to the needs of policymakers, industry actors, civil society, the general public and press.
- Provide or facilitate ad hoc advice on social, ethical and legal data and AI challenges on demand

via umbrella organisations.
- Set up networking opportunities such as lunch lectures, conferences and seminars that focus on AI technology and its societal, legal and ethical aspects.

- **Towards industry, the Centre will:**
  - provide the support of employers' associations and their members in the uptake of innovations related to data and AI technology with a particular focus on social, ethical and legal implications.
- **Towards policymakers and regulators, the Centre will:**
  - develop strategies to assess the impact of AI in Flanders in order to optimise and increase the representativity of statistics on AI activities in Flanders.
  - monitor experiences, attitudes and needs of the general public and industry towards AI.
  - provide policy recommendations on legal and ethical aspects of AI.
  - provide input on aspects related to AI whenever needed.
- **Towards the general public, the Centre will:**
  - increase awareness of AI and its societal, legal and ethical implications by benchmarking national and international initiatives on AI.
  - advise on the public outreach of AI by helping to determine the target groups and communication strategies and co-develop the basic understanding of AI that the public may/should have.
  - function as a portal and information point on AI for the general public.

Original tasks:

- Identify relevant stakeholder representatives for the Centre including policymakers, regulators, industry partners, innovators and the general public. Establish an advisory board with representatives of all different involved stakeholders to continuously identify their needs (see action point 2).
- Address the more fundamental and general legal, ethical and societal challenges caused by AI in a structural manner (see action points 2-4).
- Establish a learning environment for individual companies and organisations to eliminate false premises and erroneous suppositions about AI and data-driven applications.
- Act as coordinator on AI initiatives within the institutional structure of Flanders and aid in the division of ethical and legal competences thereby seeking for cooperation possibilities between all levels and actors (see action point 2).

Expected results:

- The Centre will become the central node for data and AI knowledge related to social, ethical and legal challenges and their possible solutions for our stakeholder representative organisations.
- The output of the Centre is integrated and distributed by different stakeholder organisations as part of their communication to their target groups.

- Advice of the Centre is used for the creation of policies, regulation and self-regulation to take social, ethical and legal challenges into account.

Existing initiatives to consider:

Name	URL	Description
Artificial Intelligence Impact Assessment	https://ecp.nl/publicatie/artificial-intelligence-impact-assessment-english-version/	An essential question for organisations relates to the legal and ethical consequences when deciding to use AI. The Artificial Intelligence Impact Assessment (AIIA) helps to answer this question.
Plattform Lernende Systeme ("Learning Systems Platform) – Germany's Platform for Artificial Intelligence	https://www.plattform-lernende-systeme.de/ki-landkarte.html	The aim of this initiative is to show how AI is transforming the economy and everyday life. The Learning Systems Platform bundles applications and development projects in which AI technologies are used in Germany across all industries, applications and company sizes. The initiative is supported by the Federal Ministry of Education and Research in cooperation with the Federal Ministry for Economic Affairs and Energy.
Cartographie des entreprises actives dans l'IA en Wallonie	https://www.digitalwallonia.be/fr/publications/cartographie-ia	As part of the DigitalWallonia4.ai strategy, and more specifically its part on awareness, a mapping of the relevant AI activities and initiatives in Wallonia is being made through the Digital Wallonia platform.
Elements of AI	https://www.elementsofai.com/	The Elements of AI is a series of free online courses created by Reaktor and the University of Helsinki. The aim of the courses is to encourage people to learn what AI is, what can (and cannot) be done with it and how to start creating AI methods.

CDEI Snapshots	https://www.gov.uk/government/publications/cdei-publishes-its-first-series-of-three-snapshot-papers-ethical-issues-in-ai	The Centre for Data Ethics and Innovation (CDEI) Snapshots are briefing papers designed to build understanding on ethical and governance issues related to the development and deployment of AI. They aim to separate facts from fiction, clarify what is known and yet unknown about a particular issue, and outline possible avenues of action by government and industry in the near future.

Part 1: Enabling structural initiatives for social, ethical and legal aspects in data and AI

In this section we focus on initiatives that provide structural support for social, ethical and legal values in data and AI. These initiatives are meant to bring together different actors, influence legislation and change how Flanders deals with social, ethical and legal challenges in data and AI.

2. Flemish Institutional Panel on AI

AI systems are evolving fast and their presence in our everyday lives is significantly increasing. Nevertheless, little seems to be known about these systems' legal, ethical and social impact. Moreover, although concepts like lawful and ethical AI are governing the AI policymaking debate, it is still unclear how these should be implemented in practice. As remarked by the High Level Expert Group on AI (AI HLEG), there seems to be an existing governance gap on the matter. The AI HLEG has, therefore, recently recommended the establishment of a new institutional structure for Trustworthy AI. This institution would be tasked, among others, with the following:

- Contributing to the development of EU's framework and policy for Trustworthy AI.
- Provide guidance to stakeholders on how their applications can comply with the law and the requirements for Trustworthy AI.
- Host and update a repository of best practices that are deemed compliant with applicable law and Trustworthy AI principles and requirements.
- Assist public and private sector in the application of a risk-based approach including providing opinions for the assessment of the intensity, probability and unacceptability of AI created risks.
- Contribute to preparing for socio-economic changes brought about by AI.

Regardless of this European requirement, an institutional panel on AI is required to advice on the different technical AI areas that exist, specific sectors and different legal areas that exist.

> We propose this point as a first step because it is part of the proposed mission for the Centre and it will help in initiating the other steps which require a connection with the EU and relevant national and regional policymakers and administrations.

Challenges:

- Although this recommendation is not binding and it is (currently) unknown whether such an institution will be created at the EU level, Flanders should closely watch developments in this direction. In doing so, it should investigate whether and how it could act as the regional contact point for such institution, translating and/or enforcing relevant policies and guidelines developed at EU level to the Flemish framework, with respect to the areas in which it has the competence to legislate.
- If such an institution would not be created at the EU level, the Centre could serve as an institutional panel nonetheless. Given the complex Belgian state structure and the transnational character of AI, cooperation and coordination with the other regions and the federal level is important. There is a

risk that fragmentation may lead to a lack of overview and participation of citizens, businesses and departments. Fragmentation may result in different and even inconsistent policy decisions.
- The Centre lacks specific domain or sectoral knowledge required to ensure that advice, tools or recommendations are relevant for specific areas.

Envisioned steps to address the challenges:

- Map and invite representatives of different regional or international bodies to establish our presence as Flemish institutional panel on AI.
- Identify relevant sectors.
- Follow up legal and policy initiatives on AI at the EU, federal and regional level.
- Consult policymakers at the supranational, federal and EU regional level to investigate the feasibility and desirability of a (European) institutional panel on AI with national/regional contact points.
- Map division on competences in the Belgian state structure.
- Map the different actors and departments working on AI at the regional and the federal levels and especially the actions they have taken or will take.
- Propose cooperation in the fields where it might be required and useful.
- Create templates for future actions and gather data from stakeholders.

Expected results:

- The Centre becomes a facilitator of AI initiatives tied to social, ethical and legal values in Flanders and Belgium, taking into account the institutional state structure and the division of competences. Cooperation possibilities between all levels and actors could be proposed thereby avoiding unnecessary similar initiatives and efforts. This can reduce the fragmentation of AI initiatives, policies and actors in Belgium.
- The Centre will contribute to the development of the AI HLEG framework and legal frameworks for Trustworthy AI.
- The Centre can become the institutional panel for Flanders and provide an overview of the relevant policy domains and different governments (regional and federal) affected by or working on AI.

Existing initiatives to consider:

Name	URL	Description
Policy and investment recommendations for trustworthy Artificial Intelligence	https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence	This document contains the policy and investment recommendations for Trustworthy AI drafted by the AI HLEG (see above and previous benchmark D.1.2)
Mandate for the International Panel on Artificial Intelligence	https://pm.gc.ca/en/news/backgrounders/2018/12/06/mandate-international-panel-artificial-intelligence	Canada and France established an International Panel on Artificial Intelligence aimed at becoming a global point of reference for understanding and sharing research results on AI issues and best practices, as well as convening international AI initiatives.
A comprehensive European industrial policy on artificial intelligence and robotics	http://www.europarl.europa.eu/doceo/document/TA-8-2019-0081_EN.html	In this resolution the European Parliament: • considered that a comprehensive Union system of registration of advanced robots should be introduced within the Union's internal market where relevant and necessary for specific categories of robots, • calls on the Commission to establish criteria for the classification of robots that would need to be registered, • in this context, calls on the Commission to investigate whether it would be desirable for the registration system and the register to be managed by a designated EU Agency for Robotics and Artificial Intelligence.

3. Innovation Hubs (on data and AI)

Digital Innovation Hubs (DIHs) can help to support the digitisation of the industry and the improvement of local innovation ecosystems. DIHs are one-stop shops that help companies to become more competitive by using several digital technologies in their business or production processes. DIHs provide access to technical expertise and experimentation (cf. "test before you invest"), support companies to find investments, give financial advice and skills development, and promote networking opportunities. DIHs are an important distributor of data and AI knowhow for the industry, for this reason we need to include them in our network to distribute Centre tools and guidelines.

Challenges:

- Many public and private efforts/initiatives to support innovation and digitisation co-exist. This can decrease clarity and lead to a waste of resources.
- DIHs run the risk of offering fragmented and highly different levels of support for social, ethical and legal challenges related to data and AI.
- Belgian DIH funding may be tied to their adherence and operationalisation of EU social, ethical and legal values.
- Ensure the proper exchange of valuable information between all the involved parties.

We propose this as a second step because DIHs require the Centre's institutional panel on AI. As a panel we will have a working relation with the European level. DIHs also require tools and guidelines that are tailored to specific DIH contexts (we elaborate on this in action point 5).

Envisioned steps to address the challenges:

- **Establishment of DIHs:**
  - Focus on a trustworthy deployment of AI technologies in the DIHs to enable compliance with the Ethical Guidelines issued by the AI HLEG. The Centre could support initiatives taken by DIHs to test and implement guidelines in Flemish industry.
  - Support the creation of sectoral hubs focusing on important (economic) sectors in Flanders or Belgium, if necessary and relevant. AI should be embedded in those hubs, just as high-performance computing (HPC) and cybersecurity. These newly created hubs could take into account or rely on currently existing clusters and initiatives.
  - Create a central hub that gathers all DIHs in Flanders to create synergies in terms of inter alia best practices to operationalise social, ethical and legal requirements. This central hub can also act as a single point of contact and redirect inquiries to the competent sectoral DIHs (cf. model in the Netherlands).

- Define necessary (regulatory) criteria and co-create requirements that DIHs have to comply with. Cooperation between for instance VLAIO, other relevant stakeholders and the Centre can be useful to define those criteria.
- Public funding resources might need to be combined with other funding resources (e.g. remuneration for companies providing innovative services). Here, the Centre could support Flemish research and innovation funders in defining criteria that include social, ethical and legal requirements that companies need to comply with to receive funding.
- The Centre is involved as a stakeholder in defining the (general) digital strategy for DIHs and give recommendations accordingly.

- **Coordination, communication and exchange of information**
  - Coordinate DIHs together with the responsible and competent partners at the Federal levels as well as regional levels (Wallonia).
  - Monitor cross-border initiatives and rely on these to develop testing-tools and provide our local industry with highly competitive services.
  - Cooperate with DIHs, support them with regard to social, ethical and legal issues and challenges, be involved as a stakeholder in determining the digital strategy for DIHs and provide tools to hubs allowing companies to test them.

Expected results:

- Identify relevant DIHs and facilitating partners to integrate DIHs.
- Co-create a framework to integrate social, ethical and legal requirements in DIHs in a uniform way for all Flemish and/or Belgian DIHs.
- Use the DIH network as a means to create sectoral panels to advice on the implementation of social, ethical and legal policies, tools and guidelines.
- Publish guidelines, recommendations, overview best practices that can be used by stakeholders.

Existing initiatives to consider:

Name	URL	Description
AI Experience Center	https://ai.vub.ac.be/ai-test-experience-center/	The AI Experience Center is a joint project of four research groups of the VUB: the Artificial Intelligence Lab, Brubotics, SMIT and ETRO. It is a high-tech test, demonstration and meeting environment in which companies, entrepreneurs, researchers, policy-makers and the general public can experiment with AI and collaborate to develop and produce technological solutions.
Smart Digital Farming, the orchestrator of SmartAgriHubs in Flanders	https://www.smartdigitalfarming.be/	Smart Digital Farming is the orchestrator of SmartAgriHubs in Flanders. SmartAgriHubs connects the dots of agri-tech innovation to catalyse the digital transformation of the European agri-food sector. The project aims to realise this by fostering a self-sustaining agricultural innovation ecosystem dedicated to excellence, sustainability and success.
Flanders' digital innovation ecosystem	https://www.flandersinvestmentandtrade.com/invest/en/sectors/digital-society/digital-society-ecosystem	Overview of the digital innovation ecosystem in Flanders (know-how and research, cluster and key organisations, business incubators).
Flanders' chemical ecosystem	https://www.flandersinvestmentandtrade.com/invest/en/sectors/chemicals/chemicals-ecosystem	Overview of the chemical ecosystem in Flanders (know-how and research; cluster and key organisations; business incubators).
Fintech in Flanders	https://www.flandersinvestmentandtrade.com/invest/en/sectors/digital-society/fintech	Overview of the financial ecosystem in Flanders (know-how and research; cluster and key organisations; business incubators).

Flanders' automotive ecosystem	https://www.flandersinvestmentandtrade.com/invest/en/flanders'-automotive-ecosystem	Overview of the automotive ecosystem in Flanders (know-how and research; cluster and key organisations; business incubators).
Energy in Flanders	https://www.flandersinvestmentandtrade.com/invest/en/sectors/energy	Overview of the energy ecosystem in Flanders (know-how and research; cluster and key organisations; business incubators).
Food in Flanders	https://www.flandersinvestmentandtrade.com/invest/en/sectors/food-nutrition/food-nutrition-ecosystem	Overview of the food ecosystem in Flanders (know-how and research; cluster and key organisations; business incubators).
Flanders' innovation ecosystem for life sciences	https://www.flandersinvestmentandtrade.com/invest/en/flanders'-innovation-ecosystem-life-sciences	Overview of the innovation ecosystem in Flanders for life sciences and health (know-how and research; cluster and key organisations; business incubators).
Pan-European network of Digital Innovation Hubs (DIHs)	https://ec.europa.eu/digital-single-market/en/digital-innovation-hubs	Pan-European network of Digital Innovation Hubs (DIHs)

4. Regulatory Sandboxing

Regulatory sandboxes allow private enterprises to test new technological developments in a supervised experimental environment. Regulatory authorities such as Data Protection Authorities (DPAs) or other authorities relax specific legal requirements for the duration of a sandbox case. Authorities can also provide advice on mitigating risks and enabling compliance by design. This process saves time and allows for learning by doing because adopting legislation is a lengthy and time-consuming process. Regulatory sandboxes, by contrast, immediately show empirical results of a specific application. Sandboxes allow regulators to identify (legal) challenges and could speed up the legislative process.

This initiative requires the integration of DIHs, different governments and regulators. For that reason, this step follows after the other two action points of this first part.

Challenges:

- There is no safe or trusted (regulatory) sandboxing environment.
- An improper balance between commercial interest and the protection of several other rights and/or the general interest (e.g. transparency vs. trade secrets or commercially sensitive information) may exist.
- There is a risk of fragmentation as companies could establish sandboxes without necessarily informing all required authorities. The latter may lead to a lack of oversight.
- Limited room for participants in sandboxes may raise competition and other concerns (e.g. wasted investments if a company is not selected to participate in the hub, information might fall into hands of competitors, …).
- Division of competence in Belgium could influence the efficiency and efficacity of a regulatory sandbox considering that different regulators might need to be involved (cf. competences and structure in Belgium).

Envisioned steps to address the challenges:

- Focus on trustworthy deployment of AI technologies in the DIHs to enhance compliance with Ethical Guidelines issued by the AI HLEG.
- Determine conditions for participating in a regulatory sandbox (e.g. compliance with certain standards or certification). Specific industrial needs can be taken as departing point in this regard. This will allow to create sector-specific sandboxes.
- Support regulatory authorities to establish an entry test based on certain pre-defined criteria (e.g. benefits to consumers, innovative aspect, …), testing parameters or conditions, evaluation methods

and exit criteria. Inspiration can be sought in best practices in other countries (cf. functional comparative approach).

- Provide a clear definition of specific AI systems that are tested, thereby taking self-learning capabilities of those systems into account.
- Examine how the division of competence in Belgium could influence a specific regulatory sandbox, and especially evaluate what can be done to remedy potential institutional problems and challenges (cf. Flemish institutional panel).
- Determine which regulatory sandboxes already exist in Flanders, Belgium and abroad. This will be useful to decide in which sectors they can be created and which companies can be included.
- Ensure that different regulatory authorities, including but not restricted to DPAs, are explicitly authorised by law to oversee regulatory sandboxes.
- Ensure accountability and means of redress in case harm or any other adverse impact is caused.
- Make sure that public and private stakeholders are sufficiently involved and informed to avoid fragmentation.
- Examine whether other innovative tools and/or (adaptive) regulatory techniques are useful in addition to sandboxes when regulating AI.

Expected results:

- Identification of sandbox opportunities with DIHs and the other Flemish AI initiatives.
- Facilitate the elaboration of particular sandbox areas between innovators and regulators
- Be part or aid in the creation of an oversight body to address challenges for social, ethical and legal dimensions.

Existing initiatives to consider:

Name	URL	Description
More room for innovation in the financial sector – Market access, authorisations and supervision: Next steps AFM - DNB	https://www.dnb.nl/en/binaries/More-room-for-innovation-in-the-financial%20sector_tcm47-361364.pdf	Financial sector, December 2016 "De Nederlandsche Bank" (DNB) and the Netherlands Authority for the Financial Markets (Autoriteit Financiële Markten / AFM) supervise financial undertakings and pension funds in the Netherlands.

Financial Conduct Authority - Regulatory sandbox	https://www.fca.org.uk/publication/research/regulatory-sandbox.pdf	Financial sector, November 2015 The Financial Conduct Authority is the conduct regulator for 59,000 financial services firms and financial markets in the UK and the prudential regulator for over 18,000 of those firms
Finland's Age of Artificial Intelligence - Turning Finland into a leading country in the application of artificial intelligence - Objective and recommendations for measures	http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEM-rap_47_2017_verkkojulkaisu.pdf?sequenc%20e=1&isAllowed=y	December 2017 Regulatory sandbox to encourage data sharing in Finland by the Ministry of Economic Affairs and Employment of Finland
Malta Digital Innovation Authority Act	http://www.justiceservices.gov.mt/DownloadDocument.aspx?app=lom&itemid=12873&l=1	"AN ACT to provide for the establishment of an Authority to be known as the Malta Digital Innovation Authority, to support the development and implementation of the guiding principles described in this Act and to promote consistent principles for the development of visions, skills, and other qualities relating to technology innovation, including distributed or decentralised technology, and to exercise regulatory functions regarding innovative technology, arrangements and related services and to make provision with respect to matters ancillary thereto or connected therewith."

Part 2: Creating tools and guidelines for social, ethical and legal values in data and AI

In this section we focus on initiatives that provide support for social, ethical and legal values in data and AI on a thematical level. For this content we identified the current strengths of the Centre.

5. AI Assessment Tools and Certification

The Guidelines issued by AI HLEG contain several requirements to ensure that trustworthy AI is created. The key requirements of the AI HLEG trustworthy AI assessment are currently being piloted and other methods to evaluate data and AI exist as well. An important issue arises: how should it be determined that AI is indeed trustworthy or ready for society? Practical tools may be developed and tested to that end. There will thus be a need for common standards and especially certification mechanisms to ensure that AI is lawful, ethical and robust.

> The need to prove adherence to social, ethical and legal values is an important prerequisite to facilitate AI acceptance in Flanders. For this reason, this step is a first content wise priority.

Challenges:

- It is unclear how social, ethical and legal AI should be operationalised. As a result, it is difficult for innovators to show that they comply with ethics guidelines.
- There is a lack of clear processes and requirements to assess how companies may establish compliance with AI assessment criteria.
- Compliance is often seen as an administrative burden, so tools should decrease this burden and be framed as product differentiators. For other forms of compliance, ethics or sustainable development goals, industry efforts increase if these goals contribute to the goals of a company.
- How can certification be used as a regulatory tool to ensure and promote Trustworthy AI?

Envisioned steps to address the challenges:

- Bring together stakeholders to discuss how to operationalise and measure trustworthy AI (ongoing process, cf. previous event organised by AI4Belgium).
- Not every social, ethical and legal method will be beneficial to any type of AI application or for each step in the development process. It is important to prioritise what to do first.
- Examine which practical tools already exist to save time, efforts and costs in the long run.
- Translate each ethical guideline into clear assessment criteria.
- Rely on existing (academic) research to evaluate how certification can be used as a regulatory tool to promote Trustworthy AI.
- Evaluate to which extent the existing legal framework can already be relied upon to find necessary answers (cf. GDPR, Product Liability Directive).

Expected results:

- Publish the advantages and disadvantages of AI assessment methods applied to cases to illustrate their appropriateness and usefulness.
- Co-create tools with relevant stakeholder representatives for the most common or pressing AI applications.
- Report on the use of certification as regulatory tool to promote AI systems (i.e. benefits, challenges, …).

Existing initiatives to consider:

Name	URL	Description
Foundation for Responsible Robotics (FRR)	https://responsiblerobotics.org/	The Foundation for Responsible Robotics (FRR) aims to shape the future of responsible robotics design, development, use, regulation, and implementation. The Foundation for Responsible Robotics has partnered with Deloitte to create the FRR Quality Mark for Robotics and AI. The Quality Mark for Robotics and AI is a label on robotics products that indicates to consumers that this product has been assessed/evaluated by an independent external expert group on responsible robotics. The aim of the quality mark is to promote important features of products that will contribute to a better world, including sustainability, integrity, safety and security, and ethical design together with a consideration of societal impact.

The Ethics Certification Program for Autonomous and Intelligent Systems (ECPAIS) Developing metrics and processes towards the implementation of a certification methodology addressing transparency, accountability and algorithmic bias	https://standards.ieee.org/industry-connections/ecpais.html	The Institute of Electrical and Electronics Engineers launched the Ethics Certification Program for Autonomous and Intelligent Systems (ECPAIS). The goal of The Ethics Certification Program for Autonomous and Intelligent Systems (ECPAIS) is to create specifications for certification and marking processes that advance transparency, accountability and reduction in algorithmic bias in Autonomous and Intelligent Systems (A/IS).

6. Chatbots in Flanders

Chatbots can be seen as autonomous AI agents that use NLP[1] and scripts to automate communication tasks previously carried out by employees. Chatbots are currently being used by innovative municipalities[2] and companies to support their customer side with 24/7 customer support.

Not all municipalities and companies have the necessary knowledge to understand how chatbots work and what their benefits and disadvantages might be. The Centre will use chatbots as a case to apply its guidelines and tools for so that chatbot development can be a social, ethical and legal by design process.

This AI application is currently on the market and being used in such a way that a broad audience will interact with these in the near future. Therefore it is an ideal case to illustrate the methods of the Centre. For this a collaboration with VVSG has been set up and EWI already expressed interest in this project.

Challenges:

- There is no understanding of the appropriate and inappropriate application areas of a chatbot. For example, to what extent should local administrations automate the granting of social rights and benefits and how do we decrease/avoid exclusion?
- There are no guidelines on how to be transparent about human-assisted chatbots, and other social, ethical and legal obligations such as how personal data should be handled.
- It is unclear if the general audience understands how chatbots function.
- Standards and recommendations are necessary to guide chatbot applications towards beneficial directions for society.

Envisioned steps to address the challenges:

- Analyse different chatbot applications in private and public services and evaluate their benefits from the customer and service side.
- Create a tool to evaluate the appropriateness of using chatbots.

[1] Natural language processing.

[2] The topic of Chatbots in Natural language processing in Flemish Municipalities is investigated by VVSG. The latter organised a workshop on November 8th.

- Create standard clauses for procurement, data processing and transparency and a label or certification to show that proper steps were taken to evaluate the appropriateness of a chatbot instance.
- Develop an inclusion test or questions to evaluate if and how humans may be replaced or complemented by a chatbot.

Expected results:

- A whitepaper based on best practices for chatbots in Flanders with voices from chatbot developers, their clients and the intended audience of chatbots.
- A tool to evaluate the appropriateness of chatbots using social, ethical and legal values.
- A recommendation for standardisation of chatbots co-created with stakeholder representatives.

Existing initiatives to consider:

Name	URL	Description
Open Standaarden voor Linkende Organisaties (OSLO)	https://overheid.vlaanderen.be/producten-diensten/oslo2	Public services to citizens and entrepreneurs in Flanders are supported by various specialized applications from various software suppliers. The aim of 'Open Standaarden voor Linkende Organisaties' (OSLO) is to ensure greater coherence, and better understandability and findability of information and services.
Smart Flanders	https://smart.flanders.be/	On 1 January 2017, Minister Liesbeth Homans launched the Smart Flanders program. It is a support program that is executed by imec researchers. The Flemish government thereby supports several cities in their development towards smart cities.
Programma Innovatieve Overheids-opdrachten	http://innovatieveoverheid-sopdrachten.be/	PIO 2019

Lawren.io chatbot	https://www.computable.be/artikel/achtergrond/start-it/6631766/5679911/chatbot-van-lawrenio-digitaliseert-advocatuur.html	Chatbot by Lawren.io digitalises the profession of lawyers.
Clever chatbot	https://www.clever.be/	Clever already developed a chatbot that is used in Aalter.

7. AI Systems and GDPR Compliance

The main purposes of the GDPR are the following:

- To ensure anyone's right to protection of his or her personal data, by putting them in control over the processing of such data and imposing processing principles such as transparency, fairness, purpose limitation, accountability, data minimisation and lawfulness;
- To create an economic and social environment in which data subjects can trust that their personal data will only be processed in compliance with their expectations and thus do not fear to share their personal data, allowing such data to be used for the economic development and society as a whole;
- To ensure and facilitate the free movement of personal data within the European Economic Area.

The requirements and the way AI systems function, such as their need for large volumes of data and the "black box" principle, are often considered to be at odds with the GDPR, especially because the regulation appears to limit some specific AI functions.

> The thematical focus on GDPR will support regulatory sandboxes, AI assessments and certifications as the GDPR will be relevant if personal data is being processed in AI applications.

General GDPR challenges:

- Processing of personal data by AI systems, including decision-making based on personal data, cannot interfere with one's right to personal data protection.
- Misinterpretation or misapplication of the GDPR must be avoided as this may distort or stifle the innovation and development of AI systems and reduce the resulting societal added value. Fear for GDPR sanctions may withhold organisations to deploy AI systems when processing personal data.
- Implementing AI business cases without taking into account the GDPR requirements, may/will lead to a "technical debt", which will require organisations to invest (even more) in privacy compliance in a later stage. Resolving this "technical debt" will hinder the development process and thus hamper competitive advantages that may have been prevented at an earlier stage. Not resolving it in time may also lead to fines for GDPR infringement.
- The intent and the scope of the GDPR are not always clear in relation to AI. The applicable provisions lead to different interpretations. This creates legal uncertainty on how AI systems can be used in a GDPR-compliant way, and forms a hurdle to develop AI systems in the EU.
- The consequences of the use of AI-systems on a person's fundamental rights might be underexposed due to the strong focus on the often rather administrative/formalistic implementation of procedures and requirements under the GDPR. Such an administrative/formalistic approach regarding

GDPR compliancy might hinder the proper development of privacy awareness and privacy culture in organisations, making it difficult to apply the GDPR in new and complex situations.

- Many organisations are already struggling with GDPR compliancy. This will only become more problematic when AI systems are deployed as these will require an even more thorough and comprehensive approach to data protection.
- AI systems can be useful to help safeguard the right to data protection of individuals and to help organisations to be compliant with the GDPR.
- AI systems can have many (societal) benefits when it comes to processing of specific categories of personal data such as health data. Processing those categories of data, however, poses a higher risk to the concerned persons. This results in a higher protection of such personal data and additional restrictions on the processing hereof, making the deployment of AI systems in these fields extra burdensome.

Specific GDPR challenges:

- By setting forward principles such as data minimisation, accountability, transparency and explainability regarding the processing of personal data, the GDPR may appear to create an environment that prevents data-driven AI systems to be fully deployed ('flourish') when they need to process and re-process as much as possible personal data.
    - Explainability appears difficult to apply when decisions are made by AI systems. The requirement of explainability of actions generated by AI systems and the interrogability of such systems may, however, contribute to the development of Trustworthy AI. The application of these principles may indeed allow to discover potential biases in decision-making processes by AI systems that would otherwise remain undiscovered for a long(er) period. Therefore, it is important to emphasize the benefits that can be derived from explainability for the development of Trustworthy AI, outside of the strict GDPR scope.
    - Transparency and information obligations apply under the GDPR and are also closely related to the question of explainability. Personal data must be processed in a transparent manner and the data subjects must be transparently informed on how their personal data is processed and for which purposes.
- The GDPR imposes specific requirements for profiling and automated decision-making. These processes are often powered by AI systems.
- There are certain restrictions under the GDPR regarding the collection of personal data can that conflict with how AI systems should ideally have access to data. For instance,
    - Personal data must be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes;
    - Personal data must be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed.
- Data subjects also have a right to object against processing of their personal data, including profil-

ing, processed in the public interest or for the legitimate interest of the controller, except if the controller has an overriding interest. Data subjects can equally object against processing for marketing purposes. In these cases, the processing of such personal data must be ended.

- There are several requirements and limitations regarding automated individual decision-making including profiling.

Envisioned steps to address the challenges:

Identification steps

- Identify common misconceptions on the (relationship between) GDPR and AI.
- Identify hurdles experienced by the stakeholders when applying GDPR to AI systems.
- Identify possible frictions and tensions between data protection requirements and the optimal use of AI system, and provide guidelines on how to remedy them.

Information steps

- Inform the public about the benefits of a GDPR-compliant use of AI systems.
- The knowledge Centre should provide information and advice to the stakeholders on their potential liability under the GDPR framework. In other words, clarity should be brought on the risk of liability faced by companies under the GDPR.
- Inform companies and other relevant actors on the interpretation and application of relevant GDPR provisions and aspects for AI systems
- Further inform the companies and other relevant actors on how to:
    - translate GDPR obligations into practice, in particular when using AI systems;
    - anonymise personal data and allow the use of such data by AI systems without entering the scope of the GDPR;
    - reduce the need of extensive datasets;
    - comply with the transparency requirements in relation to "black box decisions", such as by developing or using explainable AI ("XAI") methods;
    - avoid using AI in ways that infringe individual's data protection rights.

Development steps

- Privacy by design must be promoted and become the standard in the development and use of AI. Effective adoption of privacy by design will avoid technical debt in a later stage.
- Regulatory sandboxes must be established to allow researchers and the industry to develop AI systems in close dialogue with the data protection authorities (See action point 4).
- Create informative and guiding tools that can be directly implemented by SMEs and increase actual privacy awareness, instead of the mere formalistic check-the-box attitude towards data protection. The Centre could help translating the existing conceptual framework on privacy in more practical

tools that can actually help developers of AI to comply with privacy requirements.

- Cooperate with existing sector organisations to develop and distribute these tools.
- Promote the development and use of AI systems to enhance and monitor GDPR compliance.
- Promote explainability and interrogability as means to have a better understanding of AI systems, to ensure the quality of AI-based decisions, to avoid biases and to comply with the GDPR.

Expected results:

- Publish and communicate the results obtained from the envisioned steps.
- Advise stakeholders on the outcome obtained from the envisioned steps.
- Establish and distribute sector-specific and best practices on how to apply data protection principles when using and developing AI systems. This may include providing and developing data protection impact assessments as an example.
- Organise workshops and trainings to stakeholders ('train-the-trainer') and the public.

Existing initiatives to consider:

Name	URL	Description
Declaration on ethics and data protection in artificial intelligence	https://edps.europa.eu/sites/edp/files/publication/icdppc-40th_ai-declaration_adopted_en_0.pdf	Declaration on ethics and data protection in artificial intelligence (2018 – EDPS)
ExplAIn	https://ico.org.uk/about-the-ico/research-and-reports/project-explain-interim-report/	Project ExplAIn (2019 – ICO)
Algorithms and artificial intelligence: CNIL's report on the ethical issues	https://www.cnil.fr/en/algorithms-and-artificial-intelligence-cnils-report-ethical-issues	Algorithms and artificial intelligence: CNIL's report on the ethical issues (2018 – CNIL)
Ethics Guidelines for Trustworthy AI	https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai	Ethics Guidelines for Trustworthy AI of 8 April 2019, by the AI High-Level Expert Group on Artificial Intelligence (2019 – HLEG AI)

Guidelines on Artificial Intelligence and data protection	https://www.coe.int/en/web/artificial-intelligence/-/new-guidelines-on-artificial-intelligence-and-data-protection	Guidelines on Artificial Intelligence and data protection (2019 - Council of Europe)
Big data, artificial intelligence, machine learning and data protection	https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf	Big data, artificial intelligence, machine learning and data protection (2017 – ICO)
Guide for GDPR implementation for SMEs	http://ict-kmo.be/sites/default/files/52/brochure_ict_kmo_2018.pdf	Guide for GDPR implementation for SMEs
An artificial intelligence impact assessment	https://ecp.nl/publicatie/artificial-intelligence-impact-assessment-english-version/	An artificial intelligence impact assessment (AIIA) created by ESP in the Netherlands
Tetra	https://distrinet.cs.kuleuven.be/research/projects/TETRA-GDPR https://msec.be/GDPR/	TETRA VLAIO project. This project supports innovation of the software development process in compliance with the new privacy regulations. The project targets increasing the competitivity of Flemish software developers, integrators and consultants that focus on applications in which personal data collection and processing is required. Partners in this project are CiTiP and Distrinet, both KU Leuven.
BOOST	https://www.law.kuleuven.be/citip/en/research/projects	BOOST project - Boost Belgian SMEs' awareness of and compliance with the GDPR. Partners in this project are the Belgian Data Protection Authority, imec-SMIT (VUB) and CiTiP (KU Leuven).

8. Data Maturity and Governance

Data is essential to most AI systems and collection or preparing data are one of the first steps in development. In order to examine how AI can be implemented in a social, ethical and legally acceptable way in an organisation or specific sector, we need to understand if the collected data is managed in a way that is sufficiently mature to enable the development and deployment of AI systems. The quality of the collected data also plays an important role, while (a certain level of) data standardisation is likely required to enable the relevant actors to share and use third-party data. For the purpose of this vision paper, we divide the identified data-related challenges into four categories: Data Regulation, Data Maturity, Data Quality and Data Standardisation.

This theme will build on cases where GDPR is not applicable, but another form of data management and governance will be needed. This theme is last because expertise will be sourced from outside the Centre.

Challenges:

- Data Regulation:
  - Different types of data are covered by different legal rules. This can be confusing for Flemish organisations and businesses and may have an impact on data quality, data maturity and data governance.
- Data Maturity:
  - Without **data governance** data will not be sufficiently mature, which will hamper the uptake and development of (trustworthy) AI.
  - Data maturity also refers to the **data management** on an organisation's level and relates to the following questions:
    - is data in a company stored in a standardised way?
    - is it clear where and how the data has been collected?
    - is the collected data still accurate?
- Data Quality:
  - AI development processes spend a considerable amount of time and resources on data cleaning and structuring. This decreases the time and resources that could be spent on actually training AI.
  - Poor data quality can lead to misleading and inadequate AI systems. This can result in biases and lead to inaccurate outcomes.
- Data Standardisation:
  - Data standardisation refers to the existence and usage of standards to exchange data in the

public or private sector.

- Currently, a variety of initiatives to standardise data are ongoing. For instance, Belgian governmental authorities on all levels have been using EU standards. It is therefore important that due attention is paid to existing (cross-border) standards (e.g. Isa², OSLO, OSLO² of Core Public Services Vocabularies).
- Both the advantages as well as the disadvantages of data standardisation will need to be considered. Whereas increased standardisation may enable smoother data transfers, improve machine learning or support more competitive data markets, it can also entail negative externalities (i.e. improved profiling, increased privacy and cybersecurity risks).

Envisioned steps to address the challenges:

- Data Regulation:
  - Develop a clear and comprehensible platform or website that covers all applicable legislations (cf. CUTLER Project CiTiP).
- Data Maturity:
  - Data governance:
    - Define and distribute a method to evaluate a sector's data maturity (e.g. questionnaire).
    - Define and distribute a method to evaluate an organisation's data maturity (e.g. a questionnaire) in order to allow for the creation of a specific pattern that organisations can follow when dealing with data. This evaluation should enable organisations to set up a proper data governance framework in which they can define their processes, policies, practices and structures to optimise the collection, storage, use and dissemination of data.
    - Investigate the desirability of the introduction of (additional) data sharing obligations. The analogy and link with the former Public Sector Information Directive (PSI) and Open Data Directive is important in this regard. Alternative data governance models should also be developed and used to support a more equal division of power regarding data collection, storing and sharing.
  - Data management:
    - Collect, distribute and/or develop operational data maturity and governance models (or related best practices) that companies and organisations can apply themselves.
    - Familiarize the relevant organisations with such existing data maturity and governance models.
- Data Quality:
  - Develop quantitative and qualitative assessment criteria to define high-value datasets by using a multi-stakeholder approach (e.g. open data certificates). The data quality might also be evaluated by a supervisory authority (cf. AI Institutional Panel) before using an AI system.
  - Create user-friendly databases that use visualisation techniques. Measures might be taken to improve the accessibility, readability and interoperability of datasets.

    - Support companies and organisations in making use of their already available data. Collect and distribute related best practices and guidelines.
- Data Standardisation:
    - Enable, distribute and/or create open data or data collaborative standards and/or the related compatibility models.
    - Develop a methodology to define standards, done by individuals and communities.
    - Consider data access regime based on FRAND terms (Fair, Reasonable And Non-Discriminatory).

Expected results:

- Create or bring together an overview of relevant regulation and standards for specific data or sectors.
- Co-create a framework for data governance.
- Recommendations and guidelines (general sectorial and for regulators).
- Publish and communicate the results obtained from the envisioned steps.
- Advise stakeholders on the outcome obtained from the envisioned steps.
- Report on situation in other countries (comparative approaches).

Existing initiatives to consider:

Name	URL	Description
Data collaboratives	https://datacollaboratives.org	Data collaboratives is a new form of collaboration in addition to the public-private partnership model in which participants from different sectors -in particular companies- exchange their data to create public value.
Open Standaarden voor Linkende Organisaties (OSLO)	https://overheid.vlaanderen.be/producten-diensten/oslo2	Public services to citizens and entrepreneurs in Flanders are supported by various specialized applications from various software suppliers. The aim of 'Open Standaarden voor Linkende Organisaties' (OSLO) is to ensure greater coherence, and better understandability and/or findability of information and services.

Smart Flanders	https://smart.flanders.be/	On 1 January 2017, Minister Liesbeth Homans launched the Smart Flanders program. It is a support program that is conducted and performed by imec researchers. The Flemish government thereby supports several cities in their development towards smart cities.
Djust Connect Boer & Data	https://djustconnect.be/nl/	Focus on sharing data in the agri-food chain. Initiators are ILVO, Aveve, Boerenbond, CRV, DGZ and Milcobel.

Kenniscentrum
Data & Maatschappij

Kenniscentrum Data & Maatschappij
Pleinlaan 9
1050 Brussels

info@data-en-maatschappij.ai
www.data-en-maatschappij.ai