

Artificial intelligence solutions to accelerate impactful data-driven cancer innovation and research: a deep dive into data challenges

Nomi Byström & Robin Gustafsson

Abstract

While artificial intelligence (AI) has the capability to revolutionise cancer research and innovation, the Achilles heel is personal information: in particular its quantity, quality, data protection and data security. The issues are so severe as to hinder the critical expansion of cancer datasets for improved computational models to develop much-needed new treatments. This Policy Brief, the outcome of IN2MISSION research project funded by Business Finland, examines from the Finnish perspective both the causes of the data challenges and the means how to solve them in order to maximise the potential of AI solutions for cancer innovation and research.

Challenges and recommendations

Challenges

Challenge 1. Restrictive Data Access Under the Finnish Act on the Secondary Use of Health and Social Data

Challenge 2. The regulation of data protection – the GDPR and the Finnish Data Protection Act – hampers cancer research and innovation

Challenge 3. A restrictive approach to legal interpretation and use cases

Challenge 4. Uncertainty regarding the European Health Data Space

Policy recommendations

Recommendation 1. Reform the Finnish Act on the Secondary Use of Health and Social Data

Recommendation 2. Revisiting the regulation of data protection to find new cures

Recommendation 3. Encouraging pro-innovation interpretation of legislation and support hosting of data in the EU

Recommendation 4. Updating the proposed EHDS to permit harnessing AI for cancer research and innovation

Keywords: IN2MISSION, cancer innovation, cancer research, EU Mission on Cancer, Finnish Act on the Secondary Use of Health and Social Data, GDPR, European Health Data Space, Finnish Data Protection Act
JEL: I18, I19, I23, L50, O38

Introduction

The European Union has launched a Mission of paramount importance: the fight against cancer. As part of the Europe's Beating Cancer Plan, the goal of the Mission is to transform the lives of over three million people by the year 2030 through advancing prevention and treatments for a longer and healthier future.ⁱ

The significance of the Mission is underscored by the current trend in the union: according to the Commission, although even up to 40% of cancer cases are preventable, in 2020, 2.7 million people were diagnosed with the illness, and yet another 1.3 million lost their lives to it. The latter number includes over 2,000 young persons. While these figures give a numerical indication of those affected, they do not reveal the enormous toll of the illness on them – or their families and friends. Nor do they reveal the burden on healthcare.ⁱⁱ

Furthermore, globally Europeans are disproportionately affected by cancer. While making up only one-tenth of the world's population, still circa 25 per cent of all annual cancer cases are documented in Europe. The data suggests that only half of Europeans will go through life without, at some point, developing cancer.ⁱⁱⁱ Unless decisive action for new treatments is rapidly undertaken, cancer is estimated to become the leading cause of death in the union by the year 2035.^{iv}

To tackle the urgent need for new treatments, solutions based on personal data and the application of artificial intelligence (AI), in particular machine learning, hold great promise. And not only promise. Suffice to mention how available tumour data – by means of cross-modality integration, cross-cohort aggregation and reuse of data – is facilitating biomedical breakthroughs.^v

While AI is revolutionising cancer research and innovation, this development would not be possible without a surge in personal data. And herein lies the Achilles heel, for in the European Union and beyond, there are still major issues relating to data: particularly its quantity, quality, data protection, and data security. These challenges are hindering the critical expansion of cancer datasets for improved computational models to develop much-needed new treatments.^{vi}

This Policy Brief doubles down on the data challenges currently slowing down and even downright obstructing the development of AI-based research and innovation for new cancer cures. The overarching aim, by means of analysing the challenges, is to present recommendations so that the data bottlenecks can be removed and thereby the full potential of artificial intelligence can be leveraged to ensure that cancer will not become the leading cause of death in the European Union.

Material and methods

IN2MISSION, funded by Business Finland, is a research project of Aalto University and VTT that seeks to improve the understanding of innovation policy tools that enable advancements and innovations in data and platform economy for the acceleration of mission-oriented activities and system-level transformations.^{vii}

On 31 May 2023, IN2MISSION arranged an expert workshop on artificial intelligence solutions to accelerate impactful data-driven cancer innovation and research. Those invited to the workshop, representing research organisations and companies ranging from large universities to SMEs, are leading specialists in cancer research and care, AI and data protection.

The discussions of the workshop, together with reviews of the state of cancer in the European Union and Finland; the latter's data-driven cancer research and innovation as well as pertinent EU and national legislation form the basis of the challenges and recommendations presented in this Policy Brief.

Central EU and national legislation

Artificial intelligence solutions for cancer research and innovation require personal data, in particular health and genetic data that both belong to special categories of personal information, also known as sensitive data. The key EU law governing the processing of personal data is Regulation 2016/679^{viii}, better known as the General Data Protection Regulation (GDPR). Furthermore, in Finland two other pieces of legislation are central: first, the Finnish Data Protection Act (1050/2018)^{ix} which supplements the GDPR, and second, the Finnish Act on the Secondary Use of Health and Social Data (552/2019)^x.

Moreover, in the not too distant future, relevant EU regulation can be expected to increase as currently the Proposal for a Regulation on the European Health Data Space, COM/2022/197^{xi}, is being advanced. At the same time, as regards the upcoming Artificial Intelligence Act^{xii}, it may be worthwhile noting that scientific research is beyond its scope.^{xiii}

Finnish pioneering work in cancer care, research and innovation

How does Finland fare in tackling cancer? The country's healthcare system has been essential to the growth of cutting-edge treatment. Helsinki University Hospital (HUS), the first hospital in northern Europe to receive the Comprehensive Cancer Centre accreditation by the Organization of European Cancer Institutes, is recognised as a world leader in cancer care and research.^{xiv} The Finnish national R&D flagship program iCAN, funded by the Academy of Finland and with the founding hosts of HUS and the University of Helsinki, aims to boost breakthrough discoveries for improved treatments.^{xv}

Finnish universities, starting with the University of Helsinki and Aalto University, can boast researchers with top-notch expertise in both cancer and artificial intelligence.^{xvi} On the basis of the average number of citations per clinical cancer publication, Finland is the leading country in cancer research.^{xvii}

Finland can also take pride in the availability of state-of-the-art medical technologies, including on tackling cancer. Moreover, health tech is both the largest and among the fastest growing export segments of the country's high-tech industry.^{xviii}

Last but not least, in the promotion of national computing as well as data storage services, the significance of CSC – IT Center for Science Ltd needs to be underlined in the context of cancer research.^{xix}

Nevertheless, in Finland, too, the rate of cancer cases continues to rise.^{xx} At the beginning of the 1950s, when the Finnish Cancer Registry began operating, there were circa 8,000 new cancer cases per year. In 2017, the figure was already 34,261.^{xxi}

AI solutions for cancer research and innovation depend on data which, in turn, depend on legislation

Workshop participants, irrespective of the size or type of the organisation that they represented, underlined the pressing data-related needs, in particular hindrances regarding access to data; re-use of data; causal data and data sharing. These issues are so severe as to slow down or impede research and innovation for new cancer cures, a state of affairs which, in turn, undermines the ability to cure and support cancer patients.

The current situation derives from a legal bottleneck: both the laws themselves and how they are interpreted.

Challenge 1: Restrictive Data Access Under the Finnish Act on the Secondary Use of Health and Social Data

Initially, the proposed law on the Secondary Use of Health and Social Data held significant promise to meet the need for high-quality big data. However, during the drafting process limiting alterations were made, leading to a law that has proven counterproductive for cancer research and innovation.

- **Data access constraints.** In particular, the rule that all processing of data has to be physically carried out at Findata has restricted, even denied use of critical data.
- **Adverse consequences for university research.** The legislation has had detrimental effects on entire cancer research programmes at universities. Some programmes have even struggled to continue their work due to the restrictions imposed by the law.
- **Negative impact on SMEs.** The Act has resulted in substantial hurdles associated with accessing data and integrating it with internal datasets for small and medium-sized enterprises (SMEs). The situation has been described as dire for these companies, with data availability falling significantly short of their needs, making it challenging for them to conduct research and innovation effectively.

Recommendation 1: Reform the Finnish Act on the Secondary Use of Health and Social Data.

At the minimum, the law should be altered. Even rescinding and substituting it with entirely new legislation could be considered.

- **A new rule on the location of data processing.** It is essential that the rule in the current law that all accessing of data has to be physically carried out at Findata be repealed. Provided that agreed security rules are complied with, data access should be permitted in user's own environment.
- **End monopoly on anonymisation.** The monopoly held by Findata over the anonymisation of data should be reformed as there are actors with high-level capabilities on anonymisation.
- **Shift to technical data security focus.** The current emphasis on limiting data use should be replaced with an approach that requires state-of-the-art data security measures.

Challenge 2: The regulation of data protection – the GDPR and the Finnish Data Protection Act – hampers cancer research and innovation

- **Lack of adequate provisions for the application of AI in cancer research integration.** Regarding the processing of personal data in the context of scientific research, both laws contain special data processing rules and exceptions, for example in Articles 5(1)(b); 85 and 89 GDPR and section 6(1), paragraph 7 of the Data Protection Act that are designed to support and advance scientific research.^{xvii} Nevertheless, they are not enough to accommodate the uses of AI in data processing for cancer research, for example, the use of AI to validate biomarkers.
- **Innovation-provision gap in the data protection legislation.** Regarding the processing of personal data for innovation, there are no special innovation rules and exceptions. This constitutes a major obstacle to developing new AI-based cures for cancer patients.
- **Limitations of supervisory authorities.** Data protection authorities are also becoming AI regulators, yet all too often they lack the expertise to understand the uses of artificial intelligence for cancer research and innovation.

Recommendation 2: Revisiting the regulation of data protection to find new cures

- **Upgrading the data protection legislation for AI-based cancer research and innovation.** The rules on data processing for research should be updated to take into account the special needs to use artificial intelligence for new cancer cures. To give just one example, applying AI to validate biomarkers would have multiple benefits: lower trial costs and assist in opening the black box of explainability, etc. Likewise, the updated rules should be extended to cover cancer innovation to enable new treatments.
- **Promoting causal analysis.** Due to the significance of causal analysis in the context of cancer research, the regulation of data protection should be reformed to allow clinical trials for obtaining causal data as well as to permit more effective pooling of data. The potential of MyData-type solutions should also be developed and harnessed to this end.
- **Addressing the limitations of data protection authorities.** Since data protection authorities are also becoming AI regulators, their limitations in fields such as AI-based cancer research and innovation should not be bypassed, but rather, taken into account in the reforming of the regulation of data protection. In particular, the relevant medical and AI expertise should be consulted in the updating of the legislation.

Challenge 3: A restrictive approach to legal interpretation and use cases

The data hurdles stemming from the regulation itself are compounded by the prohibitory way in which the laws are interpreted. Such an environment is not conducive to accelerating much needed new, data-driven solutions.

- **Failing cancer innovation.** Not only are laws themselves but also their prohibitive interpretation, not least by authorities, is significantly impacting the use of data. This approach is not innovation friendly.
- **Use cases and hosting of data are not supported enough in the Union.** The interpretation of legislation stands in stark contrast, for example, the United Kingdom (cf. its Data Protection Act 2018) where hosting of data and use cases are supported. Bearing in mind that research is already heavily regulated, it is regrettable that the regulatory atmosphere in the union is restrictive and becoming even more so.

Recommendation 3:

Encouraging pro-innovation interpretation of legislation and support hosting of data in the EU

- **Encouraging pro-innovation interpretation.** Interpretation of legislation cannot be mandated by law, but it can be encouraged and recommended, not least by cancer researchers and research institutions and even the medical research community more broadly.
- **Use cases and hosting of data should be supported in the Union.** The various bodies: the supervisory authorities as well as the European Data Protection Board should take an active role in supporting use cases and hosting of data. While maintaining a robust data protection standard, this would encourage the development of new cancer treatments.

Challenge 4: Uncertainty regarding the European Health Data Space. The upcoming Regulation on the European Health Data Space (EHDS), especially due to its one-stop-shop mechanism and obligations for controllers, could improve use of data for cancer research and innovation. Nevertheless, there are also concerns.

- **The risk of replicating current data hurdles.** There is a growing concern that the EHDS may come to repeat hindrances caused by the Finnish Act on the Secondary Use of Health and Social Data. Specifically, repetition of the issues related to data access and usage as outlined in Challenge & Recommendation 1 could transform the upcoming EHDS into a missed opportunity.

Recommendation 4:

Updating the proposed EHDS to permit harnessing AI for cancer research and innovation

- **Allowing non-centralized data processing.** In the forthcoming EHDS, it is essential to avoid overly limiting the location of data processing, as this can significantly impede access to and processing of personal data. It is also crucial that cancer researchers be heard when the permitted locations will be agreed on.
- **Attention to data anonymisation.** In the field, there are actors with state-of-the-art capabilities on anonymisation and, hence, insistence on a monopoly of anonymisation would have a detrimental impact on cancer research and innovation.
- **Emphasising a technical approach to data security.** Rather than over-focusing on limiting the use of data, the EHDS should adopt an approach which underlines top-notch data security.

Contact

Project Researcher Dr. Nomi Byström

Aalto University, Department of Industrial Engineering and Management
P.O. Box 15500, 00076 AALTO
nomi.bystrom@aalto.fi

Professor Robin Gustafsson

Aalto University, Department of Industrial Engineering and Management
P.O.Box 15500, FI-00076 Aalto
Tel. +358 50 316 0981
robin.gustafsson@aalto.fi
@robingustafs #platformeconomy
#IN2MISSION

References

- i https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe/eu-missions-horizon-europe/eu-mission-cancer_en
- ii https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/promoting-our-european-way-life/european-health-union/cancer-plan-europe_en
- iii https://joint-research-centre.ec.europa.eu/jrc-news-and-updates/cancer-europe-5-things-data-tells-us-2022-01-13_en
- iv https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/promoting-our-european-way-life/european-health-union/cancer-plan-europe_en
- v <https://www.nature.com/articles/s41568-022-00502-0>
- vi <https://www.nature.com/articles/s41568-022-00502-0>
- vii <https://www.aalto.fi/en/IN2MISSION>
- viii <https://eur-lex.europa.eu/eli/reg/2016/679/oj>
- ix <https://www.finlex.fi/fi/laki/ajantasa/2018/20181050>; The English translation may be found at: <https://www.finlex.fi/en/laki/kaannokset/2018/en20181050.pdf>
- x <https://www.finlex.fi/fi/laki/ajantasa/2019/20190552>. The English translation of the Act can be found at: <https://stm.fi/documents/1271139/1365571/The+Act+on+the+Secondary+Use+of+Health+and+Social+Data/a2bca08c-d067-3e54-45d1-18096de0ed76/The+Act+on+the+Secondary+Use+of+Health+and+Social+Data.pdf>
- xi <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0197>; <https://www.europarl.europa.eu/legislative-train/theme-promoting-our-european-way-of-life/file-european-health-data-space>
- xii <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206>; https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.html; <https://www.euractiv.com/section/artificial-intelligence/news/eu-policymakers-enter-the-last-mile-for-artificial-intelligence-rulebook/>
- xiii Ibid. See also <https://sciencebusiness.net/news/AI/european-parliament-agrees-research-should-have-exemptions-draft-artificial-intelligence-law>; <https://eurotech-universities.eu/news-and-events/ai-act-by-eurotech-universities/>
- xiv <https://mediabank.businessfinland.fi/ljW9w6JLRxGm>
- xv <https://ican.fi/>; <https://www.csc.fi/-/better-tailored-care-through-advanced-data-transfer-and-analysis-methods>
- xvi See e.g. <https://www.aalto.fi/en/news/scientists-unravel-the-effects-of-new-medication-for-advanced-melanoma-with-the-help-of-ai>; <https://www.helsinki.fi/en/news/cancer/ai-model-reveals-how-bodys-defence-system-recognises-skin-cancer>
- xvii <https://mediabank.businessfinland.fi/ljW9w6JLRxGm>
- xviii <https://www.businessfinland.fi/en/whats-new/news/2023/finnish-health-technology-is-the-golden-egg-of-export>; <https://mediabank.businessfinland.fi/ljW9w6JLRxGm>
- xix <https://www.csc.fi/-/better-tailored-care-through-advanced-data-transfer-and-analysis-methods>
- xx <https://thl.fi/fi/web/kansantaudit/syopa/syovan-yleisyys>; <https://cancerregistry.fi/>
- xxi <https://www.terveyskirjasto.fi/asy00614>
- xxii <https://eur-lex.europa.eu/eli/reg/2016/679/oj>; <https://www.finlex.fi/en/laki/kaannokset/2018/en20181050.pdf>; <https://tietosuojafi/en/scientific-research-and-data-protection>