

Trine Rask Nielsen, Panagiota Katsikouli, Anna Højberg Høgenhaug, William Hamilton Byrne, Thomas Gammeltoft-Hansen, Tijs Slaats, Henrik Palmer Olsen, Thomas Troels Hildebrandt, Naja Holten Møller (2021): Confronting Asylum Decision-making through Prototyping Sensemaking of Data and Participation. In: Proceedings of the 19th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2021_p08

Confronting Asylum Decision-making through Prototyping Sensemaking of Data and Participation

Trine Rask Nielsen, Panagiota Katsikouli, Anna Højberg Høgenhaug, William Hamilton Byrne, Thomas Gammeltoft-Hansen, Tijs Slaats, Henrik Palmer Olsen, Thomas Troels Hildebrandt, Naja Holten Møller

Department of Computer Science & Faculty of Law, University of Copenhagen
trn@di.ku.dk, pk@di.ku.dk, ahh@jur.ku.dk, william.hamilton.byrne@jur.ku.dk, tgh@ju.ku.dk, slaats@di.ku.dk, henrik.palmer.olsen@jur.ku.dk, hilde@di.ku.dk, naja@di.ku.dk

Abstract. The research outlined in this paper is focusing on understanding asylum decision-making and outcome variations across the Nordic countries. As a preliminary study, we extract data from an open dataset¹ of decision summaries from the Danish Refugee Appeals Board. The approach we propose, offers points for discussion of how prototyping context and participation can help raise questions about such data and engage stakeholders. Combining the application of Machine Learning (ML) and Natural Language Processing (NLP) with participatory methods (e.g. critically designed artefacts) enable us to 1) move beyond “obvious” ML-application areas, 2) through sensemaking of data with stakeholders, and 3) co-develop approaches to data science from a CSCW-perspective.

¹ <https://fln.dk/da/Praksis>

Introduction

In CSCW and related areas of research scholars are starting to investigate how to respond to and create a shared research agenda for the refugee crisis (Talhouk et al. 2016, 2019; Almohamed et al. 2020; Tachtler et al. 2021). Increasing interest in data-driven technologies for decision-support has led to scholars asking new questions about the role of technology in relation to large-scale political issues (Dombrowski et al. 2016; Molnar 2019; Costanza-Chock 2020). As such, scholars begin to discuss the entrenched inequities of data science tools among the already marginalized and how they locate wealth and power largely in Western societies (Taylor et al. 2021). Acknowledging these deep issues of inequity “data are part of the problem, to be sure. But they are also part of the solution” (D’Ignazio et al. 2020).

The research outlined in this paper is part of the interdisciplinary research project “Data Science for Asylum Legal Landscaping” (DATA4ALL). The project leverages data science to understand outcome variations in asylum adjudications across Nordic countries. Despite decades of legal harmonization, the chance of receiving asylum for displaced people from the same country varies significantly across Europe (Goodwin-Gill et al.; Guild 2016).

According to the 1951 Refugee Convention Article 1A, a refugee is defined as a person:

“who is outside his or her country of nationality or habitual residence; has a well-founded fear of being persecuted because of his or her race, religion, nationality, membership of a particular social group or political opinion; and is unable or unwilling to avail him – or herself of the protection of that country, or to return there, for fear of persecution” (The UN Refugee Agency)

The 1951 Refugee Convention does not define how states should determine refugee status. Asylum proceedings and status determinations are left to each state. In this context, asylum claims are left for adjudicators to decide on the basis of testimonies and the disclosed material, which become adjudications dataset available for data science (New to Denmark).

To that end, the paper proposes an approach and initial points for discussion of how a CSCW-perspective can help raise questions about such dataset. Data from asylum adjudications stems from a complex and cooperative decision-making practice, which is only opaquely described as they are extracted. Contextual factors that may affect data production, include imprecise language interpretation (e.g. Weibert et al. 2019), lacking trust in authorities (Almohamed et al. 2020), as well as implicit bias of adjudicators can also shape interpretations of asylum claims (Chen et al 2017).

A CSCW-perspective centers on showing the boundaries and inadequacies of applying techniques from data science such as Machine Learning (ML) and Natural Language Processing (NLP) on an “incomplete” dataset and the context and practices that shaped it. A premise for this project is thus the combination of

methods and approaches, including critical thinking. We design our method to also consider questions such as *how* data are constructed and what counts as useful data in the formal legal process. Prototyping critical research artefacts, we propose, serves as a tool for inquiry into both practice and the wider political context of sensitive areas such as asylum decision-making. In addition, the paper offers some initial examples from the application of ML and NLP to asylum decision summaries from the Danish Refugee Appeals Board.

Machine Learning of Asylum Decision-Making

Machine learning (ML) is increasingly being used in various decision-making contexts including in public administration. ML is being endorsed to support decision-making in adjudications involving applications for asylum (Chen et al. 2017) to decrease variation in decision outcome (Ramji-Nogales et al. 2007), improve impartiality, and decrease “unfair” decisions made by human judges (Heyes et al. 2019; Chen et al. 2017). Large-scale datasets are not yet being applied to asylum-decision making in Denmark, though ML has been increasingly applied in sensitive settings (e.g., social welfare) fueling debates on its use (Flügge et al 2021). In Denmark, the government has introduced a new National strategy in 2019 with the aim of becoming a frontrunner in AI through an ethical and responsible perspective on machine learning use in public casework: “The public sector should take advantage of AI to provide a world-class service” (Ministry of Industry, Business and Financial Affairs 2019).

Meanwhile, large-scale data are gaining attention in the asylum domain. In 2020 it became public that adjudications of recognition rates of one judge of the Danish Refugees Appeals Board varied significantly from the adjudications of other judges between 2012-2019 (Flygtningenævnet 2020). Prior CSCW-research have forcefully documented the importance of unpacking the situated and cooperative aspects of decision-making in practice; thus, careful investigation of the social organization of work is critical for giving context to such data (Randall et al. 2007, Møller et al. 2020). Cakici et al. show that questions of “Othering” of displaced individuals that are not considered European are not confined to practices of asylum decision-making (Cakici et al. 2020). “Hence, engaging in situated analyses of data practices means confronting big political questions revolving around Europe, treating the different practices through which the people of Europe are brought into being as political struggles over questions such as “Who are the people of Europe?” (Cakici et al. 2020: 204).

As large-scale datasets become available, anticipation grows that these can be transformed into knowledge to inform decisions, increasing algorithmic authority (Lustig et al. 2016; O’Neil 2016). Legal decision-making more broadly has seen an increase in algorithmic systems for decision-support (Zalnierute 2019; Olsen et al. 2020). In asylum decision-making, interest in data and algorithms entail

streamlining and increasing transparency in decision-making (Molnar 2019; Pakzad 2019).

Raising questions about large-scale data from a CSCW-perspective is critical because ML and NLP are applied on “incomplete” data stemming from a complex and cooperative decision-making practice that is only opaquely described by the data available for the data scientists.

Variations Identified in Asylum Cases

The DATA4ALL project will use data science techniques for the purpose of explanatory research. As a preliminary study, we extracted publicly available data from decision summaries of asylum cases treated by the Danish Refugee Appeals Board (The Danish Refugee Appeals Board). The dataset has 3 obvious limitations. First, it only contains cases that were, at a first instance, rejected by the Danish Immigration Service. Second, not all cases treated by the Refugee Appeals Board are publicly available and it is therefore not representative of the complete set of asylum cases. Third, the dataset only contains the summaries of the decisions.

After applying NLP and regular expressions on the decision summaries, we extracted the following information for every case: applicant’s nationality, gender, religion, date of entry in Denmark, date the case was processed by the Refugee Appeals Board, marital status, asylum category, whether the applicant has previously applied for asylum in another country, whether the applicant has had involvement with political parties/organizations and the military, whether divergences were noticed in their application, whether an investigation was carried out in cases of torture, and the Refugee Appeal’s Board decision.

From this analysis, we found varying rates between cases being granted asylum vs the cases not granted it, in relation to each extracted feature. Some examples, with regards to the applicant’s nationality are highlighted in figure 1.

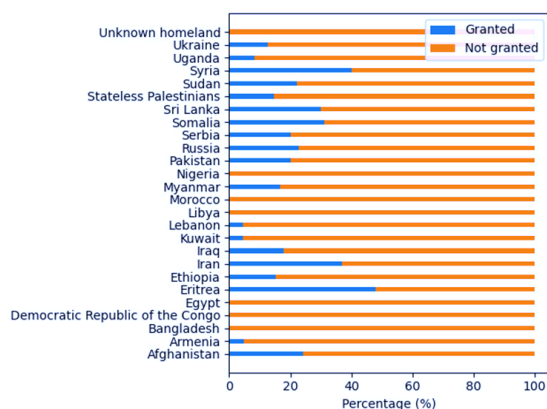


Figure 1. Percentage of applicants granted asylum according to country of origin. Where the country of origin is unknown (unknown homeland), no one has been granted asylum in the appeal process by the Danish Refugee Appeals Board between 2003-2020.

We then built a number of classifiers on all possible combinations of features in order to investigate the models’ accuracy on predicting case outcome. We received varying classification accuracies depending on the classifier used, and the combination of features, with values ranging from 43% (when SVM is used) to 82% when Random Forest was applied on the features: nationality, gender, and religion.

An initial analysis of these data indicates that applicants where the country of origin is not disclosed are not granted asylum. This points to a system built around nation states that disadvantages those applicants who do not conform with traditional expectations of formal processes through which the people of Europe are “brought into being” (Cakici et al. 2020).

The goal of our project is to provide a deeper interpretation and contextualization of data from asylum decision-making, for example, asking questions to the data on the role of the asylum application form and how the applicant is shaped through data as an asylum seeker.

Our aim with this preliminary paper is to highlight the challenges that arise from data science requests for a clean, representative and complete dataset, in order to draw confident conclusions with regards to the underlying reasons of the variations detected.

Prototyping Sensemaking of Data and Participation

To inform the leveraging of data science techniques in asylum decision-making, we propose participatory methods (e.g. Menendez-Blanco et al. 2017) that enable sensemaking of data (Neff et al. 2017) together with stakeholders. Prototyping covers a spectrum from product development (Andersen et al. 2017) to critically designed artefacts that can form tools for inquiry (Baumer 2017; Menendez-Blanco et al. 2017; Baumer et al. 2018).

In CSCW and related fields of research, critically designed artefacts serve as “containers of ideas” (Bødker 1998), meant to act as heuristics, and as catalysts for both sense-making and future thinking (Danholt 2005). Thus, such artefacts can enable adversarial design and serve as material evidence in political discourse: “By revealing the conditions of political issues and relations, adversarial design can identify new terms and themes for contestation and trajectories for action” (DiSalvo 2011:13).

Prototyping is especially useful to 1) engage stakeholders in the asylum domain and connect large-scale data to its situated context while promoting data literacy of stakeholders. Further, prototyping allows us to 2) unpack the different political struggles influencing variations in asylum outcomes that cannot be ascribed to the situated practices of asylum decision-making (Teli et al. 2018). 3) Data are transformative of future practices and prototyping enables stakeholders who are not trained in data sciences to co-develop methods and take part in the discussion and design of databased services (Seidelin et al. 2020).

In order to get a deep understanding of the Danish asylum decision-making process, an overview of the current Danish asylum procedure was created as a starting point for prototyping a critically designed artefact. Taking this approach (*following* Menendez-Blanco et al. 2017) can enable us to query into the kind of data from asylum cases that we present here as example. Thus, a first prototype is currently being developed that amplifies the role of data and how they can enter the formal decision-making process.

Prototyping critically designed artefacts serves two purposes in this project: 1) gaining a shared understanding and common vocabulary in regard to Danish refugee terms and processes, and 2) mutual learning (Kensing et al. 2013) through participation of all stakeholders to establish a platform for sensemaking of data and discussions of the work practices of asylum decision-making, such as the subtle categorization of cases (Møller et al. 2011; Møller et al. 2020).

Finally, the sensitive nature of asylum decision-making makes it urgent to amplify asylum seekers and their advocates perspectives (Talhouk et al. 2019). Through applying a critical approach, it becomes possible to also research “counter data” not available in formal decision summaries. Thus, counter data is missing data or data that are currently not collected because of e.g., bias, lack of social and political will, or structural oppression (D’Ignazio et al. 2020).

Closing Remarks

The project DATA4ALL and the approach proposed in this paper is interdisciplinary, which is pivotal to how we engage stakeholders. Our approach seeks to maximize synergies across disciplinary boundaries (Law and Computer Science), institutions (e.g., the Danish Refugee Appeals Board), and advocates for asylum seekers (e.g., the Danish Refugee Council and Refugees Welcome).

All case summaries retrieved for this study were anonymized upon retrieval. The extensive datasets of over 100,000 asylum adjudications across the Nordic countries that cannot be anonymized are securely stored and handled in accordance with GDPR. An ethical and responsible data science approach is imperative for any project engaged with asylum cases of this kind. In this respect, we adopted a strategy for amplifying stakeholder participation as advocated in ethical guidelines for AI (e.g., European Union 2019 and Dansk Standard 2020).

Consequently, we commit as interdisciplinary scholars to an ontology that recognize asylum applicants and decision-makers as active participants in the shaping of transformative data and possible near-future practice and technologies in the field of asylum decision-making.

We call on the CSCW-community to commit to researching and disclosing the boundaries and inadequacies of applying data science techniques such as ML and NLP and recognize the transformative potential of data for all.

Acknowledgements

We would like to direct a special thanks to the Danish Refugee Council in Greece, people from the Victoria Square Project and the Piraeus Open School for Migrants - as well as the people we engaged with who are subject to asylum decision-making or advocates of refugees- and asylum seekers for generously sharing their perspectives on data practices and the decision-making processes in the asylum domain. We thank colleagues Vasilis Vlachokyriakos and Yannis Efthymiou at Open Lab, Newcastle University and Open Lab:Athens who discussed the approach of the DATA4ALL project with us. Sarah Scott Ford, Raphaële Xenidis, and Irina Shklovski engaged in discussions with us on the application of ML with concern for legal- and ethical perspectives. This research was conducted as part of the Data Science for Asylum Legal Landscaping (DATA4ALL) project, funded by the University of Copenhagen Data+ program.

References

- Almohamed, A., Zhang, J., and Vyas, D. (2020): ‘Magic Machines for Refugees’. (*COMPASS '20*). Association for Computing Machinery, New York, NY, USA, pp. 76–86.

- Andersen, T. O., Bansler, J. P., Kensing, F., and Moll, J. (2017): 'From Prototype to Product: Making Participatory Design of mHealth Commercially Viable', *Stud Health Technol Inform*, 233, pp. 95-112.
- Baumer, E. P., Berrill, T., Botwinick, S. C., Gonzales, J. L., Ho, K., Kundrik, A., Kwon, L., LaRowe, T., Nguyen, P. C., Ramirez F., Schaedler, P., Ulrich, W., Wallace, A., Wan, Y., and Weinfeld, B. (2018): 'What Would You Do? Design Fiction and Ethics', *2018 ACM Conference on Supporting Group Work*, pp. 244-256.
- Baumer, E. P. (2017): 'Toward human-centered algorithm design', *Big Data & Society*, vol. 4, no.
- Bødker, S. (1998): 'Understanding Representation in Design', *Human-Computer Interaction*, vol. 13, no. 2. LEA Publishers, London, pp. 107-127.
- Cakici, B., Ruppert, E., and Scheel, S. (2020): 'Peopling Europe through Data Practices: Introduction to the Special Issue', *Science, Technology, & Human Values*, 45(2), pp. 199–211.
- Chen, D. L. and Eigel, J. (2017): 'Can machine learning help predict the outcome of asylum adjudications?', *The 16th edition of the International Conference on Artificial Intelligence and Law*, pp. 237–240.
- Costanza-Chock, S. (2020): *Design Justice: Community-Led Practices to Build the Worlds We Need*. MIT Press.
- Danholt, P. (2005): 'Prototypes as performative', *ACM Conference on Critical Computing: Between Sense and Sensibility*, pp. 1–8.
- 'Dansk Standard', 2020, 'DS/PAS 2500-2:2020' Retrieved January 20, 2021 from <https://webshop.ds.dk/da-dk/søgning/35-020-informationsteknologi-it-generelt/ds-pas-2500-22020>)
- D'Ignazio, C., and Klein, L. F. (2020): *Data feminism*, MIT Press.
- Disalvo, C. (2015): *Adversarial Design*. The MIT Press, Cambridge, Massachusetts. London, England.
- Dombrowski, L., Harmon, E. and Fox S. (2016): 'Social justice-oriented interaction design: Outlining key design strategies and commitments', *2016 ACM Conference on Designing Interactive Systems (DIS '16)*. Association for Computing Machinery, New York, NY, USA, pp. 656–671.
- 'European Union', 2019, Retrieved January 20, 2021 from <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>
- Flügge, A. A., Hildebrandt, T., and Møller, H. N. (2021): 'Street-Level Algorithms and AI in Bureaucratic DecisionMaking: A Caseworker Perspective', *ACM on Human-Computer Interaction*, 5, CSCW1, Article 40 (April 2021)
- 'Flygtningenævnet', 2020, Retrieved January 13, 2021 from <https://fln.dk/-/media/FLN/Koordinationsudvalg/Referat-at-K-udvalgsmde-den-27-oktober-2020-med-bilag.pdf?la=da&hash=5DE49861FEBAF5ADB9121A62214B4AA6A89BAC68>
- Goodwin-Gill, G., and Lambert, H. (Eds.). (2010): *The Limits of Transnational Law: Refugee Law, Policy Harmonization and Judicial Dialogues in the European Union*. Cambridge: Cambridge University Press.
- Guild, E. (2016): 'Does the EU Need a European Migration and Protection Agency?', *International Journal of Refugee Law*, 28(4), pp. 585–600.
- Heyes, A. and Saberian, S. (2019): 'Temperature and Decisions: Evidence from 207,000 Court Cases', *American Economic Journal: Applied Economics* 11, 2 (April 2019), pp. 238–65.
- Kensing, F. and Greenbaum, J. (2013): *Heritage: Having a Say*. Simonsen, J. and Robertson, T. (Eds). Routledge International Handbook of Participatory Design, pp. 21-36.

- Lustig, C., Pine, K., Nardi, B., Irani, L., Lee, K. M., Nafus, D., and Sandvig, C. (2016): 'Algorithmic Authority: the Ethics, Politics, and Economics of Algorithms that Interpret, Decide, and Manage'. *ACM CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, pp. 1057–1062.
- Menendez-Blanco, M., Bjorn, P., and De Angeli, A. (2017): 'Fostering cooperative activism through critical design'. *2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pp. 618-629.
- 'Ministry of Industry', 'Business and Financial Affairs', 2019: Retrieved January 13, 2021 from <https://eng.em.dk/news/2019/marts/denmark-should-be-a-frontrunner-in-responsible-development-and-use-of-artificial-intelligence-ai/>
- Molnar, P. and Gill, L. (2018): 'Bots at the Gate: a human rights analysis of automated decision-making in Canada's immigration and refugee system'. *The International Human Rights Program and the Citizen Lab*. Retrieved January 13, 2021 from <https://it3.utoronto.ca/wp-content/uploads/2018/10/20180926-IHRP-Automated-Systems-Report-Web.pdf>
- Møller, N. H., Shklovski, I., and Hildebrandt, T. (2020): 'Shifting Concepts of Value: Designing Algorithmic Decision-Support Systems for Public Services', *ACM Nordic Conference on Human-Computer Interaction*, Article 70, pp. 1–12.
- Møller, N. H. and Bjørn, P. (2011): 'Layers in sorting practices: Sorting out patients with potential cancer', *Computer Supported Cooperative Work (CSCW)*, 20(3), pp. 123-153.
- Neff, G., Tanweer, A., Fiore-Gartland, B. and Osborn, L. (2017): 'Critique and contribute: A practice-based framework for improving critical data studies and data science', *Big Data*, 5(2), pp. 85-97.
- 'New to Denmark', Retrieved January 11, 2021 from <https://www.nyidanmark.dk/en-GB/Applying/Asylum/Adult%20asylum%20applicant>
- Olsen, H. P., Slosser, J. L. and Hildebrandt, T. (2020): 'What's in the Box? The Legal Requirement to Explain Computationally Aided Decision-Making in Public Administration' (April 19, 2020). *Constitutional Challenges in the Algorithmic Society*. OUP, 2020, University of Copenhagen Faculty of Law Research Paper no. 2020-97.
- O'neil, C. (2016): *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
- Pakzad, R. (2019): 'Opportunities and Challenges of Emerging Technologies for the Refugee System'. *World Refugee Council*. Centre for International Governance Innovation, no. 11.
- Ramji-Nogales, J., Schoenholtz, A. I. and Schrag, P. G. (2007): 'Refugee roulette: Disparities in asylum adjudication', *Stan. L. Rev.* 60 (2007), 295.
- Randall, D., Harper, R. and Rouncefield, M. (2007): *Fieldwork for Design: Theory and Practice (Computer Supported Cooperative Work)*. Springer-Verlag, Berlin, Heidelberg.
- Seidelin, C., Dittrich, Y., and Grönvall, E. (2020): 'Foregrounding data in co-design—An exploration of how data may become an object of design', *International Journal of Human-Computer Studies*, 143, 102505.
- Tachtler, F, Aal, K, Ertl, T, Diethel, D, Niess, J, Khwaja, M, Talhouk, R, Vilaza, GN, Lazem, S, Singh, A, Barry, M, Wulf, V and Fitzpatrick. (2020): 'Artificially Intelligent Technology for the Margins: A Multidisciplinary Design Agenda'. *CHI Conference on Human Factors in Computing Systems Extended Abstracts: CHI '21 Extended Abstracts*. ACM, New York, ACM CHI Conference on Human Factors in Computing Systems (CHI 2021), Yokohama, Japan, 8/05/21.

- Talhouk, R., Ahmed, I. S., Wulf, V., Crivellaro, C., Vlachokyriakos, V., and Olivier, P. (2016): 'Refugees and HCI SIG: The Role of HCI in Responding to the Refugee Crisis', *CHI'16 Extended Abstracts*, San Jose, CA, USA ACM.
- Talhouk, R., Aal, K., Weibert, A., Krüger, M., Wulf, V., Fisher, K., Tachtler, F., Shahid, S., Ahmed, I., S., and Duarte, B., M., A. (2019): 'Refugees & HCI SIG: Situating HCI Within Humanitarian Research', *CHI'19 Extended Abstracts*, Glasgow, Scotland UK.
- Talhouk, R., Balaam, M., Toombs, A. L., Garbett, A., Akik, C., Ghattas, H., Araujo-Soares, V., Ahmad, B., and Montague, K. (2019): 'Involving Syrian refugees in design research: Lessons learnt from the field', *The 2019 on Designing Interactive Systems Conference*, pp. 1583-1594.
- Taylor, A., Rosner, D., and Wiberg, M. (2021): 'AI Activism', *ACM Interactions*, pp. 5.
- Teli, M., De Angeli, A., and Menéndez-Blanco, M. (2018): 'The positioning cards: on affect, public design, and the common', *AI & SOCIETY*, 33(1), pp. 125-132.
- 'The Danish Refugee Appeals Board', Retrieved January 13, 2021 from <https://fln.dk/da/Praksis>
- 'The UN Refugee Agency', Retrieved January 11, 2021 from <https://www.unhcr.org/what-is-a-refugee.html>
- Weibert, A., Krüger, M., Aal, K., Salehee, S. S., Khatib, R., Randall, D., & Wulf, V. (2019): 'Finding language classes: Designing a digital language wizard with refugees and migrants'. *ACM Hum.-Comput. Interact.* 3, CSCW, Article 116 (November 2019), 23 pp.
- Zalnieriute, M., Moses, L. B., and Williams, G. (2019): 'The Rule of Law and Automation of Government Decision-Making', *The Modern Law Review* 82, no. 3, pp. 425-455.