Mapping CV as an Assemblage of (Unfair) Sociotechnical Relations

Garfield Benjamin
Solent University, UK
garfield.benjamin@solent.ac.uk

Abstract

This paper presents a sketch for mapping CV as a field and set of sociotechnical practices. It takes a relational approach, centring those who are most affected by an AI system, identifying points of unfairness, situating systems in wider discourses and narratives, and thereby locating points of intervention towards the redistribution of power.

1. Introduction

There has been an increasing recognition of the potential harms of computer vision (CV) when applied to specific social contexts. But there are often difficulties in translating these concrete harms to the broader context of the societal impact of the field at large. Similar problems persists in other areas of AI/ML research, as well as to computational sciences more generally. Socially engaged - particularly Black Feminist and Queer/Trans - scholarship has shown how the colonial, racist, sexist, ableist and otherwise discriminatory narratives that define academic research and development or deployment of technologies [4, 20] can be applied to AI more generally [6] and specific applications of CV [7, 18, 22]. Building on dominant public narratives around social inequalities, these discussions have led to movement in the direction of fairness, accountability and transparency - most notably in the FAccT conferences. But the discourses surrounding these terms have limitations [15]. Fairness is often assumed to offer a measurable test of inequality, and even suggests the possibility that these problems can be solved in specific situations. But in AI and CV, this tends to lead to a focus on specific data-oriented issues like bias and the assumption that ironing out what are essentially seen as errors in the dataset will be able to resolve social problems. Moving past narratives of fairness and equality towards perspectives on inequity, justice, uneven distribution of power, harms and benefits, and other *structual* problems necessitates a much deeper, social, relational, narrative approach to analysing the field and locating wider sites of action.

This is not to discount the usefulness of locating specific areas of 'unfairness', or the need for specific interventions, providing we look beyond conceptualisations of bias in datasets, and even beyond the algorithms themselves [13, 21, 1], towards a more structural, relational and narrative scope. Moving beyond fairness entails, in part, examining *un*fairness as asymmetric power structures and the inequitable distributino of decision-making, benefits, and subjection to harms. Where, exactly, does unfairness lie? Or, where can we locate sites of potential action to reduce unfairness and the societal roots of unfairness? This paper provides a tool for mapping CV as a field - a set of people, narratives and relations - in order to assess the distribution of issues of fairness throughout the field. It builds on previous mappings of AI in specific contexts [3] to use as a tool for examining influences, narratives and power relations, assessing the limits of fairness against structural inequities, and thereby identifying points where fairness can be a useful intervention. It therefore also points to where other forms of intervention are needed in order to promote fairness or to support principles such as justice, equity and the redistribution of power.

2. MAPPING CV

2.1. Principles

Deleuze and Guattari outline mapping as a social and political action:

> open and connectable in all of its dimensions; [...] susceptible to constant modification. It can be [...] reworked by an individual, group, or social formation [...] The map has to do with performance, whereas the tracing always involves an alleged "competence". [10]

Similarly, against histories of mapping as constituting colonial patriarchal discourses, Katherine McKittrick asks:

> If our expressive demands can demonstrate a new worldview, in what ways can ethical human geographies, or interhuman geographies, be mapped? [19]

These points are particularly useful when mapping a field like CV. Maps change over time and must be constantly reworked by different groups, particularly affected communities. Mapping problematises narratives of "competence" in which technical concerns narrow our understanding of a field to reinforce unjust norms and hierarchies. While CV will inevitably remain the focus, we can destablisise it by decentring the controlling narratives, aiming for what intersectional feminism sees as "a cartography of governance that forces an official reckoning of a new way of seeing" [17]. A more ethical and equitable cartographic practice of CV therefore follows several key principles.

Relationality is fundamental to understanding asymmetric power structures. This can include ethics of justice [5] (addressing current harms) and care [12] (building new equitable structures) ethics but also leans on performativity [2] to examine the social construction of norms through different actors in asymmetric contexts. To understand the power structures and narratives of CV as a field - to assess fairness and related principles beyond individual algorithms or applications - requires such contextually embedded approaches to the relations that constitute the field.

Centring the margins is perhaps the most important step in building a tool for critically examining any field, and CV in particular, building on Black Feminism [16] and postcolonialism [23], its application to AI [5], data [11], and the act of starting from marginalised intersectional perspectives when discussing tech [14]. Intersectional approaches to breaking down categories and mapping the margins can help us to examine how dominant narratives shape conceptualisations of specific issues and thereby define, control or conceal injustices [9]. This may be less obvious for more foundational or theoretical aspects of the field than, say, a concrete technical product. But by tracing existing uses and then flipping the focus to those affected, we can retrace our steps back to see the decision-making and narrative relations that impact on people's everyday lives.

Algorithmic justice is an important principle. This includes stopping unjust systems at all levels, from refusing the use of data [8], to labour, to ecology, to culture, to harms in practice and the perpetuation of oppressive narratives and assumptions of the field. Justice highlights broader points of intervention (and resistance) within the field. It goes beyond fairness to locate points at which alternative decision-making is necessary. This may involve disrupting mainstream narratives within the field.

These principles require interdisciplinary perspectives across research, policy, industry and public spheres, elevating marginalised voices and intersectional approaches including (but not limited to) critical race, feminist, queer, trans and disability theories and communities.

2.2. Structure

Mapping CV as a sociotechnical assemblage acts as a social audit of the field. It allows us to construct different forms of objections, and the process itself can be a means of elevating different voices and narratives, with space for community or wider public consultation. Roles, components, entities and narratives can be placed on the map using information gathered indirectly from research, advocacy groups, the press, or directly in conversation with specific (affected, labour, research or user) communities. The proposed mapping method consists of concentric layers, as follows:

Centre: those affected by a CV system, particularly those with less power/choice. This is those whose data is used, those upon whom CV is used to make decisions, and all those otherwise affected by the use or development of CV;

Interface: CV as a technical objects, including constituent technologies and their interactions. The boundaries of this layer are where the interactions occur, and the placement of this layer between those affected and those using/designing/deciding the system is a purposeful and political gesture to highlight the anonymising effects of CV as an interface for social issues;

Labour: all those whose labour contributes to the designing, building and running of a CV system - operators and users of the interface layer, the researchers and engineers who design it, and the manufacture of physical components;

Decision-making: moving outwards, this includes the relative power of decision-making by researchers, businesses, organisations, clients, funders or policy-makers, that defines the development and deployment of CV;

Discourse/Narrative: also known as the principles and influence layer, this identifies the systemic forces that define the context in which the decisions are made, the broader narratives that constitute CV as a field. This may be fairly consistent across different applications - particularly in more theoretical sub-fields - and includes the press, business interests, political (including military) interests, funding body priorities, the relations between disciplines and sub-disciplines, historical and global contexts, and other aspects that shape the research community.

Individuals, groups, roles and interests are located on the map and linked to show directional relations. It should be noted that directions emphasise power structures - they can in many cases be simply reversed to denote extractive benefits. Placement within a layer is relative to other layers. This allows designation of, for example, relative levels of decision-making power or harm (such as exploitation) within labour. Figure 1 shows this method of mapping as a generic template, the structure of the concentric layers. Possible additions include identifying points of current or possible intervention - whether within fairness narratives or beyond - or highlighting different (global, application area,

sub-field, etc.) contexts as iterations of the map or elements to be adjusted depending on the specific context (such as specific funding bodies or language considerations).

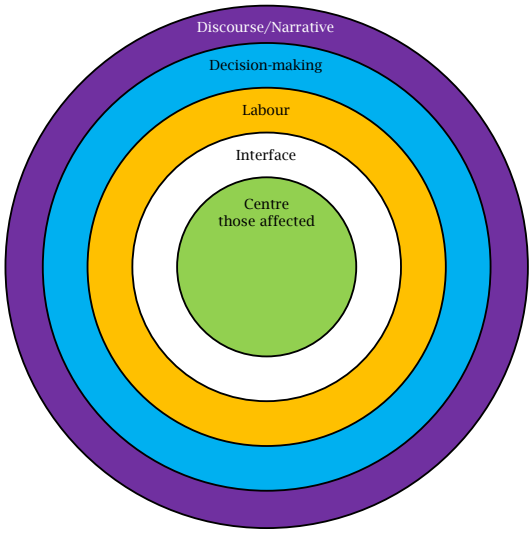

Figure 1. Mapping AI structure

3. THE CV MAP

Figure 3 shows an example CV Map. It follows the structure and principles outlined above, with some additional considerations as it is applied to a field rather than a specific application or use. The map was built simultaneously from the centre and the edge, partly for practical purposes, partly to make instant traceable links between the real affects of CV and the wider narratives that shape the field, and partly to engage in the process of moving the centre or centring the margins. Affected groups, narratives, and then actors and technological artefacts, were then filled in to enable this tracing of concrete links. Some areas include types of element, with ellipses to denote the possibility of further iterations. Relations are built up between areas and types of element to show lines of influence (and funding streams) that shape the directions and priorities that in turn continually re-constitute the narratives that form the field.

This version of the map falls into a few general areas, with major lines of influence moving between the locations of power within each area. This includes: research topics (stemming from overarching research agendas); researchers and labour leading to the algorithms themselves - echoed across the hardware and data streams that support this; application areas and the companies creating practical systems; global contexts as iteration point of the map; institutions and funders; regulators and policy-makers (linked to press and media); and (oppressive) social narratives.

The research areas given as examples were taken from range of tech company "explainers" and conference calls for papers, in order to build a broader narrative of what constitutes CV. As with other aspects, this is not meant to be exhaustive, but offers a range of possible elements: classification; identification; verification; segmentation; detection; recognition; retrieval; synthesis. We might equally have divided the list into, for example, separate strands of 2D/3D object recognition, landmark recognition, pose recognition, biometrics recognition, and so on. Like other areas of the map, at the level of the field this is meant as an example starting point for viewing how different influences and narratives interact. Therefore these are positioned between the narrative and decision-making layer, in part because the division into sub-fields is itself a decision that will impact on how the field develops. Other related areas - for example the research career pathway and relative career levels within research organisations, collaborations and project teams - would be iterated across the sub-fields. A condensing of the relations to types of relations is therefore necessary for clarity within a (relatively) simple two-dimensional visual map designed for print. More complex relations or detail could be integrated for visuals aimed at a larger scale (poster or art installation) or interactive (web-based) maps.

Key points (again, non-exhaustive) of possible intervention are identified by smaller red dots. These could be places where fairness could be used, but in the majority of cases fairness will not be enough. Fairness can be one tool in these cases, but aspects of justice and equity are also necessary. This is particularly the case in the narrative layer - and specifically the overarching contexts and agendas - where the redistribution of power and decentring of existing mainstream narratives is required.

4. CONCLUSION

This paper has outlined a method of mapping CV as a relational and sociotechnical assemblage in order to locate areas of unfairness beyond the algorithm or the dataset, and thereby contribute to the increasing range of tools for addressing structural unfairness in CV and AI more widely. One of the main contributions of this paper is integrating the decision-making, influence and narrative components that constitute the broader field and define the available, supported or excluded directions of research. The map shows the need for *fairness outside the development of specific algorithms*, and some of the locations where this is needed, particularly in disciplinary, narrative and structural inequities. By taking into account the broader sociotechnical assemblate, the map also highlights the *inadequacy of fairness alone*, particularly within the algorithms/data themselves, but also more widely. Fairness should be included, but only as one possible type of intervention. This

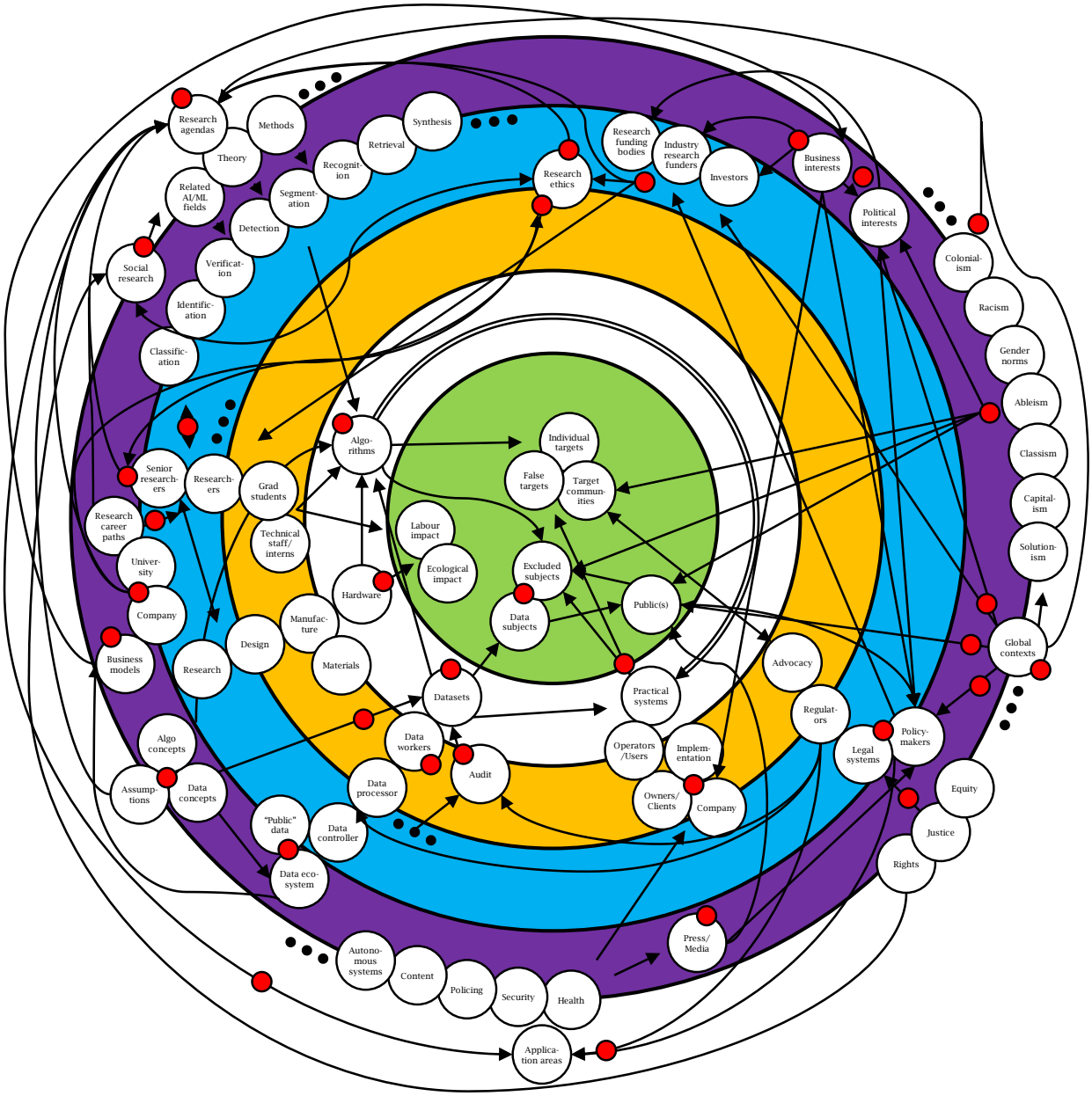

Figure 2. Mapping AI structure

shows the need for *measures beyond fairness* - such as justice, equity and redistributing power - as well as the structural systems and locations where these interventions are needed. In doing so it suggests the need for more narrative and performative approaches to mapping and intervening in a field such as CV. But any representation - visual or otherwise - has limitations. This is one of the fundamental problems with developing algorithms and using data in contexts with societal impact. As a representation, the present map is no exception. It is not an all-encompassing anatomy of the field - such a thing is impractical on a sheet of paper because of the sheer plurality of voices and experiences that need to be centred, but also impossible because it is always changing. The map as presented here is an open experimentation with locating sites for potential intervention not only in fairness but in equity, justice and the redistritbution of power across individual and collective, technical and social, practical and narrative contexts.

References

[1] J Khadijah Abdurahman. Fat* be wilin', Feb 2019. 1

[2] Garfield Benjamin. From protecting to performing privacy. *The Journal of Sociotechnical Critique*, 1(1):1, 2020. 2

[3] Garfield Benjamin. Put it in the bin: Mapping ai as a framework of refusal. In *ResistanceAI, NeurIPS2020*, 2020. 1

[4] Ruha Benjamin. *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity, London, 2019. 1

[5] Abeba Birhane. Algorithmic injustice: a relational ethics approach. *Patterns*, 2(2):9, 2021. 2

[6] Abeba Birhane and Olivia Guest. Towards decolonising computational sciences. *Women, Gender and Research*, 2020(2):60–73, 2020. 1

[7] Joy Buolamwini and Timnit Gebru. Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency*, pages 77–91. PMLR, 2018. 1

[8] Marika Cifor, Patricia Garcia, TL Cowan, Jasmine Rault, Tonia Sutherland, A Chan, Jennifer Rode, Anna Lauren Hoffmann, Niloufar Salehi, and Lisa Nakamura. Feminist data manifest-no. *Cit. on*, page 119, 2019. 2

[9] Kimberlé Crenshaw. Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Review*, 43(6):1241–1300, 1991. 2

[10] Gilles Deleuze and Félix Guattari. *A Thousand Plateaus: Capitalism and Schizophrenia*. Continuum, London, 2004. 1

[11] Catherine D'Ignazio and Lauren Klein. *Data Feminism*. MIT Press, Cambridge, 2020. 2

[12] Aristea Fotopoulou. Understanding citizen data practices from a feminist perspective: embodiment and the ethics of care. 2019. 2

[13] Alex Hanna, Emily Denton, Andrew Smart, and Jamila Smith-Loud. Towards a critical race methodology in algorithmic fairness. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*, pages 501–512, 2020. 1

[14] Anna Lauren Hoffmann. Data, technology, and gender: Thinking about (and from) trans lives. *Spaces for the Future. Routledge*, pages 15–25, 2017. 2

[15] Anna Lauren Hoffmann. Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse. *Information, Communication & Society*, 22(7):900–915, 2019. 1

[16] bell hooks. *Feminist theory: from margin to centre*. South End Press, Boston, 1984. 2

[17] Kalpana Kannabiran. A cartography of resistance: the national federation of dalit women. *The Situated Politics of Belonging, London: Sage*, pages 54–74, 2006. 2

[18] Os Keyes. The misgendering machines: Trans/hci implications of automatic gender recognition. *Proceedings of the ACM on human-computer interaction*, 2(CSCW):1–22, 2018. 1

[19] Katherine McKittrick. *Demonic Grounds: Black Women and the Cartographies of Struggle*. University of Minnesota Press, Minneapolis, 2006. 1

[20] Safiya Noble. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press, New York, 2018. 1

[21] Andrew D Selbst, Danah Boyd, Sorelle A Friedler, Suresh Venkatasubramanian, and Janet Vertesi. Fairness and abstraction in sociotechnical systems. In *Proceedings of the conference on fairness, accountability, and transparency*, pages 59–68, 2019. 1

[22] Luke Stark and Jesse Hoey. The ethics of emotion in artificial intelligence systems. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pages 782–793, 2021. 1

[23] Ngũgĩ wa Thiong'o. *Moving the Centre: The struggle for cultural freedoms*. James Currey, Melton, 1993. 2