
Where are the values? locating and reasoning 25 years on

Alan Dix

HCI Centre
School of Computer Science
University of Birmingham
Birmingham, B15 2TT, UK
and
Talis
Birmingham, B1 3HN, UK

alan@hciebook.com

<http://alandix.com/academic/papers/chi2017-values/>

**Abstract**

The author's work in the early 1990s identified issues of privacy and bias in complex data processing and black-box machine learning algorithms. Although some of these have taken time to emerge, they are now major societal issues. This early work has direct lessons, but also suggests that we need to look widely in the way computer systems are designed, produced and deployed to understand the way values become embedded in code. In addition we need new ways to reason about values, both in code and society at large.

Author Keywords

Privacy, discrimination, machine learning, social values.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous; K.4.2 Social Issues

Introduction

This paper starts by revisiting some largely theoretical work dating back to the early 1990s, before looking at the different parts of the software development process where values become embedded, and then finally at the challenges in reasoning about values in code and elsewhere in a turbulent 'post-truth' world.

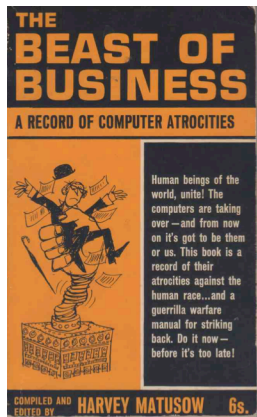
This paper is published under Creative Commons CC-BY, with the exception of the cover image of 'The Beast of Business'. You are free to share or modify, even commercially so long as the source is cited or attributed appropriately.

For full licence details see:

<https://creativecommons.org/licenses/by/4.0/>

<https://creativecommons.org/licenses/by/4.0/legalcode>





Even 50 years ago Harvey Matusow worried about the creeping influence of computers on everyday rights, founding the *International Society for the Abolition of Data Processing Machines*, which despite only having 1,500 members was reported in Time Magazine (*Frustrations: Guerrilla War Against Computers*, Friday, Sept. 12, 1969

The Beast of Business (pub. Wolfe, 1968) was his manifesto urging followers to demagnetize their cheques and to:

- WORRY a computer
- CONFUSE a computer
- WRECK a computer

Revisitation

Information processing, context and privacy [6] was the first paper in the HCI literature on privacy issues. Since 1990 there has been substantial work on privacy in HCI, including early work by Anne Adams [1] and Victoria Bellotti [2] as well as much technical work in the Ubicomp and Mobile HCI literature. However, there were two aspects of it that are still particularly relevant.

First, privacy was seen in the context of the information processing pipeline; that is not an add-on: merely an issue of anonymisation at the start, or access control at the end, but intimately interwoven with all the stages of filtering, selection and algorithmic processing. This is also likely to be true of other forms of values embedded in computational artefacts.

Second, privacy was not seen solely in terms of limiting access to certain data, but instead as the value information has to a person. Crucially this often means that retaining context and purpose is central – indeed the latter is well recognised in UK and European data protection legislation. In turn this means that losing information may make what is left more valuable or more damaging to an individual.

This less may be more even applies to fully anonymised and impersonal data. A family used to live by a busy four-lane main road with the children's school the other side. Sometimes they would cross by foot, but as the closest pedestrian crossing was more than half a mile away, they would take their lives in their hands and dash for it! Other times they would drive a mile down the road to the nearest U-turn point to drive back the other side of the road to the school and then home again. Suppose the local council placed those rubber

strips that count traffic; the trip back and forth to school creates four impressions. The council would then see how busy the road is and maybe reduce the number of pedestrian crossings to improve traffic flow, the exact opposite effect the family would have wanted.

Human issues in the use of pattern recognition techniques [7] was written following an early (1991) workshop [10] on the connections between human-computer interaction and the growing use of neural nets, machine learning and other AI pattern recognition techniques. The paper had a number of relevant features for the current discussion.

First, this was an early, maybe the first, paper to focus on the potential for gender and racial bias in black-box algorithms. Happily it has taken some time for this danger to be realised, but, of course, this has now become a major issue with, for example, suggestions of racist search suggestions in Google, or unethical surge pricing with Uber.

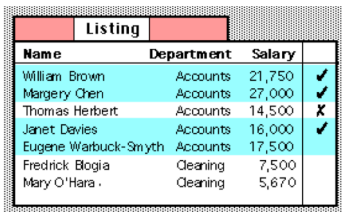
Second, bias can emerge even if there is no explicit storage of gender or ethnicity due to correlating factors not relevant for the particular purpose (e.g. job selection). That is, the choice not to use a discriminatory feature is a societal value decision independent of whether it has predictive power.

Third, where black-box algorithms (such as neural nets) are used, this may mean that it is hard or impossible to detect whether the algorithm is discriminatory, or indeed to prove it is not if challenged. This issue has now been recognised in the General Data Protection Regulation of the Council of the European Union, which mandates that algorithms

Query-by-Browsing exploring and exposing machine-learning bias

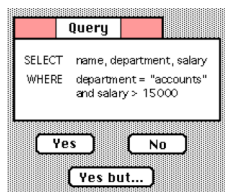
Query-by-Browsing (QbB) was initially designed as a thought experiment [7] and later implemented [8] in order to explore how AI learning algorithms could be applied to database querying, and also how this could then be exposed to user intervention.

First the user selects desired records from a listing.



Name	Department	Salary	
William Brown	Accounts	21,750	✓
Margery Chen	Accounts	27,000	✓
Thomas Herbert	Accounts	14,500	✗
Janet Davies	Accounts	16,000	✓
Eugene Warbuck-Smyth	Accounts	17,500	
Fredrick Bogia	Cleaning	7,500	
Mary O'Hara	Cleaning	5,670	

QbB then uses machine learning to infer a query and presents this to the user



```
SELECT name, department, salary
WHERE department = "accounts"
and salary > 15000
```

Yes No Yes but...

The user can examine this and iteratively modify the query or selected records

making certain critical decisions must to be able to explain their results [4, 11]

Finally, the paper made suggestions on ways to deal with some of the above problems including potential algorithmic ways to detect bias (e.g. inferring causal chains) and Query-by-Browsing, an experimental intelligent database UI that makes rules inferred by black-box algorithms visible to the user (see box).

Location

Values and assumptions about where, how, or by whom software can be used may be obvious, for example in the wording used in an interface, but may also be deeply embedded. For example TCP/IP slow-start gradually increases the rate of transmission until packet loss suggests that it has reached the limits of the channel. If packet loss further increases, TCP/IP backs off to slower speeds and begins again. With its roots in DARPA and the Cold War, this Internet protocol works well for the relatively infrequent, albeit massive, changes in network configuration caused by nuclear war, but performs very badly in rural areas where 'glitchy' connections frequently drop for a few seconds.

Values and assumptions can be found at many levels:

in deep code – for example, the TCP/IP slow-start

in the UI/UX – for example, assumptions that a user has a well-defined address, locking out the homeless

in product user impact – the 'bits leak out', TCP/IP slow-start is deeply buried, but has a crucial impact on delays and performance in rural areas

within the development/research team –

developers think "just a pretty screen", interaction designers think "just an implementation detail"

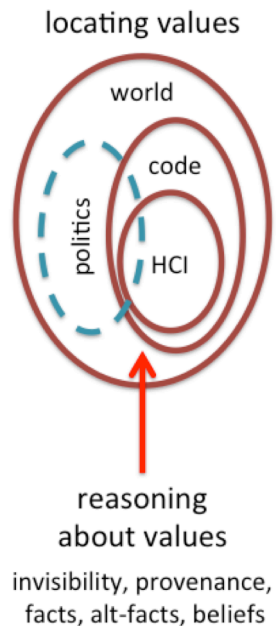
in the early design process – who to talk to, deep or superficial community engagement

in the development process – what equipment to use, agile vs long-term planning based methods

Each of these could be expanded, but we will focus briefly on the last as some of the impact may be less obvious.

Agile methods are widely adopted not least due to the way in which they rapidly react to user feedback. This is especially important in participatory or co-design practices and now the advent of affordable digital fabrication and electronic prototyping (maker-culture), have expanded the remit to physical design. However, Agile methods also implicitly decide what is mutable: typically features that are relatively small scale and localised, captured in a single user story. Broader issues can easily get missed or deprioritised due to complexity. At Talis a large portion of development effort is dedicated to infrastructure development, mainly micro-services, which address core cross-cutting concerns such as security, robustness and scalability: key values for university clients.

The equipment used during development is also crucial. In 1987 (30 years!) *The myth of the infinitely fast machine* [5] explored issues of time and delays in the user interface. Its focus was primarily on formal modelling, but also suggested that for one day a week developers should use a 2-year-old mid-range machine – now-a-days I would also add a degraded network.



Reasoning

Could suitably designed (HCI issue) tools help?

detection – The hardest thing about values, whether in software or society, is even seeing they are there; they are, by definition, invisible to us, often exposed only when we meet those from alternative cultures or belief systems: evident in recent politics, but also in debates about notification-based systems. Existing websites bring together left/right wing news items on the same events, maybe text algorithms could swap key aspects of content (e.g. race of participants). Neither is 'objective' value free; 'correct' Bayesian reasoning can create illegal discriminatory rules.

argumentation – It is maybe easier to track the role of values once recognised. Argumentation systems such as IBIS are well established [3] and in a recent overview of formal methods in HCI [9], the incorporation of values and similar qualities was identified as a key challenge for the area. The reasoning needed will inevitable mix qualitative and quantitative factors; semi-numerical non-expert tools, such iVoLVER [12], will be essential.

References

- [AS01] A. Adams, A. Sasse. 2001. Privacy in multimedia communications: Protecting users, not just data. *People and Computers XV*. 49–64
- [BS93] V. Bellotti, A. Sellen. 1993. Design for privacy in ubiquitous computing environments. *Proc. ECSCW'93 European Conference on Computer-Supported Cooperative Work*, 77-92. Dordrecht, Kluwer Academic Publishers.
- [BS06] S. Buckingham Shum, et al. (2006). Hypermedia support for argumentation-based rationale: 15 years on from gIBIS and QOC. In:

Rationale Management in Software Engineering. Berlin: Springer-Verlag, 111–132.

- [CE16] Council of the European Union. 2016. Position of the council on general data protection regulation. 8 April 2016.
- [Dx87] A. Dix. 1987. The myth of the infinitely fast machine. In *Proceedings of HCI'87*, D. Diaper & R. Winder (eds). Cambridge Univ. Press. 215–228. <http://alandix.com/academic/papers/hci87/>
- [Dx90] A. Dix. 1990. Information processing, context and privacy. In *Proc. INTERACT'90*. 15–20. <http://alandix.com/academic/papers/int90/>
- [Dx92] A. Dix. 1992. Human issues in the use of pattern recognition techniques. in *Neural Networks and Pattern Recognition in Human-Computer Interaction*. Ellis Horwood, 429–451. <http://alandix.com/academic/papers/neuro92/>
- [DP94] A. Dix, A. Patrick. 1994. Query By Browsing. In *Proc. of IDS'94*. Springer. 236–248. <http://alandix.com/academic/papers/QbB-IDS94/>
- [DW17] A. Dix, B. Weyers, J. Bowen, P. Palanque. 2017. Trends and Gaps. Chapter 3 in 7 in *The Handbook of Formal Methods in Human-Computer Interaction*. B. Weyers, et al. (eds.), Springer.
- [FB93] J. Finlay, R. Beale. 1993. Neural networks and pattern recognition in human-computer interaction. *SIGCHI Bulletin* 25, 2, 25–35. DOI: 10.1145/155804.155813
- [GF16] B. Goodman, S. Flaxman. 2016. EU regulations on algorithmic decision-making and a "right to explanation". Presented at *2016 ICML Workshop on Human Interpretability in Machine Learning*. <http://arxiv.org/abs/1606.08813v1>
- [MN16] G. Méndez, M. Nacenta, and Sebastien Vandenhete. 2016. iVoLVER: Interactive Visual Language for Visualization Extraction and Reconstruction. In *Proc. CHI '16*. 4073–4085. DOI: 10.1145/2858036.2858435