12-7-2022

Ways of Seeing: a Sensemaking Model for Sociotechnical Systems

Catherine Thompson
The University of Melbourne, ctthompso@student.unimelb.edu.au

Sherah Kurnia
The University of Melbourne, sherahk@unimelb.edu.au

Daniel Samson
The University of Melbourne, d.samson@unimelb.edu.au

Ways of Seeing: A Sensemaking Model for Sociotechnical Systems

Research-in-progress

Catherine Thompson
School of Computing and Information Systems
The University of Melbourne
Melbourne, Victoria
Australia
Email: ctthompso@student.unimelb.edu.au

Sherah Kurnia
School of Computing and Information Systems
The University of Melbourne
Melbourne, Victoria
Australia
Email: sherahk@unimelb.edu.au

Danny Samson
Department of Management and Marketing
The University of Melbourne
Melbourne, Victoria
Australia
Email: d.samson@unimelb.edu.au

Abstract

Sociotechnical systems are complex, layered and rich with relationships and interactions. It can be challenging to frame such systems in ways that allow both practical insights and access to deeper contexts and underlying beliefs. This paper proposes a simple and progressive sensemaking model for understanding these layered narratives, and uses it as a prism through which to explore the controversial Australian social welfare benefits repayment initiative known as Robodebt.

Keywords Sociotechnical systems, sensemaking, framing, complex, Robodebt

1 Introduction and Context

Digital technologies are portrayed positively as enablers of business, agents of personal connection, and tools of change. They are also portrayed as dehumanising, arbitrary, exclusionary and alienating. However, these technologies are created by humans and have no agency of their own: which is not to say that they are neutral. How such technologies are conceived and understood will in part depend on the broader sociotechnical system(s) of which they form part.

The study of these sociotechnical systems is not straightforward. Open and boundaryless, they are difficult to frame and invite a multitude of narratives. As emergent forms, only the study of the system itself will inform enquirers about its nature. Yet, these systems need to be understood, either to learn from their failure, or to support their continued smooth operation. In the words of Emery, one of sociotechnology's original researchers, "The more we know about these systems, the more we can identify what is relevant to a particular problem and detect problems missed by the conventional framework of problem analysis" (Emery 1993, p. 157).

Guided by the question: "How might we better understand the nature of complex sociotechnical systems?", this paper proposes a simple sensemaking (Weick 1995) model and uses it to explore a contemporary example of sociotechnical failure.

2 Literature: Understanding Complex Sociotechnical Systems

Sociotechnical systems were first described in the years after World War II by Tavistock Institute researchers seeking to understand how the interactions between mine workers and the technologies at their disposal shaped human work relationships, practices and structures (Trist 1981). Initially studied in team-based industrial settings, sociotechnical systems were soon taken to encompass the organisation as a whole, and then its "macrosocial" network of external relationships (Trist 1981, p. 50).

Technologies in general were seen as key contributors to the ability of sociotechnical systems to self-regulate and to maintain their ability to function under a variety of conditions (homeostasis). This importance stems from their role in managing the porous boundary conditions that exist at the edge of the enterprise and its external environment (Emery 1993). Digital technologies, predicted by the early researchers, would pervade and then profoundly reshape sociotechnical relationships (Trist 1981).

These digital technologies were – and often still are – characterised as complicated rather than complex systems. Complicated systems, unlike complex ones, are said to be fully describable (Cilliers 2002, p. 14). However, it does not follow that a technology that can be enumerated through its features, functions is merely complicated. Technologies are conceived in environments that shape their design, purpose and subsequent operation, and they in turn influence their environments. This interdependence is the very essence of sociotechnology, and is also a feature of complex systems (Cilliers 2002). Technology artefacts, and therefore the sociotechnical systems of which they form part, are neither value free (Pacey 2014) nor neutral (Kranzberg 1986).

Causal reasoning and techniques like root cause analysis are of limited assistance in understanding how these complex sociotechnical systems fail (Dekker 2011) or, conversely, how they continue to function. Multiple narratives are possible, shaped by the perspectives of participants and researchers. These narratives are necessarily partial, often arising from our imperfect knowledge of the frame (our bounded rationality) that restricts our attempts at sensemaking.

3 Framing Sociotechnical Systems

Sociotechnical systems are complex, layered and rich with relationships and interactions. It can be challenging to frame these systems in ways that allow both practical insights and access to deeper contexts and underlying beliefs.

The researchers who first described sociotechnical systems explicitly understood them to be a way of seeing and a sensemaking tool: "The first function of a socio-technical systems concept is as a frame of reference—a general way of ordering the facts" (Emery 1993, p. 159). In the spirit of Emery's work, this research aims to create a simple framing through which disparate perspectives can cohere into a continuous narrative that incorporates them all. Our approach is informed by the work of complexity theorist Paul Cilliers, who observed that: "The distinction between complex and simple often becomes a function of our 'distance' from the system (Serra and Zanarini 1990: 4, 5), i.e. of the kind of description of the system we are using" (Cilliers 2002, p. 3).

We propose a simple three-part model (Figure 1) that progressively widens the framing of the system both outwards and through time, so that practical, operational deficiencies can be addressed while also holding more complex sociotechnical challenges in mind. The model establishes an anchoring narrative whose purpose is to create a structure from which to explore the full richness of a complex, sociotechnical, digital ecosystem.

Figure 1: A Sensemaking Model for Sociotechnical Systems

The initial frame calls for understanding a digital sociotechnical system as a technology artefact or **product**. A broader second frame considers the context and environment in which the product comes into being, or the **purpose** behind its design. The third and broadest frame considers the **philosophy** of the system, being the beliefs or doctrinal underpinnings that have contributed to shaping the purpose.

4 The Online Compliance Intervention ("Robodebt")

The Online Compliance Intervention (OCI) initiative, widely known as Robodebt, was an Australian federal government benefits overpayment recovery programme that operated between 2016 and 2019. Condemned as an illustration of creeping "digital welfare dystopia" by the United Nations (UN Secretary General 2019, pp. 9, 21), the scheme sought to identify current and former welfare recipients who might have received benefits in excess of their entitlements, and to extract repayment from them.

Initially, the OCI relied on human operatives to calculate and evaluate potential overpayments and to decide on the appropriate action. However, from mid-2016, algorithms were developed to perform the assessments, and repayment demand notice production was fully automated. This automation increased the volume of assessments from some 20,000 per year to 20,000 per week.

The burden of disproving the validity of a demand notice fell to the benefit recipients, many of whom were unable to produce the records required to substantiate their defence. There was no obvious means of appealing the assessment in the absence of these records. Repayment demands were recharacterised as real "debts" owed to the state. A variety of methods were used to recover these debts, including garnishing benefits and tax refunds, and employing the services of debt collection agencies.

At the heart of the Robodebt initiative was the "income averaging" calculation used to identify overpayments. This calculation smoothed the incomes of benefit claimants who were in irregular or cyclical employment, and used the overstated income figures in the calculation of overpayments. This methodology was deemed unlawful by the Australian Federal Court in November 2019, a judgement that signalled Robodebt's death knell (Karp 2019). A class action swiftly followed and in June 2021 this too was settled in favour of the plaintiffs. In total, some AUD $1.7B has been earmarked to flow back to Robodebt's victims through a combination of refunds, debt waivers and compensation. In the 3.5 years of its operation, the programme affected ~470,000 individuals and caused innumerable hardships, reportedly including suicides (Whiteford 2021).

5 Framing Robodebt

Researchers have contributed thoughtful Robodebt analysis and commentary from the domains of policy, technology, and the law. Whatever their specialisation, these accounts identify some common failings including algorithmic bias (Akter et al. 2021; Carney 2018); lack of regulatory oversight (Carney 2018; Whiteford 2021); and indifference to the plight of disadvantaged populations (Akter et al. 2021; Carney 2018).

The interwoven perspectives of these and other researchers provide a sense of the rich interdependencies and sprawling span of this sociotechnical system – a span which is largely defined by the gaze of the individual researcher. How then might we engage with such a system to study it? And if narratives are, as Weick (1995) maintains, "a plausible frame for sensemaking" (p. 128), how might we

understand the overlapping narratives around Robodebt? Our sensemaking model was used to create a structured narrative account in which multiple perspectives can be represented and questioned.

5.1 Frame 1: Product

The initial frame concerns the design and operation of the technology artefact, and is the easiest to construct. The technology is straightforward to describe and its flaws are evident. Official government documentation, from two senate enquiries and a Commonwealth Ombudsman investigation, provides a quantitative factbase of harms, illustrated with the personal stories of victims. In this narrow framing, Robodebt presents as bad technology. It reads as a poorly thought-out and executed business process automation (BPA) initiative that optimised for operational scale and efficiency (i.e. the vastly accelerated volume of assessments) at the expense of quality and accuracy (i.e. the income averaging calculation).

In this frame, the failure of the OCI can be explained as a series of transgressions against ethics and good technology practices. Witness the BPA implementation that entirely removed the human from the loop; and the poor data governance that resulted in incorrectly addressed demand notices that never reached their recipients. The four principles of the European Union's *Ethics Guidelines for Trustworthy AI* (2019) are representative of the contents of a typical AI framework. The first principle emphasises respect for human autonomy, including the human oversight of AI systems, and asks that these systems "not unjustifiably subordinate, coerce, deceive, manipulate, condition or herd humans" (p. 12). The three remaining principles address the prevention of harms and the requirement for both fairness and explicability. It can be strongly argued that adherence to guidelines such as these would have prevented Robodebt.

The judicial process pointed to the flawed income averaging calculation as the proximal cause of the Robodebt failure, because the "debts" that it created were illusory and illegitimate. This process also ruled that the design of the system contravened administrative law by requiring recipients of demand notices to prove that their debt was invalid. Stop at this point, and middle management is to blame for failures of professional ethics and practice. This perspective underlines that "no one in government was fired for designing and operating an unlawful system for years on end" (Warren 2022, p. 34).

This first framing leads us to a causally satisfying narrative of management incompetence. It highlights opportunities for better practices informed and constrained by AI and data ethics guidance, as well as compliant to the rule of law. This is not to say that the practical turn of Frame 1 encourages a superficial approach. On the contrary, central to Frame 1 thinking are questions of whether there is good technology and bad technology, the preventability of harms, the special circumstances of no-optionality government services, and what our digital governance and regulatory choices should look like.

5.2 Frame 2: Purpose

But what of the environment in which Robodebt was proposed, tested, piloted, implemented and heralded as a success? The certainty of an identified root cause failure in ethics falls away when the frame is widened to include the context in which relevant decisions were made.

Sometimes this context includes a declarative mission statement, such as John F Kennedy's affirmation that "we choose to go to the Moon in this decade" (Kennedy 1962). Robodebt was founded on such a statement. The FY15/16 federal budget included a provision to recover $1.7B in benefits overpayments over a period of 4 years. In the context of Frame 2, the ethical failures appear not as causal, but as trace indicators of another narrative. Robodebt is no longer *what* happened: it becomes *how* it happened.

In this light, Robodebt, with all its flaws, appears purposeful. The public servants accountable for the target mustered the tools at their disposal to deliver the outcomes desired by the government of the day. Their behaviours were goal-directed or "characterized by choosing actions according to the outcomes they produce in a certain situation" (Zwosta 2015, p. 1). Their scope for action was significantly constrained by the limited target population and the large quantum of funds to be delivered. In this broader framing, the imposition of a 10% debt recovery fee, the intimidating wording of the demand letters and the difficulty in contesting an alleged debt can all be construed not as bugs (i.e. the professional and ethics failures of the "bad technology" framing) but deliberatively, as features.

This framing entertains the possibility that, even if Frame 1 practices had been better and had avoided Robodebt, any sociotechnical system constructed around the same declarative mission could foreseeably have produced an equivalent outcome.

5.3 Frame 3: Philosophy

The third frame deepens and widens to consider the underpinning, perhaps unspoken, belief structures shaping the system. Using retroduction (Jagosh 2021), we can ask: "What would have to be the case in order to believe that AUD$1.7B could be unlocked from welfare overpayments?". Robodebt researchers have pointed to a doctrinal component in the establishment of the FY15/16 budget measure. They reference an underlying sentiment that welfare recipients incur "debts" to society and are also inclined to claim benefits to which they are not entitled (Akter et al. 2021; Carney 2018; Whiteford 2021).

Counterfactual thinking (Byrne 2017) may be helpful in appraising the level of doctrinal influence in a sociotechnical system. In this case, if overpayments had been made, it follows logically that underpayments likely also occurred. The budget measure would therefore have been ethically balanced if, and only if, underpayments had been factored into any calculations and symmetrical efforts were made to detect and rectify both. This broadest framing speaks to the exercise of power and to its legitimacy.

Each of the sensemaking model's three frames has a narrative that is satisfying and complete within its own boundaries. All three frames – and perhaps others – validly coexist. Figure 2 summarises how the model can develop simple yet generative narratives that encourage future enquiry.

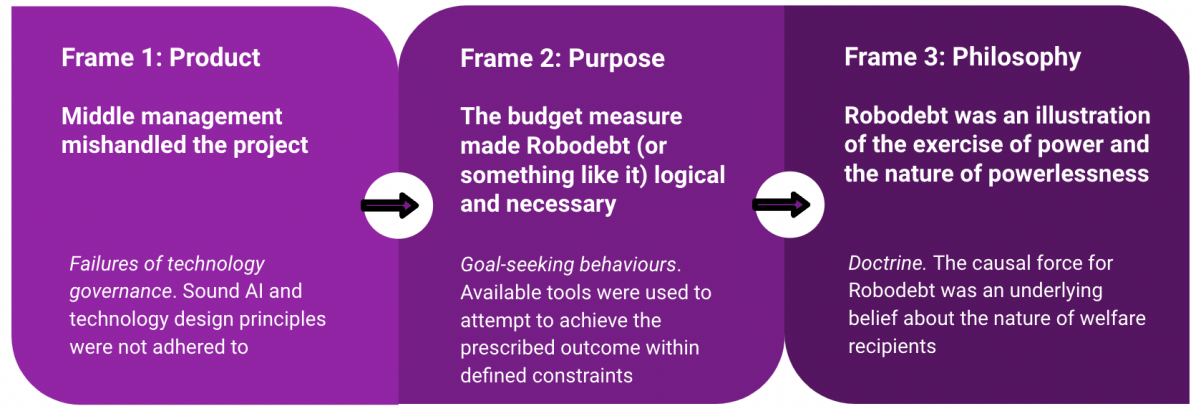

Figure 2: Framing Robodebt: Worked Example of a Sensemaking Model for Sociotechnical Systems

6 Conclusion and Future Research

One of the first challenges faced by a researcher seeking to study a complex sociotechnical system is that these systems lack defined boundaries and can be understood in very different ways depending on the frame of reference that is applied. Our suggested sensemaking model is a lightweight framework intended to help researchers swiftly structure framing decisions. It is an attempt to allow the many narratives that may surface in such a system to legitimately coexist, and thus to allow for work at different levels, from practical technology improvements to the exploration of deep causal forces.

This sensemaking model forms part of a broader research agenda into trustworthiness in the digital world and will be used in our continuing Robodebt case research and further planned case studies. These case studies will enquire into the roles and interactions that may contribute to developing, sustaining, and governing digital ecosystems that are safer for all to navigate.

We suggest that practitioners, especially those investigating systems failures in complex sociotechnical settings, might use this model to structure the initial stages of their enquiry. We hope that the model may also prove useful to researchers whose interests lie in the stratified narratives of sociotechnical systems, or alternatively to those seeking the simplicity of a high-level framework.

7 References

Akter, S., McCarthy, G., Sajib, S., Michael, K., Dwivedi, Y. K., D'Ambra, J., & Shen, K. N. (2021). Algorithmic bias in data-driven innovation in the age of AI. *International Journal of Information Management*, *60*, 102387.

Byrne, R. M. (2017). Counterfactual thinking: From logic to morality. *Current Directions in Psychological Science*, *26*(4), 314-322.

Carney, T. (2019). Robo-debt illegality: The seven veils of failed guarantees of the rule of law?. *Alternative Law Journal*, *44*(1), 4-10.

Cilliers, P. (2002). *Complexity and postmodernism: Understanding complex systems*. Routledge.

Dekker, S. (2011). *Drift Into Failure: From Hunting Broken Components to Understanding Complex Systems*. Surrey, UK: Ashgate Publishers.

Emery, F. (1993). Characteristics of Socio-technical Systems. In E. Trist, H. Murray, & B. Trist (Eds), *The Social Engagement of Social Science, Volume 2: A Tavistock Anthology--The Socio-technical Perspective* (pp. 157–186). University of Pennsylvania Press.

European Commission, (2019). *Ethics Guidelines for Trustworthy AI*. Publication Office. https://data.europa.eu/doi/10.2759/346720

Jagosh, J. (2020). Retroductive theorizing in Pawson and Tilley's applied scientific realism. *Journal of Critical Realism*, *19*(2), 121-130.

Karp, P (2019). Government admits Robodebt was unlawful as it settles legal challenge. *The Guardian*. https://www.theguardian.com/australia-news/2019/nov/27/government-admits-robodebt-was-unlawful-as-it-settles-legal-challenge

Kennedy, J. F. (1962, September 12). Address at Rice University on the Nation's Space Effort. John F. Kennedy Presidential Library and Museum. https://www.jfklibrary.org/learn/about-jfk/historic-speeches/address-at-rice-university-on-the-nations-space-effort

Kranzberg, M. (1986). Technology and History:" Kranzberg's Laws". *Technology and culture*, *27*(3), 544-560.

Pacey, A. (2014). Technology: practice and culture. In *Ethics and Emerging Technologies* (pp. 27-36). Palgrave Macmillan, London.

Trist, E. L. (1981). *The evolution of socio-technical systems* (Vol. 2). Toronto: Ontario Quality of Working Life Centre.

UN Secretary General. (2019). *Report of the Special Rapporteur on Extreme Poverty and Human Rights*. A/74/493. United Nations, Geneva.

Warren, J. (2022), Tech safety that includes the removal of harmful systems. *InnovationAus*

Weick, K. E. (1995). Sensemaking in organizations (Vol. 3). Sage.

Whiteford, P. (2021). Debt by design: The anatomy of a social policy fiasco–Or was it something worse?. *Australian Journal of Public Administration*, *80*(2), 340-360.

Zwosta, K., Ruge, H., & Wolfensteller, U. (2015). Neural mechanisms of goal-directed behavior: outcome-based response selection is associated with increased functional coupling of the angular gyrus. *Frontiers in human neuroscience*, *9*, 180.

Acknowledgements

We extend our thanks to the three anonymous reviewers for their insights; to Dr Lesley Seebeck for so generously sharing her expertise; and to the first author's colleagues for their continuing support.

Copyright