

Creative brief

Motivation

AI is increasingly used by public sector actors to support, augment and automate decision-making. Such systems lack democratic legitimacy. This can be improved by ensuring systems are *contestable*: Open and responsive to human intervention throughout their lifecycle, establishing a dialectical relationship between decision subjects and system operators.

Objective

The aim of this project is to create a **visual explanation** that enables professional interaction designers to create **concept designs for public AI** systems that are **contestable**. The contents of this visual explanation are derived from the *features* section of the “Contestable AI by Design”-framework (Figure 1) [1], and the “five contestability loops”-model (Figure 2) [2]. The envisioned use case of the visual explanation is that it serves as a source of guidance and inspiration for design practitioners in the early stages of design projects dealing with public AI systems. The visual explanation will be evaluated as part of a scientific study, using a half-day workshop in which designers are tasked with improving the contestability of a real-world public AI system that is presented to them by a representative of the municipal (city) government that owns and operates it.

Success criteria

The visual explanation should be...

1. **Effective, useful:** Supports the intended task. Concept designs created with it share properties with those described in the “Contestable AI by Design”-framework.
2. **Learnable:** Easy to understand.
3. **Efficient:** Quick to use.
4. **Complete, self-contained:** Contains all necessary information.
5. **Flexible:** Adaptable to individual designers’ preferred way of working.
6. **Generative:** Inspires novel ideas for future designs that do not simply mirror what is represented in the visual explanation or underlying framework. Has the correct level of abstraction.
7. **Delightful, attractive:** Pleasant to use and attractive to perceive.

Deliverables

The ultimate deliverable of the project is (at minimum) a single visual explanation that can be printed by designers in their studios, or easily viewed on a single display or projector while working. This implies a print size of up to A3, and a landscape orientation.

Note: It is unlikely that a single static A3-size image will be sufficient to convey all that we need to. This will have to be further determined at design time.

Content

The visual explanation should convey the following content:

- Actors:
 - Citizens (a.k.a. “decision subjects”)
 - Developers (perhaps distinguishing between internal and external ones)
 - Policy-makers (alderpersons, mayor, council members, ...)
 - AI system (data inputs, models, output predictions)
 - Civil servant (a.k.a. “human controller”)
 - “Third parties” (e.g., external oversight bodies)
- Contestability features:
 - Interactive controls – for civil servants
 - Explanations (*justifications*) – for citizens
 - Intervention requests (appeals) – channels for voice, arenas for debate, obligation to review/respond/reconsider decision
 - Tools for scrutiny – for citizens, third parties
 - (These are all taken from the framework. We leave out Built-In Safeguards because it is not central to contestations.)
- Loops:
 - Appeals (loop L1). These map onto intervention requests, above.
 - Monitoring of decision-appeals loops leading back to development (L4) and policy-making (L5). These *do* imply technical system features that are not captured by the original framework.
 - (We choose to leave out Participation in Development (L2) and in Policy-Making (L3) because these are practice-related.)

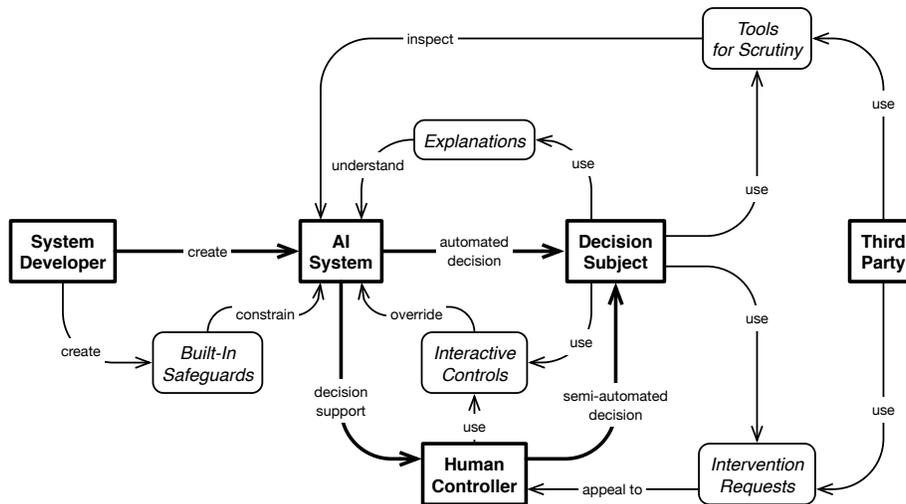


Figure 1: Features contributing to contestable AI: System developers create built-in safeguards to constrain the behavior of AI systems. Human controllers use interactive controls to correct or override AI system decisions. Decision subjects use interactive controls, explanations, intervention requests, and tools for scrutiny to contest AI system decisions. Third parties also use tools for scrutiny and intervention requests for oversight and contestation on behalf of individuals and groups.

Additional requirements

- Dynamism, temporality: The visual explanation should show how a system, under the influence of contestations, shifts from a present state to a future state.
- Context: We need one or more real-world examples, or conceptual metaphors [4], that we can use to visually communicate the abstract features and loops described above.

Appearance

Some example visual explanations that serve as starting points for visual design.

Waterwerk (Figure 3) by Carlijn Kingma is a visual explanation of the contemporary monetary system that uses water as its central conceptual metaphor. It will be necessary for us to deploy metaphor as well. Picking the right one will require some careful consideration.¹

Building Stories (Figure 4) by Chris Ware is strictly speaking not a visual explanation, but a (non-linear) comic. I like the way Ware mixes architecture and narrative in one image, and the reader can start anywhere. Relevant because

¹<https://www.waterworksofmoney.com>

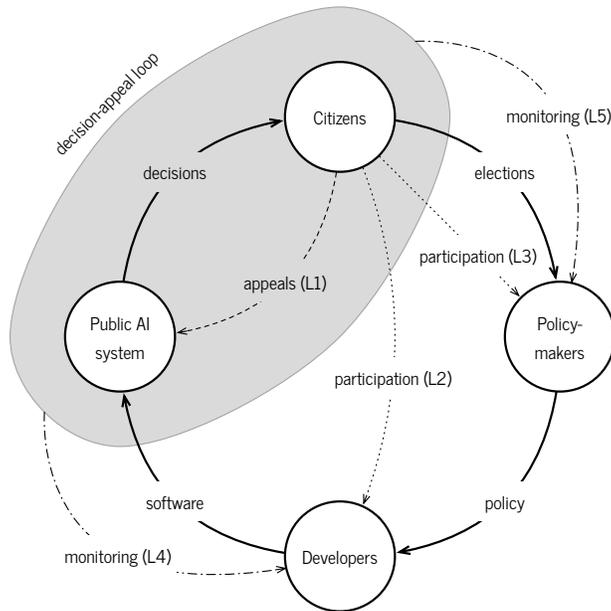


Figure 2: Diagram of our “five loops model,” showing the basic flow of policy through software into decisions (solid arrows), the direct way citizens can contest individual decisions (L1, dashed arrow), the direct ways in which citizens can contest systems development and policy making (L2–3, dotted arrows), and the second-order feedback loops leading from all decision-appeal interactions in the aggregate back to software development and policy-making (L4–5, dashed-dotted arrows).

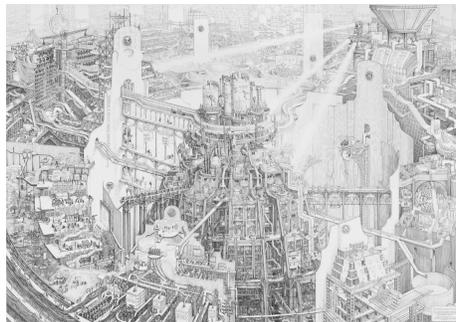


Figure 3: Waterwerk by Carlijn Kingma / Follow the Money.

we will be explaining interactions that happen between actors over time, that are in turn distributed in time and space [9].

The State of the Beaches (Figure 5) is one of many infographics by Megan Jaegerman that are described by Edward Tufte as “some of the best news graph-

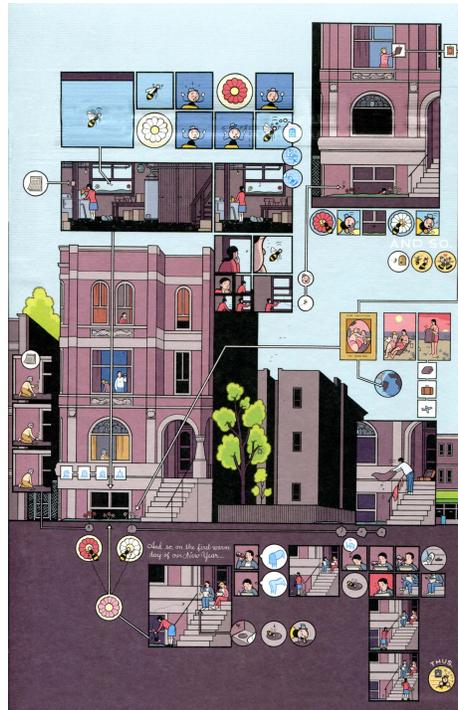


Figure 4: Building Stories by Chris Ware.

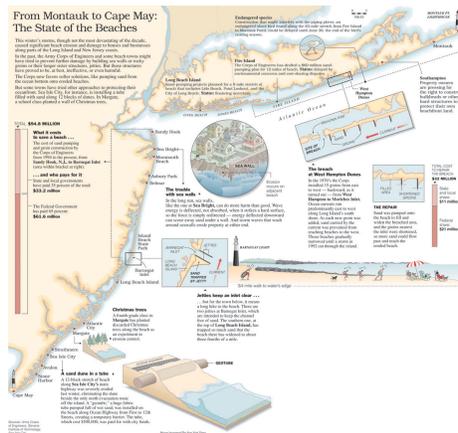


Figure 5: The State of the Beaches by Megan Jaegerman.

ics ever,”² because she mixes wit with elegance and informativeness. I like how this image has the coastline as a backbone and then various callouts that zoom

²https://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=0002w4

into particular aspects.

Definitions

- **Visual explanation:** “*Pictures of verbs*, the representation of mechanism and motion, of process and dynamics, of causes and effects, of explanation and narrative” [8]. In our case we use visual explanations as a form of intermediate-level design knowledge – i.e., somewhere between particular design instances, and general theory [3].
- **Design concept:** Portrayals of future designs [6]. As opposed to design *artifacts*.
- **Artificial intelligence (AI):** “[A] cover term for a range of techniques for data analysis and processing, the relevant parameters of which can be adjusted according to either internally or externally generated feedback” [7].
- **Public AI:** AI used by public sector actors for supporting, augmenting or automating decisions [5].
- **Contestability:** Open and responsive to human intervention, throughout the system lifecycle, establishing a dialectical relationship between decision subjects and system operators [1].

References

- [1] Kars Alfrink et al. “Contestable AI by Design: Towards a Framework”. en. In: *Minds and Machines* 33.4 (Aug. 2022), pp. 613–639. ISSN: 1572-8641. DOI: 10/gqnjcs. URL: <https://link.springer.com/10.1007/s11023-022-09611-z> (visited on 03/21/2024).
- [2] Kars Alfrink et al. “Contestable Camera Cars: A Speculative Design Exploration of Public AI That Is Open and Responsive to Dispute”. en. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 0 citations (Crossref) [2023-08-17]. Hamburg Germany: ACM, Apr. 2023, pp. 1–16. ISBN: 978-1-4503-9421-5. DOI: 10/gr5wcx. URL: <https://dl.acm.org/doi/10.1145/3544548.3580984> (visited on 06/07/2023).
- [3] Kristina Höök and Jonas Löwgren. “Strong Concepts: Intermediate-Level Knowledge in Interaction Design Research”. In: *ACM Transactions on Computer-Human Interaction* 19.3 (2012), pp. 1–18. ISSN: 10730516. DOI: 10/f225d4. URL: <http://dl.acm.org/citation.cfm?doid=2362364.2362371>.
- [4] George Lakoff. *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago: University of Chicago Press, 1987. ISBN: 978-0-226-46803-7.

- [5] Sem Nouws, Marijn Janssen, and Roel Dobbe. “Dismantling Digital Cages: Examining Design Practices for Public Algorithmic Systems”. In: *Electronic Government: 21st IFIP WG 8.5 International Conference, EGOV 2022, Linköping, Sweden, September 6–8, 2022, Proceedings*. 0 citations (Cross-ref) [2023-09-17]. Berlin, Heidelberg: Springer-Verlag, Sept. 2022, pp. 307–322. ISBN: 978-3-031-15085-2. DOI: 10/gsqvzf. URL: https://doi.org/10.1007/978-3-031-15086-9_20 (visited on 09/17/2023).
- [6] Erik Stolterman and Mikael Wiberg. “Concept-Driven Interaction Design Research”. en. In: *Human-Computer Interaction 25.2* (May 2010), pp. 95–118. ISSN: 0737-0024, 1532-7051. DOI: 10/c6qftv. URL: <https://www.tandfonline.com/doi/full/10.1080/07370020903586696> (visited on 05/18/2021).
- [7] Lucy Suchman. *Corporate Accountability*. June 2018. URL: <https://edu.nl/aqp4u> (visited on 08/12/2021).
- [8] Edward R. Tufte. *Visual explanations: images and quantities, evidence and narrative*. Cheshire, Conn: Graphics Press, 1997. ISBN: 978-0-9613921-2-3.
- [9] Chris Ware. *Building stories*. eng. New York: Pantheon Books, a division of Random House, Inc, 2012. ISBN: 978-0-375-42433-5.