Serials Review

ISSN: 0098-7913 (Print) 1879-095X (Online) Journal homepage: https://www.tandfonline.com/loi/usrv20

It's Made of People: Designing Systems for Humans

Kristen Wilson(Recorder) & Andreas Orphanides(Presenter)

To cite this article: Kristen Wilson(Recorder) & Andreas Orphanides(Presenter) (2018) It's Made of People: Designing Systems for Humans, Serials Review, 44:3, 211-216, DOI: 10.1080/00987913.2018.1525244

To link to this article: https://doi.org/10.1080/00987913.2018.1525244

SERIALS REVIEW
2018, VOL. 44, NO. 3, 211–216
https://doi.org/10.1080/00987913.2018.1525244

Routledge
Taylor & Francis Group

CONFERENCE REPORT

It's Made of People: Designing Systems for Humans

Kristen Wilson, Recorder[a] and Andreas Orphanides, Presenter[b]

[a]Project Manager/Business Analyst, Index Data, Boston, Massachusetts, USA; [b]Associate Head, User Experience, North Carolina State University Libraries, Raleigh, North Carolina, USA

ABSTRACT
Technical services staff, along with programmers, supervisors, and frontline librarians, participate in all sorts of systems. Whether they recognize it or not, they are used to interacting with the world through the lens of the systems they work with. In this presentation from the North Carolina Serials Conference, Andreas Orphanides looks at some of the challenges of interacting with the world in terms of systems, discusses the human costs of failing to recognize the limitations of systems, and provides a framework for thinking about systems to help ensure that our systems respect the humanity of their human participants.

KEYWORDS
design ethics; human-centered design; models; systems analysis

Everything we do is systems

Understanding systems is essential to the work of libraries. Nearly all of our work revolves around implementing systems, so being able to make sense of their workings is critical to achieving organizational goals. This imperative applies not just to coders and project managers but to technical services staff, supervisors, and frontline librarians—all of whom are participants in all sorts of systems. Understanding those systems can help everyone who works in libraries understand their constituents better.

A *system* is broadly defined as a group of interacting components within a particular boundary. We get to decide what is inside the system and what is not, but the system is only meaningful if the boundary defines a context that makes sense. Often, some components of the system are simplified to make it more conceptually manageable, and things that are outside of the system get ignored. The end result is a representation of part of the world that happens to capture our interest and can be considered in a more or less self-contained way.

In libraries, our focus is mainly on human systems, where humans endeavor to accomplish useful work. Humans play the role of actors in the system, but system actors are not necessarily humans. They may be, for example, databases, applications, and other automated components. The actors participate in workflows, moving the system into different states. They aim to get the system into particular states that we call goals. It is worth noting that some of these goals may conflict. Sometimes it makes sense to say that the system has its own goals as well. For example, the goal of a shipyard is to build ships, but it is unlikely that any single person would say their specific task is to create a ship (Gall, 2002). This is the basic structure of a human system: actors, workflows, and goals.

There is one important additional aspect of human systems that differentiates them from systems more generally: They are often dealing with information. This could be information about the actors, the workflow, or the goals—or information about the system as a whole. But since information does not have a physical presence, we often create a representation for our convenience. That representation is called a model, and it serves as a surrogate for information.

Because they are representations, models exclude, simplify, or abstract components that are not immediately relevant. They selectively embody our knowledge in a way that is easier to manage than the thing itself. A subway map is a good example of a model. It provides station names and adjacencies, transfer points, and the rough physical locations of the stations. But things like precise distances between stations are not useful to the model, so they are excluded. Useful models must be simplifications. A map of Chapel Hill, North Carolina, for example, would not be useful if it were the same size as Chapel Hill. A model can also help us better understand things that would be difficult to get a handle on. The map of Chapel Hill lets users make sense of spatial

CONTACT Kristen Wilson kristen@indexdata.com Index Data, 8 Faneuil Hall Marketplace, 3rd floor, Boston, MA 02109; Andreas Orphanides akorphan@ncsu.edu NCSU Libraries, 2 Broughton Dr., Box 7111, Raleigh, NC 27695.

relationships that would be hard to recognize up close. MARC records are another type of model that allows us to understand physical and conceptual information about books without having the books themselves physically present.

It is important to choose the right way to simplify a model. The statistician George Box (Box & Draper, 1987) said, “All models are wrong, but some are useful” (p. 424). Their utility comes from providing and omitting detail in a way that facilitates understanding rather than hinders it, given a particular need. The right kind of wrongness makes a model useful.

To review: A model is a strategic, simplified representation of something, designed to make it more conceptually manageable and easier to work with. A system is a strategic, simplified representation of reality designed to make reality more conceptually manageable. In other words, a system is a model of reality.

Understanding this fact—that a system is a model—is critical to ensuring that the systems we design do the work that we actually want them to do. Just like other models, systems are necessarily imperfect representations. And, as in models, the success or failure of our system hinges on how we draw boundaries, what we abstract, and what we ignore. Our understanding of the missions, goals, and structure of our organizations help to inform the system’s design, but the system is still, fundamentally, a simplification of reality. In this way, all systems are wrong. So how do we make sure they are useful?

The systems challenge

For those who work within a system, it is their model of reality and therefore an incomplete representation. But from the perspective of the system, where most library workers spend their time, the system represents their complete understanding about the state of reality for the purpose of achieving organizational goals. Every decision made from within the system is filtered through how it represents the world. This situation presents a paradox, perhaps the very central paradox in systems analysis and design: The system is (necessarily) wrong about reality, but the system is our source of truth about reality.

This gives rise to what systems analyst John Gall (2002) called the Fundamental Law of Administrative Workings: “Things are what they’re reported to be” (p. 45). That is, if we are relying on the system as the source of truth, and the system’s knowledge is wrong or incomplete, our own knowledge is wrong. But try telling that to a system that disagrees with you. A classic example of this is when the IRS erroneously decides that you are dead (Crockett, 2015). Good luck convincing them otherwise. Likewise, getting too close to a system—and forgetting its fundamental limitations—can lead people to fail those they are trying to serve. This danger becomes clear through some examples that highlight how shortcomings in defining the system lead to bad system output.

Framing the system

The first point of failure in a system is how it is scoped–it’s boundary. The framing of the system has a critical influence on how the system’s goals are defined. If you get the framing wrong, you get bad goals. For instance, “antihomeless spikes” are installed in protected alcoves of buildings to prevent indigent people from sheltering there. “Antihomeless spikes” clearly represent a failure of ethics in system design. They also represent a failure in system scoping. In system terms, the spikes achieve a goal—they prevent homeless people from bedding down in a spot. But they fail at the more fundamental goal of addressing the problem of homelessness. The system also excludes from consideration the notion that homeless people are actors in the system with agency, needs, motivations, and intrinsic value as humans. The system only considers them to the extent that their presence conflicts with the desire of the property owner.

In a similar example, journalist and media technologist Shane Snow (2015) wrote an article in which he proposed a comprehensive solution to the problem of prisons. Snow identified three key problems in the current U.S. prison system: the incidence of violence and antisocial behavior among prisoners, the high cost to the taxpayer of prison operations, and the general inhumanity of the prison system for both those who are incarcerated and its employees. The first element of the proposed solution is to put everyone in solitary confinement. Snow argues that this will be safer for prisoners and administratively simpler. However, since solitary confinement brings its own antisocial and inhumane aspects, the next part of the solution is to provide a virtual reality experience where prisoners can remotely interact with other prisoners, take classes, etc. This would safely provide the social outlet needed, give a perception of freedom, and offer an opportunity for self-improvement for prisoners. A final cost-saving component is to replace all prison food with Soylent, a liquid meal replacement. This would remove the overhead of food preparation, the need for a common dining area, and the risks associated with large groups of prisoners with access to potential materials for weapons. Snow estimates his proposal would cost $10,000 a year per prisoner, a savings of two to six times the current cost.

The issues with Snow's proposal are myriad and, hopefully, obvious. The key systems-analytic failure is, once again, that the system is framed to overlook the humanity of the human actors, with individuality, motivations, and desires—despite aspiring to address the inhumanity of prisons! The system as described provides for prisoners' needs, but only in the most superficial, reductionist way, rather than recognizing prisoners' needs and desires as intrinsic, fully realized aspects of actual human persons. Once again, this proposed solution attempts to solve the wrong problems. Snow's proposal never meaningfully examines the underlying causes of prison costs, considers what aspects of incarceration lead to prison violence, or investigates what it means for an experience to be inhumane. Without understanding the genesis of these problems with the system, we cannot meaningfully propose a solution. Any such attempt is just chasing down symptoms.

Conversely, framing a system correctly is critical to facilitating a working solution. In the late 1950s, the Royal Aeronautical Society and the British aeronautical engineer Henry Kremer wanted to challenge the limits of aeronautical engineering (Raskin, 2011). They offered a prize to the first engineers who created a plane flown completely under the pilot's power. Many competing teams attempted to claim the prize. They would spend a couple years designing a plane, attempting a flight, crash, go back to the drawing board. Progress was crushingly slow. The prize was finally won when mechanical engineer Paul MacCready recast the problem from one of performance benchmarking to one of process design. By developing techniques where prototypes took just hours or days to construct instead of years, he was able to iterate and improve quickly enough to make the task tractable. His Gossamer Condor won the £50,000 prize in 1977.

The lesson here is that the way we frame the system influences the definition of the system's goals, and the quality of those goals can spell the success or failure of the system. It is essential to frame a system in a way that aligns with its goals (Raskin, 2011).

Goals are also models

Even within a well-framed system, it is important for actors to define the right goals, since those too are simply models for the outcomes we desire. A popular aphorism, frequently misattributed to management consultant Peter Drucker, says "What gets measured gets improved." This is meant as a kind of inspirational platitude: If you want something to improve, start measuring it! But this notion has a dark side, too: Whatever it is you happen to be measuring, that is the thing that will improve—possibly in isolation. In other words, you may forget that the system goal is a proxy for a real-world outcome and mistake achievement of the metric target for achievement of the outcome itself. This is the supply-side version of mistaking the symptom for the problem.

Take, for example, a possibly apocryphal story from India during the British occupation of the late 19th and early 20th centuries. The story goes that British rulers were uneasy with the prevalence of the venomous Indian cobra in the subcontinent, so they offered a bounty to the locals for each cobra head turned in. Naturally, in need of income, the locals took to breeding cobras, and then slaughtering them for the bounty. When the British realized their system had failed, they canceled the program. Breeders then let their captive cobras free, exacerbating the original problem.

The lesson here is that of the perverse incentive. A system goal that seems to reflect your actual goal, such as equating cobra heads with exterminated cobras, ultimately results in a counterproductive system output. For the Indians participating in the system, it was clear that the best way to achieve more cobra heads was to create more cobras. What gets measured gets improved.

The problem of goal definition gets even more complicated in the presence of conflicting goals. In the automotive world, the manufacturer's goal is to sell automobiles, which it attempts to do by creating cars that have features customers want, such as fuel efficiency and performance. At the same time, they are bound by limits imposed by the government, which essentially acts as an adversarial agent to keep the automakers in check. The government's goals include such things as safety and emissions compliance. Some of these goals overlap, but some of the government goals inevitably come into conflict with the manufacturer's goals.

This was the case with the recent Volkswagen emissions scandal. To summarize, Volkswagen programmed the engine control computers of some of its diesel vehicles to comply with emissions standards *only when the car was actually being tested* for emissions. At other times, the vehicles would operate in a higher performance mode, which released emissions at levels that violated pollution standards. Volkswagen accomplished this by programming their vehicles to pass emissions tests only under certain conditions. A government emissions test is highly standardized, so it has a predictable driving pattern. The vehicle stays in the efficient mode during the entire emissions test, passing the standard, despite releasing illegally high emissions with near certainty during normal driving (Lange & Domke, 2015).

What is the problem with this? The emissions test is a system goal—that is, it is a model, naturally an imperfect one, of what emissions behavior should be like in the real world. The intention is for performance under the model to reflect, in some standardized way, the performance in practice. By "teaching to the test" and seizing on the limits of the model, Volkswagen betrayed this implicit understanding that the system goal is supposed to represent the real world—revealing one set of behavior under scrutiny while engaging in a totally different, and noncompliant, way when not being evaluated.

The problem in these examples is summed up nicely in a concept originated by economist Charles Goodhart (as cited in Strathern, 1997): "When a measure becomes a target, it ceases to be a good measure." This is precisely the dark-side version of Drucker's apocryphal command to improve something by measuring it. As soon as the metric becomes valued as an outcome in itself, we forget that the metric is just a model that helps the system know the real world.

These examples show the importance of assessing system goals against real-world outcomes. It is a problem that is easy to succumb to, even unintentionally. System goals are models of real-world goals, so we must continuously scrutinize the system goals to ensure they are an accurate reflection of the desired real-world outcome. Is a six-year graduation rate really a good proxy for student success? Does cost per use accurately capture the relative value of a subscription? We will not know unless we audit our goals.

Data is a model too

The truth is that every part of system development involves the creation of models, and the way we select and define our models drives the success or failure of our system. The model-based nature of systems is perhaps clearest with data, which serve as a model for quantitative information that is important to the system. Since computation often underlies our systems, data have an obvious appeal, especially as we often need to deal with more stuff than we could manage by hand. Data also have a veneer of neutrality and correctness, since data are considered, more or less by definition, to be factual. But like all models, data are a simplification of what they represent—that is, when viewed as a model, data are wrong, and for data to serve the system, the wrongness needs to be compatible with our system needs.

The wrongness of data is highlighted in cases where we assume that they are intrinsically neutral. Consider: In 2016 an algorithm called beauty.ai was deployed to evaluate the attractiveness of human faces. The developers staged an online beauty contest, with over 600,000 people submitting headshots. The idea of whether aesthetic value can be determined algorithmically is a classic (and thorny) one in machine intelligence, and you might agree or disagree with the algorithm's findings or, fairly, bristle at the entire concept. Nevertheless, there was a noticeable trend in the algorithm's findings. Only one of the 44 winners had a dark complexion, despite the fact that the half-million submitted faces were reasonably diverse. Somehow, dark skin was selected against in the algorithm, even though it was not a parameter in the evaluation. What happened? The algorithm was presumably neutral. The data set was also supposed to be neutral. How is it that a computer can be racist? The problem is algorithmic bias (Levin, 2016).

In this case, when the artificial intelligence (AI) was being trained, it was given mostly faces of European origin. So when the algorithm was put into practice, and it came across non-European faces, it was not well trained to evaluate them, and they scored poorly (Levin, 2016). This highlights a key weakness of AI. A learning algorithm is only as good as the data set it is trained on. When we lean on algorithms to provide "objective" evaluations of historically subjective topics, the algorithm will reflect the same biases that are inherent in the data we give it. Bias in the data—the wrong kind of wrongness in our model—leads to a bad system output.

In a data-driven age, algorithmic bias can have genuine consequences for those caught up in it—so it is critical to inquire whether these systems are based on valid models. This is especially true as organizations and governments move toward algorithmic policies (such as predictive policing and data-driven financing), meaning that we are constantly being assessed as risks by algorithms, often without our knowledge. And these algorithms do in fact demonstrate inherent bias. One technology used in data-driven policing is the "crime predictor score," where demographic information about defendants is used to predict their risk of reoffending.

One such tool, called COMPAS, was investigated in depth by the research group ProPublica (Angwin, Larson, & Mattu, 2016). This algorithm, which is used in Broward County, Florida, as well as other locales, gives defendants a score of 1–10, indicating risk of reoffense. This score is used in part to determine whether bail should be offered. ProPublica's investigation into the product uncovered some troubling trends, finding that COMPAS systematically gives higher risk scores to Black defendants than White defendants. According to the vendor, in addition to criminal background, demographics such as education, employment, and residence are also used in the calculation of a defendant's risk score. But race is specifically omitted. How can an algorithm display systemic racism when race is not part of its calculation?

The problem is that just naively omitting a dimension does not guarantee that it is gone. If we exclude some factor—such as race—to try to prevent it from influencing the algorithm, we need to ask the following question: If you look at all the factors that the algorithm does consider, could you reasonably predict what you left out? That is, if you look at what is included—such as education, employment, location of residence, etc.—could you make a decent guess at the user's race? If the answer is yes, then you have not excluded race from the algorithm.

This is related to an idea from mathematics called linear independence. Due to deeply ingrained interplays between race and social structures in society, race is simply not independent of these other social factors, so it is almost trivial to make accurate predictions of race from other metrics that on the surface appear unrelated. Such relationships will inevitably surface in algorithms even if they are not programmed in. The systems takeaway is once again that data are a model, and especially with data, it is not always clear where the imperfections in the model reside. Integrating a model into a system without fully understanding its shortcomings leads to bad system outputs.

Another real risk of leaning too heavily on data comes in the moment of converting analysis into meaning. In the 2016 election, almost every major pollster overpredicted the probability of Hillary Clinton winning the election. When the votes came in, many spectators were shocked at her loss. Part of the failure came from an overreliance on data, combined with a naive lack of skepticism (Lohr & Singer, 2016). These components played into a narrative that made sense to pollsters and the public, so underlying assumptions were never critically examined. This is essentially a form of confirmation bias: The outcomes of the polls seem to support what we expect, so we do not investigate more closely.

Even in the face of changing voter demographics, pollsters steadfastly held to their standard methods for recruiting participants, leading to a selection bias that further exacerbated the problem (Lohr & Singer, 2016). At the heart of this problem is this: Polling agencies had too much faith in what systems analyst Gerald Weinberg (1975) winkingly called the Axiom of Experience: "The future will be like the past, because in the past, the future was like the past" (p. 141). The dubiousness of this so-called axiom should be apparent. Indeed, in fellow systems analyst John Gall's (2002) rejoinder: "In dealing with the shape of things to come, it pays to be good at recognizing shapes" (p. 139). The lesson here is that using the data of the past as a model for the present is not without its risks, and if we forget those risks, our systems may suffer.

The drive to examine complex problems in terms of data analysis is sometimes called "datafication"—essentially, viewing entire problem spaces through a lens of big data. But as designer Cennydd Bowles (2016) points out, committing too strongly to "datafication"—valuing data above all other problem dimensions—can lead to the sort of ideological blindness that we have seen here. When we entrust too much to the data, when we commit the fallacy that data are truth, our decisions will be flawed. We must not forget that data are simply a model for the real world, and like all models, the utility is limited. And so the final lesson: Data, like all models, can only inform our understanding of the system, not represent the whole thing. So, do not fall into the trap of chasing the data without contextualizing them properly against the system as a whole and against the real-world significance.

The takeaway

It is essential to recognize that all these system components—including the systems themselves—are incorrect, incomplete, and inherently limited. Once again, all systems are wrong. So as the environment changes, we must constantly reassess the wrongness of our systems to ensure that their wrongness remains useful. How do we ensure that the systems we make stay useful and stay true to our needs?

The first thing we need is the ability to assess systems. To do this, we need to be able to think in terms of systems. Being able to define a system, to recognize its boundary, to judge whether its components are defined correctly and its models are effective: These skills are central to critically evaluating the system. You can foster systems thinking in your organization by seeking out training in systems analysis, by evaluating new and unfamiliar environments in terms of systems, and by going through the exercise of modeling your own systems in detail.

Second, as we create and maintain our systems, we need to include opportunities for the systems themselves to be self-critical. Build in assessments of system goals against real-world goals. Create opportunities for skepticism within the system. Ask yourself, "What could possibly go wrong with this?" and then imagine worst-case scenarios. How can you minimize the gap between intended behavior and worst-case behavior?

Finally, we must recognize that human systems are human, and as such they are designed to further human purposes. Be on the lookout for cases where in-system expectations are being valued above human interests that are not completely represented. Create teams with a diversity of backgrounds and experience to better recognize assumptions, omissions, and bias in systems design. Remember that a key reason you may be ignorant of your system's oversights is that you are also ignorant of your own oversights. Ultimately, challenge yourself to consider the human outcomes of your system's decisions. After all, the system exists to support humans, not the other way around.

ORCID

Kristen Wilson http://orcid.org/0000-0002-3509-3417
Andreas Orphanides http://orcid.org/0000-0003-4441-8532

References

Angwin, J., Larson, J., and Mattu, S. (2016). Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica. Retrieved from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Box, G. E. P., & Draper, N. R. (1987). *Empirical model-building and response surfaces*. New York, NY: Wiley and Sons.

Bowles, Cennydd. (2016). Datafication and ideological blindness [Blog post]. Retrieved from https://www.cennydd.com/writing/datafication-and-ideological-blindness

Crockett, Z. (2015). What it's like to be declared dead by the government. *Priceonomics*. Retrieved from https://priceonomics.com/what-its-like-to-be-declared-dead-by-the/

Gall, J. (2002). *The systems bible*. Walker, MN: General Systemantics Press.

Lange, D., & Domke, F. (2015). *The exhaust emissions scandal ("Dieselgate"): Take a deep breath into pollution trickery*. Presented at the Chaos Communication Congress, Hamburg, Germany. Retrieved from https://media.ccc.de/v/32c3-7331-the_exhaust_emissions_scandal_dieselgate

Levin, S. (2016). A beauty contest was judged by AI and the robots didn't like dark skin. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2016/sep/08/artificial-intelligence-beauty-contest-doesnt-like-black-people

Lohr, S., & Singer, N. (2016). How data failed us in calling an election. *The New York Times*. Retrieved from https://www.nytimes.com/2016/11/10/technology/the-data-said-clinton-would-win-why-you-shouldnt-have-believed-it.html

Raskin, A. (2011). You are solving the wrong problem: When solving a difficult problem, re-ask the problem so the solution helps you learn faster. *UX Magazine*. Retrieved from https://uxmag.com/articles/you-are-solving-the-wrong-problem

Snow, S. (2015). How soylent and oculus could fix the prison system (a thought experiment) [Blog post]. Retrieved from https://blog.shanesnow.com/how-soylent-and-oculus-could-fix-the-prison-system-a-thought-experiment-e26be8b21a42

Strathern, M. (1997). 'Improving ratings': Audit in the British University System. *European Review*, *5*, 305–321. Retrieved from http://conferences.asucollegeoflaw.com/sciencepublicsphere/files/2014/02/Strathern1997-2.pdf

Weinberg, G. M. (1975). *An introduction to general systems thinking*. New York: Wiley and Sons.