Ethics and Information Technology (2021) 23:69–73
https://doi.org/10.1007/s10676-018-9468-2

ORIGINAL PAPER

Science fiction as a value scenario for historical technology

Ian S. King[1]

Published online: 11 July 2018

Abstract
The value scenario is a useful tool in the sheaf of methods within value sensitive design. When envisioning new technology, this tool supports the designer in speculatively considering relevant stakeholders, values expressed or rebuffed by an artifact's design, and tensions that may exist between those values. This paper explores how science fiction stories can serve as value scenarios to supplement traditional historical methods, especially when informants are no longer accessible.

Keywords Value sensitive design · Value scenarios · History of technology · Sociotechnical systems

The process of design requires envisioning that which does not exist. The many different theories of design have proposed mechanisms and methods to support the necessary creativity while focusing that same imagination upon the task at hand. One such focused envisioning method is the *value scenario* (Nathan et al. 2007). In the design of new technology, the value scenario supports discovery of potential stakeholder roles, those stakeholders' values, and potential value tensions, that is, how values may conflict. A classic example of a value tension is between and liberty and safety (i.e., freedom vs. security), as stated by Benjamin Franklin in 1755: "Those who would give up essential Liberty, to purchase a little temporary Safety, deserve neither Liberty nor Safety" (Franklin 1963). Value tensions may be resolved in a number of ways, or remain unresolved but acknowledged.

The *empirical investigation*, the realm of data collection through various research methods, is one of three methodological elements of value sensitive design (Friedman et al. 2002). The *technical investigation* explores the expression or experience of values within the made artifact. The *conceptual investigation* connects human experiences within a given *stakeholder role*[1] and one or more *values* important to that stakeholder role. There is no prescribed ordering of investigation and is it is common for results of one investigation to prompt revisiting an already-considered investigation.

The value scenario describes a sociotechnical system[2] that does not exist, but is nascent within the mind of the inventor-designer. Nathan et al. describe the value scenario as an extension of scenario-based design (Rosson and Carroll 2003) "that supports envisioning the systemic effects of new technologies" (Nathan et al. 2007). The extension of scenario-based design is the value scenario's consideration of a technology beyond its intended interactions with target users. The origins of the value scenario and its relationship to scenario-based design are described in detail in Nathan et al. (2008).

The value scenario is a crafted, speculative account of a proposed technology's interaction with society. But rather than a free-ranging speculation, the value scenario is premised upon five concrete elements: stakeholders, pervasiveness, time, systemic effects and value implications (Nathan et al. 2007). This methodological approach to new innovation in technology has been discussed in literature (cf. Czeskis et al. 2010; Woelfer et al. 2011), but this paper proposes an analogous analysis of existing technological artifacts, i.e., post hoc analysis of sociotechnical integration of a given innovation. One piece of that approach is the employment of extant fiction, in particular science fiction, in place of the purpose-built fiction of the value scenario.

Does a science fiction story look like a value scenario? Literary critic Darko Suvin describes science fiction as works of *cognitive estrangement* (Suvin 1979). A world

✉ Ian S. King
isking@uw.edu

[1] University of Washington, Seattle, WA, USA

[1] I frequently use the phrase *stakeholder role* rather than the simpler *stakeholder* because it emphasizes that a 'stakeholder' is not an individual.

[2] While some literature defines the *sociotechnical system* within the bounds of organizational science (Pasmore and Sherwood 1978, cf.), I have chosen a broader view such as that of Bijker (1997) and alternatively presented by Edwards (2003).

(or sociotechnical system) is described that differs from the empirically observed status quo, that is, a *novum*. The definition of the novum is based in cognition, that is, not as fantasy but instead premised as either a change in scientific law (e.g., antigravity, extraordinary energy sources) or technological expression (a new or changed made object). The story unfolds within the novum and explores the implications of the cognitively-defined transformation. A value scenario shares this nature of a cognitive estrangement, proposing a novum premised on cognitive change (the technological innovation) and exploring the implications of the resulting estrangement. The science fiction story and the value scenario are conceptually congruent.

Not all science fiction stories will easily lend themselves to such an analysis, but the prescribed elements of a value scenario can be derived in many stories. As an example, the following consideration of a novel of science fiction, "The Moon Is A Harsh Mistress" (Heinlein 1966), offers a mapping of the elements proposed by Nathan et al. onto an arbitrary work of science fiction.

"Mistress", published in 1966, was a critical success for its author, at the time perhaps more so than his earlier "Stranger In A Strange Land" (Heinlein 1961). The novum is a penal colony on Earth's Moon (called 'Luna' by residents), managed with a central computer system. Residents of Luna successfully rebel against rule by Earth through the entity of the Lunar Authority; this rule seems classically colonial in nature, protecting the delivery of hydroponically-grown foodstuffs to a needy Earth. "Mistress" has a strong libertarian influence, questioning the role of the nation-state (pp. 63–66).[3] While space travel and colonization of another world are technological aspects, the key technological focus of "Mistress" is the role of 'Mike' aka HOLMES IV, a powerful computer system that controls myriad aspects of the built environment of the main settlement of the penal colony, Luna City.

The following analysis of "Mistress" follows Nathan et al., with my interpretation of each criterion.

- *Stakeholders* Identify the roles of people who are either *direct* stakeholders, interacting directly with the technology, or *indirect* stakeholders who have no such contact but are affected by the technology. "Mistress" has a well-developed description of society, which aids but is not necessary for a stakeholder analysis. While the majority of Luna's inhabitants do not interact with the HOLMES IV as a computer, their lives depend on the computer's management of the Luna City environment, positioning them as indirect stakeholders.[4] A key character is a computer technician ('Mannie'), a direct stakeholder who is 'Mike's valet' (p. 9) and both maintains and enhances the system. The core group of revolutionaries, including an activist and a college professor, become direct stakeholders during the story, illustrating how stakeholder roles may be transformed by innovation or social interaction.

  Indirect stakeholders include a small privileged class of the 'Protector' *qua* prison warden and his staff, and the Lunar Authority, an Earth-based organization chartered with managing the penal colony; and a small transient staff of research scientists limiting their stays to avoid physiological adaptation to Lunar gravity. Finally, other smaller colonies, constituted primarily of former prisoners and their offspring, are not within the HOLMES IV's control. It is interesting to consider whether the smaller colonies become indirect stakeholders during the conflict with the Lunar Authority, when the safety of those colonies (unknowingly) rely on intervention in the conflict by the HOLMES IV.
- *Pervasiveness* Rather than considering a technology in a social 'silo', consider its effects if it is broadly present in society. Interpreting the description offered in Nathan et al., there is not a stated requirement of broad *acceptance* of the technology—there may be a population of stakeholders who are adamantly opposed to adoption but are nonetheless affected by that adoption.[5] The HOLMES IV controls nearly every aspect of Lunar intrastructure, down to individual residences (p. 101). Outlying settlements (most notably 'Hong Kong in Luna') have less direct dependency on HOLMES IV but would not be viable without the existence of Luna City.
- *Time* Consider the possible effects of the technology over a period of months or years after its introduction. The HOLMES IV system had been adopted by the Lunar Authority to the degree that it had become 'invisible' and (as noted above) increasingly pervasive. When Mike develops an emotionally intimate connection with Mannie, the invisibility of the HOLMES IV in society becomes a blind spot for the Lunar Authority as Mannie becomes a revolutionary and can leverage the HOLMES IV's control of infrastructure in service of the revolution (pp. 145–146).
- *Systemic effects* How does the technology interact as one component of a larger system, rather than in isolation? The HOLMES IV is integrated with all of Luna City's infrastructure. This provides economic benefit for Lunar Authority by obviating distributed computing infrastructure. Although other colonies are not directly integrated

[3] Page numbers are taken from the 1966 original paperback edition.

[4] The identification of stakeholders is a topic of ongoing discussion and debate.

[5] Consider the social dialogue regarding nuclear weapons, despite the fact that most people do not own any.

with HOLMES IV, they are dependent on Luna City as the center of Lunar government and commerce. Consider that at the time "Mistress" was written, actual computer systems were quite expensive. It should be noted that there are smaller computers on Luna that are employed by the revolutionaries in critical systems not requiring conversational interaction (p. 238). In the 1960s, so-called minicomputers[6] were beginning to challenge the adoption of centralized computer resources such as the fictional HOLMES IV.

- *Value implications* Identify the ways in which the technology and stakeholders interact, as stakeholders perceive it. The value scenario rests on the exploration of values that may be enabled by the technology as well as those that may be frustrated by it, and both those values perceived by stakeholders as positive as well as those felt to be detrimental. A key value driving the narrative is sustainability: a principal character states the current state of food exports will exhaust Lunar resources to ecological collapse. The Lunar Authority's reliance on a single central system served a value of economic frugality but was demonstrated to be in tension with its effective control. Mike's emergence as a personality was originally hampered by the computer's perception of humans as "stupid" (p. 13), and his growing emotional intimacy with Mannie and subsequently with the other core revolutionaries is key to subsequent social developments. The trust expressed by the humans in Mike's loyalty is instrumental to the implementation of the revolution.

The estrangement of "Mistress" is primarily based on two cognitive propositions: a Lunar colony and the related spaceflight technology, and a computer system that becomes sentient. The estrangement serves the narrative of the story by positing a political and social conflict between Earth and Luna that closely parallels historical narratives of colonial rule and revolution. A consequence of the nature of the colony is a scientifically stated motivation for sociopolitical change: the potential for collapse of the fragile Lunar ecology. The story element of the HOLMES IV computer illustrates the potential for pervasive presence of an information system to serve as a lever for sociopolitical change.[7]

"Mistress" describes a highly technological society whose technology, in the specific implementation of Mike, contributes to the well-being of its citizens, through self-governance and the prevention of ecological crisis. The message is not completely utopian, however. One character states that Mike is the greatest enemy of a free society because the computer represents total control over the flow of information (p. 205). The sentient nature of the machine also enables it to act with human agency and challenge human ethics, such as when Mike deliberately crashes an attacking spacecraft, ending the lives of all aboard (p. 257). Other value scenarios focus almost completely on ethical dilemmas, and are referred to as *design noir*. Stories such as "Colossus" (Jones 1966)[8] and "This Perfect Day" (Levin 1970) are examples of design noir.[9]

In emerging design, the role of a value scenario is to inform and advise, not to prescribe. In the historical analysis posited here, the temptation is to look for causal relationships between science fiction and technological innovation. Another possibility is to consider science fiction to be post hoc descriptive or documentary. To help avoid a 'whig history' analysis (Butterfield 1965), it seems wise to consider the temporal relationship between the science fiction story *qua* value scenario and the recorded history of the innovation as informed through the technical investigation. The following categorization is an initial attempt to address how to interpret this temporal relationship.

- *A considerable amount of time before the innovation—on the order of decades* In this case, the story may represent a pervasive and strongly-held range of interests and values. For example, the time between the 'logics' described in Murray Leinster's "A Logic Named Joe" (Leinster 1946) and fundamentally equivalent pervasive client/server information systems was approximately five decades, with initial implementations of conversational interaction with computer information systems occurring less than twenty years later than "Joe".
- *Contemporaneously with the innovation—within a decade of the innovation* Claiming causality may be most tempting here, but this scenario is more concretely interpreted to represent concurrent threads of social and technological development. My observation is that for a given sociotechnical system there is a cyclical exchange of influence and meaning between the social and technical elements, which one may posit is expressed within this temporal relationship. It must be recognized that there is a potential for either a naively descriptive

[6] There is general agreement that the first public use of the term 'mini-computer' was between 1967 and 1970.

[7] The use of pervasive information systems, i.e., social media, for economic and political purposes is a conversation occurring in the early twenty-first century.

[8] "Colossus" was written during the US/Soviet Cold War and is perhaps primarily a statement of political rhetoric, but its analysis of consequences within its scope is no less meaningful. Relevant to the narrative of "Mistress", the computer in "Colossus" also takes independent action to end human lives in furtherance of its self-interpreted purpose.

[9] Nathan et al. cites Dunne and Raby (2001) as a defining source of the design noir.

account or intentional advertising in support of the innovation. "Mistress", written within the same decade as the introduction of the JOHNNIAC Open Shop System (JOSS) (1963) and the Dartmouth Time Sharing System (DTSS) (1964), suggests another indicator of veracity: while both the authentic and fictional systems support conversational interaction, JOSS and DTSS are first steps along the path to systems such as Mike (which have not been fully realized in the twenty-first century). Even if Heinlein had been privy to JOSS or DTSS, "Mistress" was not about what is, but what might be, that is, Suvin's distinction between the author's empirical reality and the novum (Suvin 1979).

- *Subsequent to the innovation* Unless the author can be shown to have no connection to the social construction within which the innovation occurs, the story is most likely documentary. This does not immediately dismiss its value, for the story may use the innovation as a launching point for either a social analysis of the innovation or a launching point for yet another value scenario suggesting further innovation. The value of the analysis to a given discussion will depend upon the scope of that discussion.

  In the event that the author has no connection to the relevant social construction, such a story may provide interesting insight into the external visibility of the innovation within that construction, or in the alternative a perspective on the interpretation of that innovation in a different construction from the one in which the innovation originated. The potential contribution of this 'disconnected' viewpoint merits further work to understand how to evaluate the relevance and possible bias introduced by the viewpoint's disconnected yet post hoc nature.

The above analysis of "Mistress" demonstrates a value scenario as a lens illuminating the values and visions of the social construction encompassing a given sociotechnical system. "Mistress" was published in 1966 and presumably written within a small number of years prior. In this era, history observes the emergence of conversational interactive computer systems such as JOSS and DTSS. Unlike previous systems that employed non-interactive interfaces such as punched cards or tape, and interactive systems that allowed for remote job entry (RJE) or computer-mediated content delivery, JOSS and DTSS supported a back-and-forth exchange that built context with each query-response pair. While there is extensive literature describing these systems and their functionality, little is available about the values that promoted adoption of these innovations *from the perspective of the adopters*, i.e., stakeholders.

The nature of Mike in "Mistress" demonstrates an evident human desire to communicate with ever increasing intimacy with our companions in life. Works such as "Mistress" did not prescribe design or feature sets for systems such as JOSS or DTSS, but can be read as reflecting a vision of computers with which we can 'talk', contributing to a sense of intimacy with a non-human intelligence, writ large. Humans have ascribed human-like behavior to even early conversationally interactive systems (Friedman 1995; Nass et al. 1994) reinforcing the idea that we seek value-based connection with our creations. Science fiction such as "Mistress" explores the implications of these value-based connections, in the manner ascribed to the value scenario of Nathan et al.

In conclusion, the value scenario shares key definitional elements with science fiction that compose Suvin's cognitive estrangement (Suvin 1979). Employing science fiction stories as value scenarios can provide insight into the social construction of the era in which an innovation occurred, in a structured and accessible way. In addition to support of historical narrative, a conceptual congruence between science fiction *qua* historical value scenarios and value scenarios created in support of new innovation offers a common language of discourse between historical and contemporary technological transformations. For example, a similar approach to contemporary innovation, *design fiction*, is proposed by Sterling (2009).

I want to acknowledge the thoughtful and valuable feedback I received on this paper from Christiane Grünloh, Sarah Spiekermann, and the review team.

References

Bijker, W. E. (1997). *Of bicycles, bakelites, and bulbs: Toward a theory of sociotechnical change*. London: MIT Press.

Butterfield, H. (1965). *The Whig interpretation of history*. New York: WW Norton & Company.

Czeskis, A., Dermendjieva, I., Yapit, H., Borning, A., Friedman, B., Gill, B., & Kohno, T. (2010). Parenting from the pocket: Value tensions and technical directions for secure and private parent-teen mobile safety. In *Proceedings of the sixth symposium on usable privacy and security* (p. 15). New York: ACM.

Dunne, A., & Raby, F. (2001). *Design noir: The secret life of electronic objects*. New York: Springer.

Edwards, P. N. (2003). Infrastructure and modernity: Force, time, and social organization in the history of sociotechnical systems. In T. J. Misa, P. Brey, & A. Feenberg (Eds.), *Modernity and technology* (pp. 185–226). Cambridge, MA: MIT Press.

Franklin, B. (1963). Pennsylvania assembly: Reply to the governor, November 11, 1755. *The Papers of Benjamin Franklin*, *6*, 242.

Friedman, B. (1995). It's the computer's fault: Reasoning about computers as moral agents. *Conference companion on human factors in computing systems* (pp. 226–227). New York: ACM.

Friedman, B., Kahn, P., & Borning, A. (2002). Value sensitive design: Theory and methods. *University of Washington technical report* (pp. 2–12).

Heinlein, R. A. (1961). *Stranger in a strange land*. New York: G. P. Putnam's Sons.

Heinlein, R. A. (1966). *The moon is a harsh mistress*. New York: G. P. Putnam's Sons.

Jones, D. F. (1966). *Colossus*. New York: Putnam.

Leinster, M. (1946). A logic named Joe. *Astounding*, 139–155.

Levin, I. (1970). *This perfect day*. New York: Random House.

Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 72–78). New York: ACM.

Nathan, L. P., Friedman, B., Klasnja, P., Kane, S. K., & Miller, J. K. (2008). Envisioning systemic effects on persons and society throughout interactive system design. In *Proceedings of the 7th ACM conference on designing interactive systems* (pp. 1–10). New York: ACM.

Nathan, L. P., Klasnja, P. V., & Friedman, B. (2007). Value scenarios: A technique for envisioning systemic effects of new technologies. In *CHI'07 extended abstracts on human factors in computing systems* (pp. 2585–2590). New York: ACM.

Pasmore, W. A., & Sherwood, J. J. (1978). *Sociotechnical systems: A sourcebook*. San Diego, CA: Pfeiffer & Co.

Rosson, M., & Carroll, J. (2003). Scenario-based design. In J. Jacko & A. Sears (Eds.), *The human computer interaction handbook: Fundamentals, evolving technologies, and emerging applications* (Chap. 10, pp. 1032–1050). Mahwah, NJ: Lawrence Erlbaum Associates.

Sterling, B. (2009). Cover story: Design fiction. *Interactions*, *16*(3), 20–24.

Suvin, D. (1979). *Metamorphoses of science fiction: On the poetics and history of a literary genre*. New Haven: Yale University Press.

Woelfer, J. P., Iverson, A., Hendry, D. G., Friedman, B., & Gill, B. T. (2011). Improving the safety of homeless young people with mobile phones: Values, form and function. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1707–1716). New York: ACM.