Insolvent

How to Reorient Computing for Just Sustainability

By: Christoph Becker

Citation:

Insolvent: How to Reorient Computing for Just Sustainability
By: Christoph Becker
DOI: `10.7551/mitpress/14668.001.0001`
ISBN (electronic): `9780262374668`
Publisher: The MIT Press
Published: 2023

The MIT Press

1

THE DESIGN OF SUSTAINABILITY

> The dominant discourses about the nature of the climate threat are scientific and economic. But the deepest challenge is ethical. What matters most is what we do to protect those vulnerable to our actions and unable to hold us accountable, especially the global poor, future generations, and nonhuman nature.
>
> —Gardiner (2014, xii)

SOFTWARE HAS BECOME PART OF THE FABRIC OF OUR SOCIETIES

Technologies are social, and modern societies are technological. It is now widely acknowledged that there is a mutually constitutive relationship between the social construction of scientific and technological artifacts and the technologically mediated social relationships that make up what we call "society." But we are still struggling to grasp its implications. In computing, the question *how computing shapes societies* has been welcomed: It resonates with the widely accepted role of computer science as a source of innovation, and it aligns with the demand for impact.[1] The reverse question *how societies shape computing* has been largely left to other disciplines. Those who responded to it often come from critical and feminist perspectives in sociology, science and technology studies, information, communications, media studies, geography, and the humanities.[2] I call these researchers *critical friends* of computing.

Within computer science, the development of algorithms and especially of software systems is the primary purview of the discipline of *software engineering*. The field sees its focus as the development of technical systems with clear boundaries and identifiable parts and connections, modules, and dependencies. But fifty years after the founding of software engineering as a field,[3] the boundaries between software and its social and environmental contexts are rapidly dissolving. Software systems now have become part of our societies' fabrics and shape the relationships that constitute them, through their information storage, collection, aggregation, and routing; their algorithmic sorting and filtering; their communication and control capacities; and their ability to learn and predict patterns. Communication systems, dating platforms, travel booking services, and procurement systems influence the private, social, economic, and natural environment through far-reaching effects on how we communicate and form relationships, how we travel, and what we buy. The indirect effects of these systems generally remain invisible in the software development process, and developers routinely believe that the effects of their designs are not their responsibility. Their responsibility, as per their job descriptions, is to create "software systems." But the boundaries that matter for understanding *what system has been designed* in each case become increasingly difficult to ascertain. Since every relevant software system is deeply interlinked with people, economic processes, social relationships, and other elements, what we need to consider *as system* inevitably transcends the boundaries of the software.

Software development then always creates sociotechnical, socioeconomic, sociocultural systems.[4] When this book refers, for the sake of readability, to "software systems," it will be always based on that understanding. But when software became social, software engineering research did not make the leap of fully incorporating social theory into its foundational body of literature, and neither did software engineering education (Dittrich, Klischewski, and Floyd 2002; Ralph, Chiasson, and Kelley 2016). As a consequence, both lack the theory, conceptual tools, and people to ask the critical questions needed to understand software technology's role in our societies (Leticia Duboc, McCord, et al. 2020).

Now that this role has become formative, we are beginning to recognize the need for change. For the sake of future generations, the software systems we design in the next fifty years must advance social, economic, and environmental causes simultaneously with technical ones. The ability to meet urgent needs of the present without compromising future generations' ability to flourish requires that all stakeholders must jointly understand and address the social, human, and technical sides of systems design and must explore systemic effects across the physical, societal, economic, and human environments that software is entangled in. But as we will see, they are currently unlikely and ill-equipped to do so.

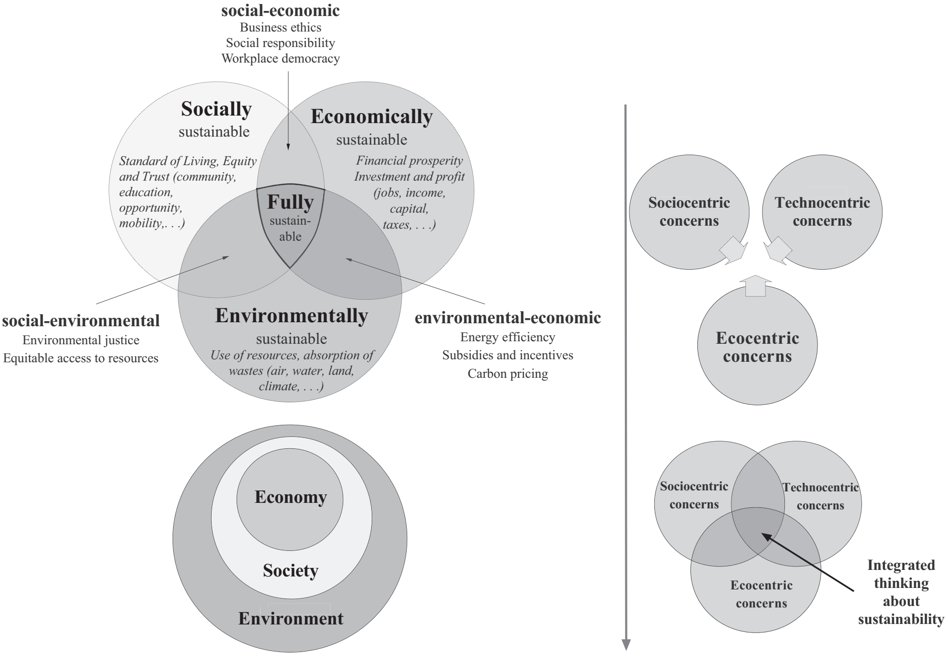
Two implications of the pervasive role of software and information technology are central to the argument of this book: the long-range concerns of sustainability and the wide-ranging concerns of justice. Both have become urgent as computing now pervades our societies. The unprecedented breadth, depth, and length of its entanglement means that the choices made in systems design have farther-reaching implications than ever before. These distant effects are often uncertain and ambiguous. This complicates the political role of technology design, the social dynamics of participation in design, and the cognitive factors of design decisions. The book grapples with each of these implications. To establish its domain, this chapter will trace the evolving perspectives on the relationship between computing and sustainability, loosely following its historical evolution.

COMPUTING AND SUSTAINABILITY HAVE A CONFLICTED RELATIONSHIP

Sustainability at its heart is the "capacity to endure" (Fowler and Fowler 1995) or, sometimes, the "ability to *be maintained* at a certain level" (*Oxford English Dictionary* 2021). Its primary prominence arose through the concept of *sustainable development* as defined in the United Nations' "Brundtland report"—"development that meets the needs of the present without compromising the ability of future generations to meet their needs" (Brundtland 1987). The use of the concept has evolved considerably: The initial view focused on avoiding the depletion of a stock of natural resources held in identified ecosystems. It drew on a global view of the planet at large and its limited resources. As famously outlined and demonstrated in the classic *Limits to Growth* report in 1972, human civilizations were already then on course to exceed the *carrying capacity* of our planet (D. L. Meadows and Club of Rome 1972). Today, the focus lies on a systemic view of the interlinked dimensions of social, environmental, and economic processes that are now seen to constitute sustainable development. For example, the UN's Sustainable Development Goals (SDGs), which now provide the framework for most large-scale initiatives and policies (International Council for Science 2017), contain "public participation" as a core goal. In North America, many businesses now use a version of the "triple-bottom-line," a model that asks companies to balance economic, social, and environmental accounts (Elkington 2004).

Sustainable development may still be the hegemonic paradigm, but it is a problematic framing, far from innocent. We will revisit it in chapter 8. In this book, I focus on the role that systems design plays in sustainability, rather than sustainable development, while acknowledging that design always implies a form of "development": it aims to shape the world according to the designers' intentions. The crucial issue is: whose intentions?

Two competing views of sustainability are illustrated in figure 1.1.[5] At the bottom left, the strict hierarchical "strong" view of sustainability places the economy inside a society, which is placed inside the natural



1.1 Sustainability visualized as hierarchy, interacting domains, and spheres of concerns.

environment. This view maintains that real limits of each of these strictly hierarchically contained systems must be respected for the continued viability of the whole. Each level of the hierarchy must remain within established limits, or else it damages its surroundings. This approach distances itself from what its proponents call "weak" sustainability, shown on the top left. This paradigm assumes that some resources within one domain may be substitutable by resources from another domain. For example, fossil fuel energy has substituted human labor, and economic profits could be used to plant trees. In this interpretation, some resources may be depleted as long as other resources continue to fulfill their functions. While this may appear reasonable, there are serious objections to this view.[6]

It is important to recognize the contrasting epistemological basis of different models. The perspective on the top left focuses on the *domains* to be addressed and organizes these according to their intersections and overlaps. For the purpose of systems design, this book follows the argument that Dodds and Venables make in the context of civil engineering (2005): The key focus is to integrate the *thinking about* these relevant spheres within a process that invariably involves a multitude of stakeholder perspectives and knowledge areas. The crucial emphasis then is one of *concerns*, as illustrated on the right (Becker et al. 2015). This shift sidesteps the debate between the views shown on the left and provides a more encompassing perspective than a resource-based view.

Integrating these concerns presents at least three types of challenges to individual and collective action. First, the effects of human choices on the environment are dispersed in space and time. Each individual action may make only a small, uncertain, opaque contribution to climate change, but their effects interact and accumulate over time. Future generations will be disproportionately affected, but they cannot speak up about their expectations of inheriting a livable planet. This creates what moral philosopher Stephen Gardiner (2014) calls *asymmetric vulnerability*. We will later argue that it is present not just in the global challenge of climate change, but in many seemingly "smaller" issues of systems design, where dispersed causality, fragmented agency, and asymmetric vulnerability combine to create a dangerous undertow.

Second, the dispersion of cause and effect across spatial and temporal timescales also raises difficult questions about the nature of human cognition and reasoning. As judgment and decision-making researcher

Elke Weber has argued, climate change “doesn’t scare” the people making choices yet because it happens, for many in the developed world, at a distance, so our knowledge of it lacks the viscerality of personal experience (Weber 2006). This *psychological distance* affects our choices, but little is understood of how this plays out in systems design (see chapter 11).

Finally, the integrated consideration of social, environmental, economic, and technological notions of sustainability raises serious political and epistemic challenges, since it is difficult to reconcile the diverse assumptions and worldviews from which the concerns arise. Specific trade-offs across dimensions can be agreed on and have been formalized—for example, it requires an economic investment to decarbonize specific industries, and decarbonization will also result in direct and indirect economic benefits. But the stakeholders will also bring their mutually incompatible worldviews on what constitutes such foundational things as social life, individual prospering, economic activity, and nature. These epistemic differences frame what constitutes a valid measure within each dimension and how the dimensions should relate to each other. As a result of this epistemological conundrum of incommensurability, sustainability becomes “immeasurable” precisely when the aim is to evaluate it.[7] Bell and Morse (2008) respond with a shift to a subjectively conceived and intersubjectively negotiated system of sustainability indicators developed through a participatory approach: “Systemic Sustainability Analysis [is] the participatory deconstruction and negotiation of what sustainability means to a group of people, along with the identification . . . of indicators to assess that vision of sustainability” (147).[8]

Computing researchers recognized early that IT has a role to play in sustainable development.[9] Several areas took an interest. Each sees the topic through its own disciplinary lens, which shapes how researchers frame and approach it.[10]

GREEN IT: ADDRESSING THE ENVIRONMENTAL IMPACT OF COMPUTING

Computerization has led to a rise in mineral extraction, energy consumption, and electrical and electronic equipment waste. A range of fields tackle these issues, driven by environmental impact assessment research (see Widmer et al. 2005; Robinson 2009). Research partnerships with IT were soon labeled “green IT” and “green computing” (Jenkin, Webster, and McShane 2011; Murugesan 2008).

On the production side, the extraction of rare minerals and other ingredients for IT products has turned into a supply chain challenge[11] and, more importantly, a human and nonhuman rights question.[12] The public’s attention rarely focuses on where the material and labor come from that make possible the global IT infrastructure, and where those materials go after they cease to be useful. But the material roots that link modern IT devices to the material and social conditions of their production span massive spatial and temporal scales. The brilliant visualization *Anatomy of an AI*[13] drew attention to just how far one has to go to identify the sources and inputs to one “smart home” device. The media’s reception of this piece demonstrated vividly how surprising this realization was to a mainstream audience, even though scholars have for many years emphasized the invisible labor behind IT infrastructure and its uneven global distribution.

At the other end of the life cycle stands obsolescence. E-waste is the general term for “all items of electrical and electronic equipment (EEE) and its parts that have been discarded by its owner as waste without the intent of re-use” (Kuehr 2014). In 2017, the amount of e-waste generated globally reached 44.7 million metric tons, as about half of the planet’s population has started using internet services. Only a fifth gets recycled, while the rest remains unaccounted for and is most likely not recycled (Balde et al. 2017). While developed countries have reached high levels of saturation in terms of technology use, consumption in countries with lower purchase power is growing by around 20 percent annually (Balde et al. 2017). Because obsolescence is not a law of nature but a sociotechnical phenomenon, we can design to avoid obsolescence or to generate it (Slade 2009). Unfortunately, the latter is the norm. Despite technological advances, the lifespans of electronic equipment continue to be extremely short. In 2005, a study estimated an annual disposal rate of 11 percent for personal computers (Widmer et al. 2005); in 2017, a similar level was estimated for laptops (Balde et al. 2017). Regulation efforts such as Extended Producer Responsibilities have made progress in placing some of the burden on the production side, but most global e-waste is still moved from rich countries to poor countries, where it is recycled under extremely unsafe conditions (e.g., Rifat, Prottoy, and Ahmed 2019). The manual extraction of the residual valuables inside electronic equipment (such as copper and gold) exposes laborers to a wide range of toxic chemicals and metals, including lead and mercury, with significant damage to their physical and mental health.[14] Unsurprisingly, these effects are very unevenly distributed, since production, consumption and disposal happen in distinct parts of the world separated by their socioeconomic development status.

Computing has responded in varied ways. Industry and research have continued miniaturization and cloud computing, based on the claim that efficiency gains would reduce waste accumulation. With the growth of mobile device use and cloud computing, the share of consumption has shifted from desktop PCs to mobile devices and data centers, while overall consumption continues to increase (Belkhir and Elmeligi 2018). Researchers and advocates have collaborated with policy makers to measure and make visible the burden of e-waste. The methods, often built on established environmental impact assessment frameworks such as life-cycle assessment, faced challenges of evidence collection, indicator definitions, and measurement (Guinée et al. 2011; Balde et al. 2017). The evidence base has led to a slow change in policy and regulation, which has not changed the direction of the overall trend: e-waste numbers continue to grow. Outside the mainstream, advocates, activists, community organizers, and some researchers have argued for extended technology lifespans through maintenance, repair, reuse, and modularity (S. Jackson 2014). These interventions certainly reduce waste, but they remain marginal efforts.

Between production and disposal stands usage, and with it the steadily increasing energy use of computing. IT accounts for a growing fraction of the world's rising electricity demand (Van Heddeghem et al. 2014; International Energy Agency 2013). The training of one single machine learning model emits as much CO_2 as five average cars in their entire life cycle (Strubell, Ganesh, and McCallum 2019). A significant portion of research in green computing, green IT, and green software engineering has thus focused on "the problem of energy use," with the declared aim of reducing the environmental impact of software systems by increasing efficiency in software production and use (Calero and Piattini 2015). Energy-efficient algorithm research is based on the claim that "algorithmic solutions can help reduce energy consumption in computing environments" (Albers 2010). At first glance, this seems direly needed. Advances in the energy efficiency of algorithms have not however reduced aggregate energy consumption. On the one hand, the share of data centers in global energy use has remained steady while their scale has rapidly expanded (IEA 2019). This is attributed largely to the efficiency gains made by extremely large-scale hardware and facility design (Shehabi et al. 2016), including the algorithmic control enabled by deploying sensors in large-scale data centers (Jones 2018; Hölzle 2020).

The energy efficiency of algorithms themselves may seem a promising research subject, since on the level of processing, there is a clear link between algorithmic complexity, memory operations, processing cycles, and energy use. But evidence is scarce that improvements on algorithmic efficiency have tangible effects. Numerous papers published in computer science conferences each year promote their algorithms as more energy efficient than previous algorithms. The strongest business driver for this trend is cost (Albers 2010), but many researchers in energy efficiency motivate their work by environmental sustainability. Few consider that their work may push an important lever in the wrong direction. The reason lies in a paradox already observed in the nineteenth century: increased technological efficiency often *increases* overall resource use (Alcott 2005). Why? Aggregate energy consumption is a product of average efficiency multiplied by total use. The expectation may be that efficiency gains result in lowered consumption, but use is not independent of efficiency. An increase in technology efficiency or capacity often induces increased usage. As a result, aggregate consumption *rebounds* against these expectations. For example, historically, despite a *one hundredfold* efficiency increase from the first light bulb to a contemporary LED bulb, increase in electricity demand *for light bulbs* has entirely offset these gains. A landmark study concluded that "global energy use for lighting has experienced 100% rebound over 300 years, six continents, and five technologies" (Saunders and Tsao 2012; Tsao et al. 2010). Just take a look around your home and try to count the bulbs.

The rebound effect is relevant for any comparable situation and often exceeds the efficiency gains, as has been observed for coal use in railways, for highway capacity increments, and for countless other cases (Alcott

2005; Sorrell 2009; Freeman, Yearworth, and Preist 2016). The compound outcome depends on economic, social, and cultural factors. It is not possible then to make a direct claim from individual-level efficiency to aggregate contribution without assessing the change of aggregate energy use. But despite the fact that rebound effects have been well studied and published, few studies in energy efficiency in computing even mention them (Knowles 2013, 4). This severely limits the relevance and value of energy efficiency work in particular (Coroama and Mattern 2019) and green computing research in general (Knowles 2013, 89–91).

Consider a bitcoin miner using a software running an algorithm that, on their hardware, produces $1 worth of bitcoin per hour with an energy cost of 90 cents, for a return on investment (ROI) of 11 percent (ignoring for simplicity the hardware costs). Decreasing the amount of energy required to produce 1$ worth of bit coin by 20 percent will increase the ROI to 27 percent. Will the miner mine the same amount of bitcoin (reduced energy), or expend the same amount of energy to mine more bitcoin (even energy balance), or invest in additional hardware as a response to the changed incentive structure (increased energy and environmental waste)? Considering the scalability of their enterprise and the fact that their ROI almost tripled, the answer seems clear. In fact, the increased attractiveness to investors may well spur new entrants into the market.[15]

If research and development into energy efficiency aims to improve environmental sustainability, the “system in question” to consider is not the bitcoin mining algorithm in isolation but the bitcoin mining algorithm *in use by a person*. Neglecting to do so commits what Churchman (1979a) calls the *environmental fallacy*: It takes too narrow a perspective to understand the effects of an intervention.

In the system of interest, increased efficiency *increases energy demand*. The outcome is often an increase in aggregate consumption. Just as with lighting, the real effect of increased efficiency in bitcoin mining is not reduced usage but increased productivity.[16] Resolving this paradox thus requires a reframing of the system of interest, and the extension of system boundaries comes with an expansion of the knowledge domain required to perform this type of research, from a mathematically and computationally grounded domain of algorithms to a broadened *systemic* understanding of relationships that span algorithmic complexity, hardware efficiency, economic models of supply and demand, and the incentives of purposeful social actors.[17] Such a paradigm shift from computational to systemic thinking does not come easy.

GREENING THROUGH IT: HOW COMPUTING CAN GREEN SOCIETIES

If e-waste and energy consumption research starts from the question “how does computing pollute the world?” and reduces that impact, *green by IT* research asks the complementary question “how can computing green societies?” directly and indirectly, and aims to create and amplify that impact. In *Greening through IT*, Tomlinson (2010) provides a comprehensive account of this research program. In his view, IT provides ways to “compress time, space, and complexity” through its capacities to store, retrieve, and analyze information and to support visualization, modeling, and simulation (9). Green IT then can support the challenging leap from human to environmental scales of thinking and action. Tomlinson illustrates this in three example contexts. In education, IT designs can help learners understand ecological concepts through simulation-based interactive systems. But this idea reaches far beyond formal education and can involve quite visceral explorations of nature at a scale not generally accessible to human senses (e.g., Driver 2022). On a personal level, IT can support data collection and visualizations about an individual’s footprint and the factors contributing to them. On a collective level, platforms can coordinate, encourage, and mutually reinforce individual actions for sustainability. Tomlinson positions this work carefully in respect to broader critiques that place capitalism and its hunger for growth at the center. I agree with him that that independently of the role of capitalism, IT will play a role in addressing climate change (Tomlinson 2010, 25).

ENGINEERING SOFTWARE FOR SUSTAINABILITY

Within software systems research and practice, the challenge of incorporating long-term considerations effectively into engineering practice has been a central concern since the founding days of the field (Becker 2014). The long-term costs of software systems were one of the central themes at the 1968 NATO conference (Naur and Randell 1969), and Lehmann's laws of software evolution were developed in the 70s (Lehman 1979; 1996). Two decades later, Parnas (1994) lamented "software aging." Another two decades later, the practical urgency of long-term thinking had not diminished (Neumann 2012).

For the most part, however, these long-term considerations were focused on internal perspectives motivated by cost reduction. When the subfield of software maintenance and evolution adopted the language of sustainability, it did so without acknowledging the broad concerns inherent in the wider sustainability discourse. Sustainability language became prominent about a decade ago (Becker 2014; Durdik et al. 2012; Koziolek et al. 2013; Zdun 2013), and sustainability-focused work in software architecture receives growing attention (Venters et al. 2018).

With the emergence of software engineering for sustainability, several research groups began to explore the role of software systems in environmental sustainability. While some focused on work that can be categorized as green software engineering[18] thanks to its focus on energy efficiency, a workshop series on requirements engineering for sustainable systems explored the intersection of the social and the technical perspectives of software engineering, especially the domain-specific challenges of eliciting requirements for software explicitly designed to support environmental sustainability (e.g., see Mahaux, Heymans, and Saval 2011; Chitchyan et al. 2015). Penzenstadler positioned sustainability as "the non-functional requirement of the 21st century" (Penzenstadler et al. 2014). Like safety, security, or usability, sustainability is not simply located in particular features but presents a concern that cuts across functionality, largely independent of the primary purpose of the system under design. Like usability, the implications of technical design choices are varied and must be considered carefully. A review of nonfunctional requirements in the past shows that the emergence of a new concern is generally followed by the growth of knowledge, techniques, measures, and models to address it. But the scale and complexity of sustainability make it more difficult and challenging to address.

Some have proposed to consider sustainability a system "quality"—that is, a property of the system under design (Lago et al. 2015)—but it is important to keep in mind that this reduces the frame of design again to internal and purely technical perspectives. This focus in turn marginalizes the more profound implications of systems design on the broader environment of software technology and treats technical aspects as fully separable from social aspects. As a result, it treats these questions as solvable problems and positions conventional engineering as the sole applicable method. Because sustainability emerges from the *interactions between* the designed technology and its environment, however, it cannot be reduced to a technical property (Becker 2014; Becker et al. 2015). We must resist the reductive framing implied in sustainability as quality and recognize it as a stakeholder concern: a matter of interest to those affected. This rejection of sustainability as system quality has far-reaching implications for systems design methods, models, and practice. It implies that the evaluation of a system's sustainability ultimately must rest with all those who are affected, and that their concerns may translate into a range of system features and qualities.

ICT FOR SUSTAINABILITY

The contours of the *ICT for Sustainability* field (short ICT4S) took shape in 2013 with an inaugural conference (Hilty and Aebischer 2015b), but the conflicted relevance of ICT for environmental sustainability was articulated earlier (Hilty et al. 2006) at about the same time as the focus on energy consumption emerged elsewhere. ICT4S aims to integrate many of the above concerns in an overarching framework that recognizes ICT both "as part of the problem" and "as part of the solution" (Hilty and Aebischer 2015a) and distinguishes between its direct, indirect, and aggregate effects,[19] defined as follows:

- **Life-cycle impacts** refer to the direct environmental and social effects of producing, using and disposing of ICT, which are negative not by definition but in practice (until the production of ICT equipment can be made carbon-negative). The green IT research summarized earlier focuses on these impacts.
- **Enabling impact** is located on the micro level of individual and organizational IT adoption and use. The work of Tomlinson and others explores this space, with a focus on potential positive applications. But enabling impact often has negative implications too, including

the effects of induced consumption and obsolescence. For example, the shift of music distribution from CD sales to online streaming has reduced the environmental impact of CD production but increased the energy consumption of data centers and internet infrastructure (Devine 2019). According to one calculation, the carbon footprint of streaming a song exceeds that of its CD counterpart once it is played more than twenty-seven times (McKay and George 2019).

- **Structural impact** refers to aggregate, macro-level changes induced by large-scale adoption of a product or service. These changes are enabled by micro-level effects, but because of the emergent nature of complex sociotechnical, socioeconomic and sociocultural systems, it is empirically difficult to link them conclusively.

To understand how these effects are linked, consider the well-known online rental platform Airbnb.[20] Its life-cycle impacts refer to the direct environmental and social effects of production and use—that is, the labor of creating and operating the platform and its physical effects, including energy consumption and e-waste. Its enabling effects result on two sides. First, the software platform has allowed travelers to change their habits by offering a new and attractive way of locating and booking accommodation. This attractiveness of lower prices and new types of experiences has induced travel, and it has allowed private and business travelers to shift their booking practices. Second, short-term rentals are more profitable than long-term rentals, and investors around the globe jumped on the opportunity. With increasing scale, more induction effects appear: new business ventures develop and manage condo tower buildings purposely designed for short-term rentals.[21] As an aggregate result of the rapid growth of IT-enabled short-term rentals, metropolitan areas have lost tens of thousands of residential housing units (Combs, Kerrigan, and Wachsmuth 2020), which has contributed to gentrification, increased rental costs, driven out residents, and changed the character of urban life (Wachsmuth et al. 2017; 2018). When municipal governments caught on and began to regulate short-term rentals, they encountered a third enabling effect: The technological affordances of the software platform support the circumvention of regulation, both on the side of individual owners and on the side of business operators, because the commercial database of transactions is under the control of Airbnb and thereby outside the jurisdiction of most national authorities. As a US company with enormous cash reserves, Airbnb invests significant sums into its legal department and has little incentive to play along when small countries request access to business transactions for tax purposes. In addition, the platform deliberately obfuscates the precise location of rental properties (Wachsmuth and Weisler 2018), which hampers the enforcement of regulation.

The story of Airbnb illustrates that the distinction between direct, enabling, and structural impact is crucial for two reasons. First, because it relates to a central feature of information technology: scale. Online services such as Google, iTunes, Netflix, Airbnb, Uber & Co have been able to have such a marked influence on our lives because they are able to scale more flexibly and faster than traditional businesses. For example, Airbnb has grown faster than any hotel chain ever could. Second, this illustrates that structural impact does not arise from enabling effects out of aggregation or multiplication. Instead, macro-level structural effects *emerge* as a result of large-scale adoption due to the convergence of many dynamic factors, and they are of an entirely different nature. In the case of Airbnb, they are not restricted to a disruption of the hospitality industry but also affect many countries' abilities to collect taxes, exacerbate global cities' municipal housing shortages, and alter the character of neighborhoods around the globe. The example also illustrates that this framework of effects can be used to analyze both positive and negative effects of ICT.

An early mapping of a published collection in ICT4S to types of effects suggests that a majority of work in this field focuses on issues such as energy efficiency that can be categorized as life-cycle effects, while some address enabling effects of production (Hilty and Aebischer 2015a, 32). This limited focus on efficiency may be misguided because a range of *rebound effects* often offsets efficiency gains, and to identify these, we need to examine behavioral and structural change (Hilty et al. 2006). As Coroama and Mattern conclude in their review of the evidence about these rebound effects, "digitalization will not redeem us from our environmental sins" (Coroama and Mattern 2019).

SUSTAINABILITY IN HUMAN–COMPUTER INTERACTION: DESIGN WITHIN LIMITS

The conflicted relationship of IT design to sustainability has been a defining topic of work in human–computer interaction (HCI) (DiSalvo, Sengers, and Brynjarsdóttir 2010). An influential paper by Eli Blevis (2007) positioned the importance of sustainability in interaction design and evaluated design values and methods from that perspective. The focus is firmly on the direct life cycle effects of interaction design as it relates to resource use and waste, with some attention to the induction and obsolescence effects of different design features. But the discussion raises profound questions, citing design scholar Tony Fry's (2005) critique of design practice:

> Currently industry . . . is still overwhelmingly deaf to those voices that speak of the complexity of unsustainability, the poverty of current responses to it, the misplaced faith in technological solutions, the myopia of present political and corporate leadership and the extent of changes that are required if a psychology, culture and economy of sustainment are to ever arrive. (23)

Fry (2005) further argued that designers "need to learn . . . how to design in a far more complex and critical frame" (23). In Blevis's (2007) words, "Fry's statement acknowledges the value tensions between sustainability goals and those of enterprise, while prescribing an ethical imperative for designers to confront such tensions" (504).

Blevis acknowledges the prevalent anthropocentrism in "human-centric" HCI research and places significant emphasis on the mutual relationships between interaction design and sustainability. Nevertheless, a majority of work on sustainability in HCI has focused on designing ambient and persuasive technology to influence consumer behavior (Knowles 2013; DiSalvo, Sengers, and Brynjarsdóttir 2010). This broadly falls into the area described earlier as *Greening through IT*. But critical voices and perspectives have been prominent in HCI. Strengers (2014), for example, eloquently argued that many designs of persuasive technology implied gendered stereotyped assumptions about how individuals make choices about their consumption, and Dourish (2010) highlighted that the common mode of uncritical design-oriented research "obscures political and cultural contexts of environmental practice that must be part of an effective solution" (1). He emphasized the central importance of ecological and political perspectives in designing with sustainability in mind. DiSalvo and colleagues have shown that there are substantive disagreements within HCI researchers: about the appropriate scale of design (from the individual to social infrastructure); about the tension between seeing "users as the problem vs. solving users' problems"; about the tension between incremental and disruptive change; and about the question whether HCI as usual is an adequate mode for addressing sustainability (DiSalvo, Sengers, and Brynjarsdóttir 2010, p. 1979). Bran Knowles's (2013) values analysis shows that persuasive tech-based intervention is likely to reinforce environmentally harmful behaviors and values (86–89). Even worse, its discourse "reifies consumerist tendencies that have driven much of the environmental destruction to date, and absolves individuals from having to make more significant behaviour changes" (87).

The computing within limits community emerged within this context.[22] It argues that the computing mainstream is "deeply problematic for ecological and social reasons" (Nardi et al. 2018, 86). In taking a "strong sustainability" view, this group of researchers argues for closer attention to planetary boundaries. From this perspective, they highlight the importance of rebound effects and criticize that classical green IT research in its narrow focus on energy efficiency ignores them. Three principles are derived from focusing on planetary limits: (1) *Question growth* encourages computing researchers and practitioners to avoid work that ultimately depends on economic growth and instead find alternative angles of work that do not encourage it; (2) *Consider models of scarcity* encourages researchers and practitioners to abandon the focus of designing for abundance, as the work in *Collapse Informatics* does;[23] (3) *Reduce energy and material consumption* emphasizes the need to minimize the footprint of computing.

Computing within limits has made significant contributions to the discourse: It firmly embedded computing within societies and the planet at a long-term and global scale and irrevocably demonstrated the responsibility that this places on the field; it brought findings from archaeology and ecological economics to bear on questions of technology design; it compellingly advocated for transformative change; and it has demonstrated that a different kind of computing can at least be envisioned.

SUSTAINABILITY DESIGN: CONVERGING PERSPECTIVES

Within this evolving landscape, it became clear that the communities addressing sustainability-related concerns through their disciplinary lenses had grown in fragmented paths with limited interactions. In 2014, an attempt was made to unite some of these concerns by providing a common ground for the above communities in the form of the “Karlskrona Manifesto for Sustainability Design.”[24] The manifesto recognizes the conflicted relationship between software systems and sustainability discussed here and emphasizes the responsibility of those who design, understood broadly as “the process of understanding the world and articulating an alternative conception on how it should be shaped, according to the designer’s intentions” (Karlskrona Initiative 2015; Becker et al. 2014). At its center is a set of principles and commitments to a transdisciplinary, systemic, long-term view (Karlskrona Initiative 2015; Becker et al. 2014). Its conceptual framing of sustainability as a concern distinguishes among the five dimensions of individual, social, economic, environmental, and technical sustainability, which are defined and illustrated loosely. The manifesto further takes a decidedly pedagogical stance in highlighting a series of common perceptions, framing them as *misperceptions*, and offering alternative viewpoints. Table 1.1 lists some and maps them to key themes discussed throughout this book.

THE LEVERAGE OF REQUIREMENTS

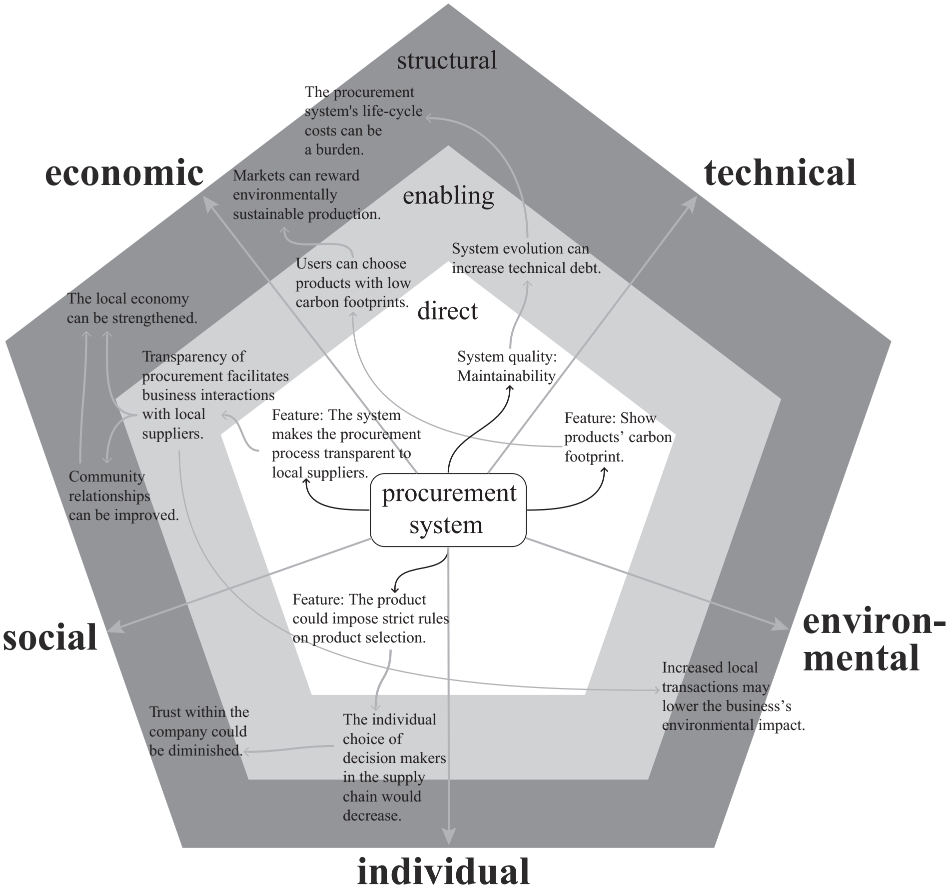
The Karlskrona Manifesto kickstarted an international initiative of researchers and took important steps toward uniting researchers across communities.[25] The convergence of five dimensions with systemic effects allowed the development of a framework to capture the possible effects of system features and qualities visually. Figure 1.2[26] shows how an instance of this framework represents the range of potential effects of a procurement system and demonstrates the entangled nature of dimensions. For example, making visible each product’s carbon footprint in a procurement system allows users to make more responsible choices. Widespread adoption of more responsible choices would influence market dynamics to incentivize more carbon-friendly products. The affordances designed into individual systems have a limited but tangible role to play in the overall shift.

Importantly, the arrows in the diagram represent causal contributions that connect micro-level decisions to aggregate effects. The diagram links these effects across the dimensions using a visual canvas that may allow a range of participants to engage in the design process. Recent work

Table 1.1 Selected misperceptions countered in the Karlskrona Manifesto

(Becker et al. 2014)

There is . . .	Whereas . . .	Theme
There is a perception that there is a tradeoff to be made between present needs and future needs, reinforced by a common definition of sustainable development, and hence that sustainability requires sacrifices in the present for the sake of future generations.	Whereas it is possible to prosper on this planet while simultaneously improving the prospects for prosperity of future generations.	Tradeoff decisions
There is a tendency to overly discount the future. The far future is discounted so much that it is considered for free (or worthless). Discount rates mean that long-term impacts matter far less than current costs and benefits.	Whereas the consequences of our actions play out over multiple timescales, and the cumulative impacts may be irreversible.	Psychological distance
There is a tendency to interpret the codes of ethics for software professionals narrowly to refer to avoiding immediate harm to individuals and property.	Whereas it is our responsibility to address the potential harm from the second- and third-order effects of the systems we design as part of our design process, even if these are not readily quantifiable.	Ethics
There is a desire to identify a distinct completion point to a given project, so success can be measured at that point, with respect to pre-ordained criteria.	Whereas measuring success at one point in time fails to capture the effects that play out over multiple timescales, and so tells us nothing about long-term success. Criteria for success change over time as we experience those impacts.	Temporal dispersal
There is a narrow conception of the roles of system designers, developers, users, owners, and regulators and their responsibilities, and there is a lack of agency of these actors in how they can fulfil these responsibilities.	Whereas sustainability imposes a distinct responsibility on each one of us, and that responsibility comes with a right to know the system design and its status, so that each participant is able to influence the outcome of the technology application in both design and use.	Responsibility
There is a tendency to think that taking small steps toward sustainability is sufficient, appropriate, and acceptable.	Whereas incremental approaches can end up reinforcing existing behaviors and lure us into a false sense of security. However, current society is so far from sustainability that deeper transformative changes are needed.	Transformative change



1.2 A visualization of systemic effects of one procurement system.

explores its use in pedagogy (Penzenstadler et al. 2018) and industry practice (Duboc et al. 2019).

In advocating for the importance of sustainability in SE and ICT, the initiative shifted the language toward a conception of "design" and highlighted professional responsibilities. Many of the principles and counterpoints reproduced above will resurface in later discussions of this book on the challenges of long-term choices in systems design and the values, politics and ethics of systems design. On their own, however, these counterpoints did not prove sufficient to enact transformative change, and it is unclear whether the initiative successfully shifted the wider community's perceptions. Instead, it appears that its impact in academic terms to date remains largely restricted to the communities of software engineering, ICT for sustainability, and requirements engineering.[27]

Still, in raising these counterpoints, and in emphasizing professional responsibility, the Karlskrona Manifesto made a significant step from research to advocacy, and its articulation of dimensions and effects is widely cited and has effectively provided a common definitional ground for researchers.[28] A range of academics and professionals have taken the opportunity to support the statement with an online signature, but its language speaks explicitly from and to a standpoint of software professionals: "As designers of software technology, *we are responsible* for the long-term consequences of our designs" (Karlskrona Initiative 2015). One might question the drawing of this boundary between those who claim responsibility and, by implication, those who do not or cannot, and chapter 5 will.

The conceptual framing of sustainability effects, and the arguments brought forward in the manifesto, carry an important consequence: central attention should be paid to the space between the technical and the social where stakeholders in technology development and design projects establish system boundaries and success criteria. "A software system's impact on its environment is often determined by how the software engineers understand its requirements. This impact's foundation is set in the decisions on which system to build (if any at all), the choices of whom to ask and whom to involve, and the specification of what constitutes success" (Becker et al. 2016, 57). In addition to choosing a system purpose, engaging with stakeholders and specifying success criteria, a whole range of activities take place in the design of any system, sometimes explicitly and often implicitly, that "reconcile the technical with the social" (J. Goguen 1994).

Whether explicitly attended to or implicitly performed, decisions about requirements arguably exert stronger leverage than specific techniques such as algorithmic approaches to energy efficiency:

> For example, techniques for increasing technical sustainability abound, ranging from architectural design patterns to documentation guidelines. Yet, because applying these techniques often involves an up-front investment of effort, it occurs only when a longer life expectancy of a system is recognized and expressed. On the other hand, a stated requirement for which no technique yet exists will lead to an identified gap in technological ability. This means that in practice,

> systemic changes to [requirements] activities will dominate the effects of whatever techniques we develop to support these activities. (Becker et al. 2016)

It is in the space of sociotechnical reconciliation represented by requirements that we can find room to act on issues of sustainability. Without a spokesperson who articulates a concern, technical approaches to address it will simply not be introduced into design practice. That places specific opportunities and responsibilities on those engaged in requirements practice. While the responsibility for requirements often does not constitute a separate role on a team, it forms a dedicated practice. What have professionals in computing made of it thus far?

A study probed the level of awareness and attention that current practitioners with this responsibility have, and the obstacles they see in the way of making sustainability a central concern in requirements practice (Chitchyan et al. 2016). The findings suggest an emerging awareness of the concerns regarding sustainability, identify a set of systemic barriers and obstacles, and group these to identify leverage points for practical change. Obstacles are identified on the level of individuals, their professional environments, and prevailing norms in software engineering practice.[29]

The study provides only a speculative outline of possible approaches, but the findings usefully highlight the potential for transformative change, the range of obstacles to be overcome, and the fact that interventions on all levels are required to make sustainability a central, accepted, and established concern in systems design practice. As the authors conclude,

> Significant barriers remain to [be] overcome before Software Engineering can claim to routinely advance not just technical and economic, but also social, individual and environmental needs simultaneously. Critical reflection is needed at the individual, organizational and community level to advance the profession's ability and commitment to do so. (Chitchyan et al. 2016, 541)

This reference to "simultaneously advancing" goals across the range of dimensions of sustainability, beyond technical and economic aspects, directly contrasts the professional roles and competencies of systems design professionals with an older, more established profession that has taken its role in sustainable development seriously for a while: engineering. In fact, the *UK Standard for Professional Engineering Competence* states that professional engineers must "undertake engineering activities in a way that contributes to sustainable development," including the "ability to [. . .] progress environmental, social and economic outcomes simultaneously" (UK Engineering Council 2014, 12). In addition, the licensing process places the burden of proof on professionals by expecting them to demonstrate they do so, for example by providing "examples of methodical assessment of risk in specific projects" or demonstrating "actions taken to minimise risk to society or the environment" (12). Unfortunately, this "explicit commitment to sustainability . . . is presently amiss within software organizations and their regulating and guiding bodies" (Chitchyan et al. 2016, 540).

A lot remains to do then to "shift the needle" on sustainability in ICT, as Mann, Bates and Maher (2018) call it. They examine the corpus of ICT4S proceedings in light of a "transformation mindset" for sustainable development, encapsulated in the following ten principled priorities.

1. Socioecological restoration over economic justification
2. Transformative system change over small steps to keep business as usual
3. Holistic perspectives over narrow focus
4. Equity and diversity over homogeneity
5. Respectful, collaborative responsibility over selfish othering
6. Action in the face of fear over paralysis or wilful ignorance
7. Values change over behaviour modification
8. Empowering engagement over imposed solutions
9. Living positive futures over bleak predictions
10. Humility and desire to learn over fixed knowledge sets. (Mann, Bates, and Maher 2018, 213)

Their stark conclusion is that ICT4S research is "unfortunately, insufficient to deliver a meaningful change" (Mann, Bates, and Maher 2018, 222). I share their hope that researchers will "examine their own research and ask themselves how they could contribute to a shifting of the needle towards ICT4S truly contributing to a positive socioecological transformation" (Mann, Bates, and Maher 2018, 222). The remainder of this book aims to encourage and facilitate this reflection.

PLANETARY BOUNDARIES, GROWTH, AND DECOUPLING

Much of the conversation around designing for sustainability continues to take place within a framing of sustainable development, inherently connected to the dominant economic paradigm of continued growth. The limits to growth argument made clear that endless growth is not possible. Its concept of limitations has evolved into a conceptual framework of "planetary boundaries" (Steffen et al. 2015; Rockström et al. 2009) within which human civilizations can operate without destroying the home they share with the rest of nature. Of the updated nine parameters—climate change, stratospheric ozone depletion, atmospheric aerosol loading, ocean acidification, biochemical flows, freshwater use, land-system change, biosphere integrity, and the introduction of novel entities—five were exceeded in 2022 (Persson et al. 2022).

In response to the recognition that the endless exponential growth of *material* consumption and destruction is clearly incompatible with finite resources, the proponents of what is now called *green growth* argue that IT offers the potential to *dematerialize* the economy and offer immaterial growth. In other words, it supposedly allows us to decouple economic growth from material consumption (resource decoupling) and environmental destruction (impact decoupling). For example, by increasing efficiencies, IT can reduce the material impact of existing activity; with improved monitoring and feedback, IT can support a better understanding and control of complex processes; and by substituting digital services for physical counterparts, as in video conferencing, IT can eliminate the need for material consumption (Royal Society 2020). This green growth argument has dominated economic policy discourse related to sustainable development in the European Union and elsewhere over the past decade and has substantively shaped policy priorities (Parrique et al. 2019; Hickel 2020; Hickel and Kallis 2020). For example, the European Commission's environmental policy for 2013–2020 targets an "an absolute decoupling of economic growth and environmental degradation" (Parrique et al. 2019, 10), and the 8th Environment Action Plan for 2021–2030 continues to pursue the vision that in 2050, "citizens live well, within the planetary boundaries in a regenerative economy where nothing is wasted, no net emissions of greenhouse gases are produced and economic growth is decoupled from resource use and environmental degradation" (European Commission 2020, 10).

But comprehensive reviews of the evidence behind decoupling come to stark conclusions. "The conclusion is both overwhelmingly clear and sobering: not only is there no empirical evidence supporting the existence of a decoupling of economic growth from environmental pressures on anywhere near the scale needed to deal with environmental breakdown, but also, and perhaps more importantly, such decoupling appears unlikely to happen in the future" (Parrique et al. 2019, 3). Parrique and colleagues assess two sides: the historical evidence of decoupling in the past and the logical argument for the feasibility of decoupling in the future. They cannot identify evidence on either side for large-scale, absolute decoupling—decoupling happens only in local, isolated cases, often remains a temporary phenomenon, and in fact, is often an illusion. Reports, for example, that highly developed countries are achieving material decoupling are identified as flawed conclusions: What these countries have achieved is an offshoring of the resource-intensive material extraction and production facilities on which their economies depend. They have externalized their environmental footprint to lower-income countries, yet continue to materially cause excessive resource consumption and material impact on the planet (Parrique et al. 2019, 21; Krausmann et al. 2017; Vadén et al. 2020). On aggregate global levels, the correlation between material use, material impact, and economic activity holds remarkably steady (Wiedmann et al. 2015; Hickel and Kallis 2020). Overall, "there is no empirical evidence supporting the existence of a decoupling of the type described as necessary . . . —that is an absolute, global, permanent, and sufficiently fast and large decoupling of environmental pressures (both resources and impacts) from economic growth . . . it is safe to say that the type of decoupling acclaimed by green growth advocates is essentially a statistical figment" (Parrique et al. 2019, 31). This assessment is confirmed by two large surveys assessing the evidence accumulated by 179 (Vadén et al. 2020) and 835 (Haberl et al. 2020) studies on decoupling.

To address the perennial counterargument that superior technology will bring about the elusive salvation, Parrique and coauthors move beyond historical evidence. In assessing the evidence of future feasibility, they identify significant barriers. Rebound effects form one of seven factors that will continue to prevent decoupling from becoming a reality. Again, the assessment is unequivocal: "we have found no trace that would warrant the hopes currently invested into the decoupling strategy." In other words, "green growth . . . is not possible" (Parrique et al. 2019, 10). The emerging *Degrowth* movement has forcefully shown that continued economic growth is not only unsustainable and destructive but also irrational (Hickel 2020; Kallis 2018; Demaria, Kallis, and D'Alisa 2015; Raworth 2017).

CONCLUSIONS

IT has a role to play in our societies' transformation toward a sustainable life form within the finite boundaries of planet Earth. Computational systems enable the coordination of cooperative and collective action, can *in principle* facilitate the partial decoupling of some economic activities from resource consumption under certain conditions, and offer new ways of living, working, and playing. But there is ample evidence that overall decoupling on an aggregate level is impossible now and in the future. So, IT development must be mindful of its role in perpetuating the exponential growth of material extraction and accept the importance of planetary boundaries.

As the history of technology shows, technology is "neither good, nor bad, nor neutral" (Kranzberg 1986). Nor is software "technology" an amorphous, shapeless whole—on the contrary. In practice, every systems design effort offers myriad moments where choices can work toward or against sustainability. Not all these choices are visible, and their effects, spread across time and space, may remain obscure, uncertain, and ambiguous. Despite its potential, the accounted effects of IT on sustainability hardly present a stellar track record. The best we may be able to claim is that the use of IT in climate modeling has allowed us to better understand the causal relationships between human activity and the climate crisis.

I want to draw attention to several aspects of the overview provided above. First, by and large, this work has focused on "how computing shapes societies" rather than examining the inverse or adopting an explicit view of mutual shaping and coevolution. In doing so, researchers rarely attend to the indirect, slower, but profound and lasting effects of their interventions. Despite this focus, researchers who argue for transformative change have acknowledged the need to better understand the social forces that shape computing, both in theory and in practice.

Second, the importance of *systems thinking* pervades many of the above fields. Different communities draw on different strands of the diverse branches of systems thinking, but the central attention in sustainability has been on system dynamics, used in the first World Model underpinning the *Limits to Growth* report (D. H. Meadows 2008; D. L. Meadows and Club of Rome 1972; D. H. Meadows, D. L. Meadows, and Randers 2004; Wiek, Withycombe, and Redman 2011). When Easterbrook (2014a) argues that a shift is needed from computational thinking to systems thinking, he primarily refers to system dynamics too. But the debate between the perspectives represented in figure 1.1 reflects the epistemological shift that underpins the break from hard to soft systems thinking in the 1970s (Checkland 1981), and the further break with soft systems thinking performed by critical systems thinking in the 1980s and beyond has not been fully considered (Flood and Jackson 1991c). Chapter 5 will return to the rich history of systems thinking in search of a critical perspective on its role in systems design.

Third, I have identified *requirements* as a central locus of attention with strong leverage over the outcomes of design projects. Because requirements shape technology development at the space where the technical meets the social, they influence the outcomes greatly. They offer significant room for actors in the design process to maneuver and substantial opportunity for innovation in practice. But existing requirements theory and practice appear ill-equipped to tackle the challenges thrown by sustainability.

Finally, the need for transformative change and critical reflection pervades many of the conversations. Those who take a step back to assess the state of work in IT typically come to stark conclusions. And, as the next chapter will show, these are only the tip of the iceberg. Not all is well in computing.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Becker, Christoph (Director of the Digital Curation Institute), author.
Title: Insolvent : how to reorient computing for just sustainability / Christoph Becker.
Description: Cambridge, Massachusetts : The MIT Press, [2023] | Includes bibliographical references and index.
Identifiers: LCCN 2022038283 (print) | LCCN 2022038284 (ebook) | ISBN 9780262545600 | ISBN 9780262374651 (epub) | ISBN 9780262374668 (pdf)
Subjects: LCSH: Electronic data processing—Social aspects. | Computer systems—Environmental aspects. | Information technology—Social aspects. | Sustainable development.
Classification: LCC QA76.9.C66 B435 2023 (print) | LCC QA76.9.C66 (ebook) | DDC 303.48/34—dc23/eng/20221121
LC record available at https://lccn.loc.gov/2022038283
LC ebook record available at https://lccn.loc.gov/2022038284