Fake Data Leads to Fake Insights: The Challenges of Prototyping Energy Dashboards

Christina Bremer
c.bremer@lancaster.ac.uk
Lancaster University
Lancaster, United Kingdom

Christian Remy
c.remy@lancaster.ac.uk
Lancaster University
Lancaster, United Kingdom

Adrian Friday
a.friday@lancaster.ac.uk
Lancaster University
Lancaster, United Kingdom

ABSTRACT

In efforts to combat climate change, reduce an organisation's carbon footprint, and ultimately save resources, energy management has become an important aspect of non-domestic building management. The practice of energy managers can involve analysing vast quantities of energy-related data, mostly quantitative, to assess current demand and forecast future consumption of often complex sets of buildings to identify targets for efficiency measures. In recent years, we have worked closely with energy managers to create better tools for them to not just uncover energy savings potential, but find better ways of representing and understanding this data. In doing so, we have faced several unexpected challenges, holding back the creation of prototypes, with problems unique to this specific application domain. We reflect on why this has proven so difficult and point out the challenges of the prototyping process for energy management. We use key HCI methodology to explain what we believe the reasons for these challenges are, and discuss potential ways to overcome them, offering recommendations for future researchers to more effectively engage with energy management.

KEYWORDS

Energy Dashboards, Prototyping, Human-Computer Interaction

1 INTRODUCTION

A significant footprint of businesses and organisations relates to the energy consumption of their infrastructure, such as the electricity, gas, and water usage of buildings and the people living and working in them [13]. The opportunities of modern technology in terms of gathering information about energy consumption through increased resolution and frequency of sensing as well as computing advances such as dashboards, AI, and digital twins, i.e., virtual representations or replica of physical products, systems, or processes, offer promising avenues for reducing this footprint. In reality, these opportunities are rarely fully explored and realised. While there is a plethora of research into visualising energy consumption [29], most of the focus is on the domestic, whereas the public and commercial sector is relatively under-explored. The reasons for this are manifold: the difficulty and complexity of stakeholder engagement [32]; the different drivers behind energy managers' work [15], making it more difficult to design for [2, 28]; and challenges in the design process that ask for more fundamental rethinking [1, 17].

Many of those challenges we have experienced first hand, as we have conducted our own research in the broader field of sustainability and energy management in various projects over the last 5–10 years with a range of stakeholders and data sources. In our most recent project, dubbed Net Zero Insights[1], we aim to analyse exactly the sort of complex multi-building data to gain insights into how organisations can reduce and transform their energy consumption. We aim to provide tools for them to make more informed decisions using contemporary statistical and computational techniques, to help arrive at a more sustainable future. Our approach includes an interdisciplinary team and data from a range of sources, both qualitative and quantitative. In the process we have encountered many unexpected roadblocks that slowed down our efforts. In this paper, we revisit some of these problems to highlight the limits of energy management research and share our lessons learned with the LIMITS community.

In particular, we focus on the barriers to conduct work in our own core discipline, user experience (UX) and HCI. One of the initial goals we set was to explore the current practices and tools being used, and redesign them to better help triage the growing volumes of data, making repeated analysis to find energy savings more tractable. For this, we worked closely with stakeholders, including energy managers and facility managers, who oversee the installation, configuration, and maintenance of energy-related infrastructure. In this paper, we focus on our work with the campus's energy data, as it is most comprehensive, our relationships are most developed, and thus our experiences are as holistic as possible. We outline our approach and the obstacles to reaching our desired outcomes, in particular problems around data sources and quality, and how these have translated into challenges for the prototyping process so vital for our HCI practices. We finish with a series of suggested strategies for future researchers engaging in this space how to tackle the domain appropriately. Our key takeaways for research into energy management are to (1) embrace interdisciplinarity and participatory design, (2) design iteratively, (3) create known valid subsets of real data for prototyping and demonstrations, (4) use stakeholders' existing tools if possible, and (5) empower users and foster critical thinking. While the takeaways at this high level would be reasonable guidelines for many projects across domains, Section 5 describes how they unfold in not-always-obvious ways during the prototyping of energy dashboards. We also want to emphasise that we use the term "fake data" in this paper as a synonym for synthetic data—as opposed to data that is deliberately misleading (as could be assumed from the connotation of the word in fake news).

2 BACKGROUND & RELATED WORK

In non-domestic and commercially managed environments such as office buildings, energy settings are usually maintained by a

LIMITS '24, June 18–19, 2024, Online
2024.

[1]https://www.net0i.org/

central organisational facility, commonly referred to as energy managers or facilities managers. These types of roles typically have several responsibilities, including maintenance and adjustment of infrastructure to deliver gas, water, heating, and electricity reliably and affordably to their estate. Modern technology, such as sensing and controlling units, displays, feedback mechanisms, controllers, and automation, are of crucial importance for those stakeholders to make informed decisions and go about their daily responsibilities. As Beucker and Hinterholzer noted, ICT can play an important role in helping create a more sustainable future, such as towards decarbonising buildings [3, 4]. Deploying sensors meaningfully in those buildings [25], as well as being able to monitor and analyse collected data [27] is an important effort to better support such objectives.

Data collection at LIMITS has so far mostly looked at the urban scale, such as Hamm's collection of civic data to improve air quality [18], or Bhardwaj et al. [5] and their insightful lessons on data curation for city-scale data collection, pointing out the importance of a rigorous design process that takes the community on board to enable user empowerment. Ringenson et al. [34] report on the limits of a smart sustainable city, concretely the difficulty of implementing and maintaining digital infrastructure—a key issue that we observed in our work on the building and organisational scale as well, and also discussed more broadly for technology in general at LIMITS (e.g., Rasoldier et al. [31]). And in a study of water consumption specifically, Pargman et al. [28] highlight the complexity of designing ICT for data collection and analysis purposes, in particular the complexities of design, and the different views of stakeholders, including policymakers.

Engaging with policymakers is an important aspect when considering the design of ICT for building systems management, as energy managers are often pulled in different directions by trying to create more comfortable surroundings for the inhabitants and users of buildings, sometimes in direct response to handling complaints. They are often underresourced and at the same time are under pressure to save money (e.g., Goulden and Spence [15], Beucker and Hinterholzer [2]), especially in the recent global energy context. This comes as no surprise as the training of energy managers and the general guidelines for starting investigations into energy use often starts with translating energy consumption into money spent [16]. Therefore, research into how to navigate the landscape of policymaking is vital to gain an understanding of the complexities and limitations [23], including the less visible implications of so-called invisible energy policies [14, 35], which describe non-energy policies with unacknowledged or insufficiently acknowledged impacts on energy demand. This is the complexity of designing for energy managers; there is more than one party of stakeholders involved, often competing goals and limited resources, and important research includes stakeholder engagement as well as organisational decision-making [32, 39].

An early survey by Pierce et al. [29] provides an overview of eco-feedback visualisations, although the focus here lies within the domestic context. This is a trend that continued into more recent research, where efforts into public contexts can be found, such as thermal comfort in managed apartments [12] or studies of the individually perceived differences of thermal comfort [11], but a majority of the work remains focused on the home (e.g., [20, 21, 36, 40, 41]. One reason for this might be that studies of energy management in larger organisational contexts encounter a multitude of additional complexities and challenges not normally found in the home [1, 22]) or are work in progress [10, 37]. Most recently, Gregg and Strengers questioned the usefulness of common dashboard designs for energy management [17], and the LIMITS community has put forward more general design guidelines for ICT for sustainability [7, 42]. Our work aims to build up on those lessons learned and, based on our own experiences, provide concrete recommendations for the LIMITS community how to more effectively engage with and design for energy stakeholders.

3 OUR WORK WITH ENERGY MANAGEMENT STAKEHOLDERS

A major part of the work in this paper stems from observations made in our project Net Zero Insights. The goal of this project has been to analyse statistical data regarding energy demand in non-domestic buildings of organisations at a larger scale and generate insights in how to reduce energy consumption towards Net Zero. While we have several commercial stakeholders as partners, we will focus in this work on the data we know best from the Lancaster University campus, coupled with unique insights from facilities and management from our long-running collaboration with energy managers on this estate.

Over the past decade, the campus system has seen several changes and updates in technology to sense and control energy supply and demand, as well as the usual cycles of change in maintaining any large and complex estate infrastructure [9]. This means that there are buildings from the foundation of the campus in the 1960s to complete new-builds. Many of the older buildings have seen a retrofit/refurbishment and several changes of use with the ebb and flow of organisational evolution at some point. While there is a central district heating system, it does not feed all buildings, sometimes due to "value engineering" (i.e., minimising costs) preventing connection to it as new buildings are built and delivered. There is no single standard across all buildings for energy-related infrastructure, be it radiators, vents, insulation, or sensors, although generally newer buildings have a higher density of in-built sensing and metering than older estate. Building standards, like the buildings themselves, also evolve over time affecting the construction and performance of each building. Similarly, much of the technology behind the sensing infrastructure to log, capture, and process any resource consumption data or other building information has been installed and updated over the decades as new technology becomes available and old technology becomes obsolete or is no longer supported or meets modern constraints (e.g., for cyber-security or network capability). To further complicate the sensing landscape, various projects by either the facility management or associated research projects across different University groups have started to deploy new concepts, and not always finished to a state that is robust enough for years of maintenance by the facilities teams. Certainly, extensive use is made of external contractors, and even the core staff involved often move into new roles or new employment over time; so the knowledge concerning these projects is often lost. This has resulted in a piecemeal of heterogeneous infrastructure and data, making it difficult to navigate and maintain, or even fully understand. While this is to some extent exacerbated by the fact that the organisation is also simultaneously a teaching and research facility, offices, shops, and accommodation; we know from conversations with energy stakeholders in the commercial sector that it is not unique and this kind of inconsistency and difficulty to arrive at a streamlined energy data solution is a common occurrence in energy systems environments.

To better understand this difficult landscape within the organisation and arrive at insights towards energy savings for the project as well as other research work carried out by the authors, we engaged in a series of qualitative and quantitative data gathering activities. This entailed interviews, focus groups, surveys, and observations across the campus site and with stakeholders on all levels of the organisation, from higher-level management, to facility managers and engineers, to supporting and research staff frequenting the buildings. All activities were carried out in a mixed-methods approach depending on the nature of enquiry, from in-depth interviews, informal coffee chats, or even just emails and instant messages. As part of the first author's PhD work, a series of expert interviews were conducted to gather different stakeholders' perspectives on and experiences with digital interventions for energy savings, including the use of energy data, interaction with building occupants, and the tailoring to the existing digital infrastructure. Building on this work, another study focused on the impact of COVID-19 on the overall energy consumption of the campus, researching the potential for energy savings in changing people's behaviour. Besides quantitative analyses of both campus-wide and building-specific energy data, this study also involved a focus group interview and energy data visualisations annotation exercise with experts.

3.1 Examples of the problematic data

Among the various data streams, one that we focused on was high frequency meter readings[2]. This data comprises consumption of electricity, heating, gas, and water supply sampled as time-series. This totals to longitudinal data from almost 1,500 meters, with a data point supplied every ten minutes. Some meters report their results in rate format (count of energy units used per unit time), while most yield cumulative readings of the count of total units used since some starting point. For analysis purposes we turn cumulative data into rate data ourselves. One commonly used approach to do this, also supplied as a built-in mechanism in most dashboards and building management systems, is to calculate the "non-negative difference", i.e., subtract the current from the previous reading for each data point to get the change in energy use for each time step. While in theory this works great, real-world energy data provides plenty of unexpected data challenges. Firstly, energy data is rarely clean since it relates to complex real-world networked infrastructures, with significant opportunity for error. For any gaps in logged data, the non-negative difference naturally is significantly higher (the higher the longer the gap between readings), resulting in large spikes in the data. It is not straightforward to distinguish these unintended spikes from "real" outliers such as a peak in electricity consumption (cf. Blazquez-Garcia [6]).

An example for this can be seen in Figure 1, which shows a typical line graph of energy consumption of a single campus building (in kWh). This problem is even more complicated for naturally non-monotonic readings such as water use that naturally spike based on demand and patterns of use, making it more complicated to detect outliers based on extreme values. While solutions exist to condition the data and look for outliers, these can require manual intervention and data treatment, such as storing information about identified gaps in a separate database and removing the first value following any gap in the transformed data stream. However, due to the sheer size and quantity of data (thousands of channels, millions of data points) such processing can introduce delays in returning from API or database calls, causing lags and a reduced responsiveness impacting UX; not even talking about the inevitable potential for error for any manual data cleaning process. This is also just one of many issues in the data pipeline; others are that sometimes loggers hit their limit and "roll over", e.g., after reading 99999 the next value returned is 0, or numbers from loggers communicating via serial protocols might be truncated due to communication errors resulting in a series of data points resembling this: 3415, 3416, 34, 3418. And this is just the tip of the iceberg!

Another problem with energy data and its use arose in one of our user testing sessions. One of the commonly requested views of energy consumption is to normalise it based on the size of the building, e.g., divide energy consumption by square meter/square foot of the building (totalled across all floors) to look at the ratio between energy use and area comparatively. In a hackathon-style activity we did this to show to the relevant stakeholders, see Figure 2. The discussion that followed was not as we intended around the insights gained from seeing this graph of the building's performance, or centred on the visualisation format (which was a fairly "quick hack" prototype)—instead, the focus became entirely about finding out why one building was topping the charts (an unexpected outlier by an order of magnitude). To some extent there was an "insight" generated from this, as the outlier had gone previously unnoticed. It was clear that this was an outlier though and confirmed by the energy manager because the building is one of several similar buildings and all readings show similar patterns, just had not been investigated before in such a comparative manner. Due to the sheer amount of data pulled from various sources into this graph, it was not trivial to identify the possible source of the error. What followed was an investigation into the past of the associated logger and whether it was faulty (failing, badly calibrated, a data pipeline or software issue, etc.; a conclusion has yet to be made at the time of writing). From an HCI perspective—an unfulfilled prototyping session regarding potential UX lessons learned!

3.2 Goals and ambitions

An important point to emphasise are the different goals of parties involved on either side of this project. We, as sustainability-motivated researchers, are interested in saving energy, and so are the energy management stakeholders. However, the underlying reasons could not be any more different: our interest is in helping create a more sustainable future, and be able to share the road to this future with a wider audience of academic and energy stakeholders to then hopefully adapt and benefit from our learnings as well. A bit of a caricature, but energy managers also seek savings—but have a clear responsibility first and foremost to keep services running and to

[2] e.g., from data loggers such as https://synetica.net/datastream/

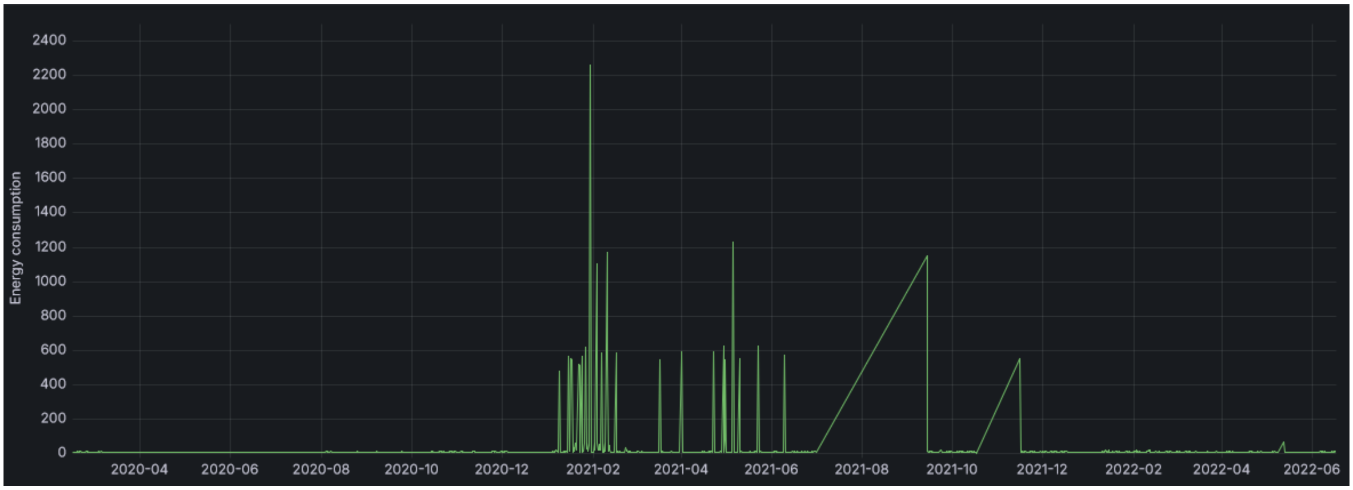

Figure 1: Line graph of raw energy consumption of a typical building. The spikes in August 2021 and 2022 come from outages, elevating the first reading after the loggers came back online. The spikes in April 2019 are from errors in the data pipeline. Removing such faulty readings is not always trivial and difficult to automate, and in less obvious instances those spikes are hard to differentiate from other anomalies; but without removing such spikes, the real data is almost uninterpretable as it results in a flatline.

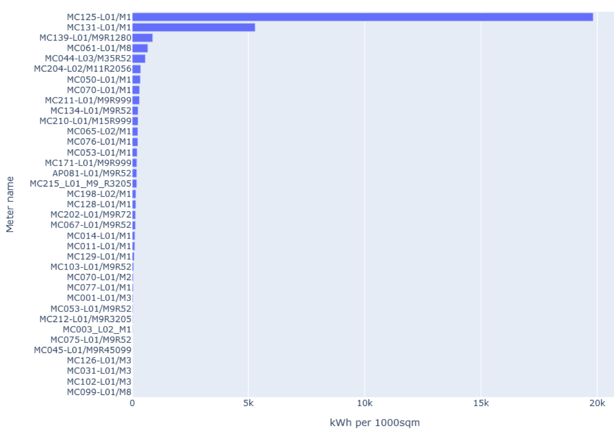

Figure 2: Overview of energy intensity of a typical day's consumption, in kWh normalised over square meters. The top meter shows an outlier, presumably due to an erroneous logger, dwarfing other results, even making some buildings appear to have zero values, and derailing the discussion during user testing.

save money. This is not because they are cold-heartedly or narrow-mindedly focused on nothing but monetary savings, but simply because it is often their most important job requirement as set by their superiors, and core to their training for the role. The energy consumption of large institutions, such as a university campus, is so massive that even tiny percentage savings can result in significant financial savings, making room to invest in other improvements to the estate. In particular during the energy crisis, saving money on energy spending *became crucial* for businesses to stay viable.

That said, many energy managers also see it as their goal to make changes towards a more sustainable future and may well have ambitious longer-term plans to decarbonise their estate, requiring significant investment and programmes of infrastructure change. In particular for businesses, it becomes more and more important to not just save money, but also take initiatives to save the environment, e.g., as part of the environmental reporting guidelines for publicly traded and other large companies[3].

4 PROTOTYPING CHALLENGES

During our work on energy dashboards, and in particular with energy managers, we encountered the following six key challenges:

(1) Energy managers do not always know what they need but often have urgent priorities. It is commonly known in HCI that users do not necessarily know what they need or want [26]; a consideration that is reflected in the infamous quote attributed to Henry Ford: *"If I had asked people what they wanted, they would have said faster horses".* The same applies to energy managers, who might not "have room" to think beyond the strategies they are currently deploying or may have somewhat idealised expectations regarding the extent to which energy data and dashboards could support their work (not least promulgated by software companies high on promises eager to sell them digital solutions!). Across our research studies, we have observed a tendency among some energy stakeholders to (incorrectly) believe that taking the human "out of the loop" is key to solving their problems, leading them to focus on tuning building parameters and restricting the control available to occupants. One stakeholder even argued that occupants *"could be a problem"* for successful energy system automation. What adds to the challenge is that energy managers might not fully or accurately

[3]https://www.gov.uk/government/publications/environmental-reporting-guidelines-including-mandatory-greenhouse-gas-emissions-reporting-guidance

understand the problems they are facing based on the data alone. In one instance, unusual meter readings in an office were not as it first appeared due to technical difficulties with the sensing, but rather to the operation of a new high-performance workstation with GPU accelerators paired with the occupant's preference for high office temperatures; in the absence of this important contextual information, those who are responsible might invest in digital tools, data, or strategies to then find that they cannot solve this underlying problem through the application of more technology.

(2) (Accurate) data might not be available. Energy dashboards and other digital solutions depend on high-quality data—which often is not available. Among the problems we encountered were the inconsistent formatting of data, a large number of missing data points due to an unnoticed database storage problem, as well as the non-existence of relevant datasets (see above). What makes this challenge even more profound is that scale of the required data grows proportionally to the size of an estate and the assets that are contained within it (e.g., controllable devices, measurable things). Secondary data such as weather data is also required for a range of energy management analyses such as degree day analysis, a technique for understanding energy required for heating or cooling from outdoor temperature to a given setpoint temperature. The building management system (BMS) on our Lancaster University campus, for example, collects about 70 million data points per month; including other, non-BMS sources those add up to over a billion data points every year. For such a large amount of data, a simple outlier removal script does not exist; and if it did, it would likely run over a long period of time, and its results would be almost impossible to verify for accuracy by hand. Assuming that a verified format of the required data exists, its inaccessibility can pose another problem: across our studies, we were confronted with difficulties relating to limited data storage, failing pipeline components, formatting and configuration errors, mislabeled data sources, surprise network configuration and IP address changes, and cycling of dynamically assigned universally unique identifiers (UUID) to data loggers that change over time (e.g., if a device reboots or after a power outage). All of these challenges must be resolved before it becomes possible to even start testing any off-the-shelf solutions or external scripts.

(3) Synthetic data leads to "synthetic" results. As outlined above, real-world data can be unavailable, inaccurate, and is normally complex to understand and explain. To limit the negative consequences of this challenge, is it common practice to use synthetic data, also informally referred to as fake data, for the design and evaluation of dashboards and user interfaces (UIs), meaning that the data has been artificially generated rather than taken from real-world sources [30]. For our energy research we also tried adopting this approach. However, while the availability of synthetic data enabled us to present dashboard prototypes to our stakeholders, we had to acknowledge its limitations in allowing the stakeholders to give us useful feedback. Similar to how unspecific prompts in interviews (e.g., *"Please describe a typical week."*) often lead to responses that reflect memorable highlights or socially desirable practices, synthetic data in dashboards gathers results for an idealised, but non-existing environment. To counter response biases, prompts, questions, and data need to be precise, trustworthy and contextualised—for interviews *and* dashboards. This is especially relevant as small changes in energy consumption or generation can indicate larger problems or trends, which need to be interpreted carefully. For example, when analysing the impact of COVID-19 lockdown measures on the electricity consumption of selected campus buildings, we found the maximum drop in energy demand due to the absence of building users tended to be between 20% and 23%. In comparison, the realistic impact of digital or behaviour change interventions supported by energy dashboards (outside of lockdown periods) are thought to be significantly lower than this, and thus relatively small compared with "always on" equipment relating to teaching and research—which should lead to more fundamental questions about what is driving energy demand [14]. The generation of synthetic data often relies on assumptions and simplifications, which itself introduces biases that do not reflect real-world conditions. For example, if the synthetic data is based on the assumption that energy consumption patterns are uniform across different campus buildings, it might not accurately represent the diversity of actual energy usage (which depend on a variety of factors, including building characteristics, policies, mixed building uses, and occupant behaviour). Synthetic data fails to capture such nuances, leading to inaccurate insights and, eventually, unusable dashboards.

(4) Dashboard data is often not put into real-world organisational context. When a suitable energy dataset for a dashboard prototype has been identified, this immediately gives rise to yet another challenge, i.e., the lack of data contextualisation within the prototype to help explain it. The majority of energy dashboards, including ours, only display quantitative data, e.g., energy consumption, cost, and efficiency metrics. While such data is crucial for tracking performance and identifying consumption trends, its presentation in isolation without relation to the broader organisation context can be problematic, as we have repeatedly seen stakeholders struggle to interpret the significance of the patterns in the data they are presented with. For instance, a dashboard displaying a spike in energy consumption might prompt investigation, but often needs contextual information such as external weather conditions, operational or practice changes, new building uses, without which the underlying cause of the possible anomaly will likely remain unclear. From an HCI perspective, the lack of explicit contextualisation also means that the needs of building occupants can get overlooked—which can lead to frustration for them and, ultimately, unexpected ways of them trying to regain control, e.g., using electric heaters within their offices, even though they are not supposed to. This is not to say that the contextual information necessarily does not exist somewhere in the organisation, it just is not typically recorded and integrated within the dashboard or may not even be formally or electronically recorded in an accessible way. We have seen a lot of relevant knowledge being held informally in e.g., energy managers' heads, stand-alone Word and Excel documents, or being distributed across "notes" columns of spreadsheets in unstructured ways. This knowledge is critical for the accurate interpretation of energy data but often is only available to a select few who are well-integrated into their organisations and departments and have worked there for significant time periods. What follows is that when such a staff member leaves, their in-depth knowledge, i.e., contextual information, is usually lost as well.

(5) Dashboard integration often remains an afterthought. Once a stand-alone dashboard has been designed and developed, it can be difficult to meaningfully integrate it with existing digital tools, systems, and databases on the one hand and workflows and business processes on the other. For this reason, it is important to understand the landscape of systems and workflows that are used in the stakeholders' organisations, and to develop a realistic plan as to how the energy dashboard will fit once it is ready to be deployed. This thorough integration increases the stakeholders' ability and thus likelihood to use the dashboard for their everyday work while enabling them to better interpret the data. For example, when they stand alone, the bar charts of an energy dashboard that compare buildings with concrete line graphs of meter readings for a certain time period can be difficult to make sense of. But combined with, for example, a map that allows the user to "zoom in" and "zoom out", see other meters of similar days in surrounding buildings, or during other days/weeks, it becomes possible to understand if an insight is just an anomaly, outlier, or even just broken data. This also became obvious during one of our user tests, in which we showed selected buildings: the energy manager struggled to identify similarities as the visualisations were static rather than dynamic, connected, and coherently linked. Only integrated dashboards will fully allow their users to understand the complexity behind the consumption patterns they see; and the more seamless the integration with their current tools and working practices, the less resistance can be expected from the user group—who understandably are rarely keen to deviate from their organisation's digital ecosystem, their work routines, and established business processes.

(6) Stakeholders have surprising faith in digital tools without necessarily evidence of success. During our work with energy managers, energy consultants, and those responsible for energy policy, we have often been met with visualisations, dashboards, and Excel spreadsheets. This seems to be the default starting point for this collection of stakeholders. Their lens is one of accounting for energy and cost, and related key-performance indicators (KPIs), which can lead to a particular chain of enquiry, focusing on numbers supported by fact finding visits. In addition, a majority of the stakeholders we have worked with are often surprisingly optimistic about the energy-saving potential of digital efficiency-related interventions, in particular energy system automation, as a key to solving their problems. They often imagine a future in which these interventions are used to seamlessly save energy without as much manual intervention as their current systems require. While such interventions undoubtedly have the potential to improve the energy efficiency of the systems they are intended for, the experts' vision does not seem to relate to actual positive real-world experiences of such technologies. In fact, evidence of the efficacy of these seems rather thin on the ground. Many stakeholders even directly report acute challenges with regard to past digital efficiency projects; these challenges range from the regularly required maintenance and high risk of failure to insufficient data storage, lack of staff training, and difficulties interpreting the data. But these experiences do not seem to lower their faith in technological solutions. One of them described it as follows, *"I mean my idea is [...] that we would be able to tie everything together, to room booking, to occupancy, to sensors [...] all done without really anyone intervening with it. I'd like to think we can get to a point as well where we predict a building."* This somewhat optimistic stance aligns with a well-established concept in the literature that describes the unjustified belief that technological progress will fix existing problems [24]. While dashboards are not equivalent to algorithmic controls or necessarily engineered solutions, the data they display can be inaccurate, incomplete, complex, and hard to interpret, and so it will be essential to encourage critical reflections among their users.

5 SUGGESTED STRATEGIES

In order to tackle the six key challenges outlined above, we propose the following strategies for the design and evaluation of energy dashboards; an overview of recommendations is also provided in Table 1.

Do	Don't
• Carry out participatory design	• Assume users know what they need
• Design iteratively to meet requirements	• Gather feedback only from a vocal minority
• Create subsets of real data for prototyping	• Base user studies on synthetic data
• Make use of easy-access tools	• Leave dashboard integration until the end
• Empower users and foster critical thinking	• Ignore stakeholders' different objectives

Table 1: Overview of recommendations for the design and implementation of energy dashboards.

(1) Embrace interdisciplinary research and participatory design. While this has been a well-established pillar of many successful HCI projects [19, 38], we want to highlight it again in this context because it is critical: we need to bring stakeholders of all levels on board early on and keep them engaged during the design process; this includes energy managers as well as people higher up in an organisation's hierarchy, such as C-level managers, policymakers, directors, but at the very least the line managers that energy managers report to. This approach will drastically improve the chances of having a direct line to policymaking impact and the needed backing to ensure such systems are integrated and supported longer-term. However, it is equally important not to forget the people below the management, e.g., the engineers working for energy managers, occupants, people living in or frequenting the buildings that are being analysed. A third group of people who need to be included are data scientists and analysts, who can help make sense of the data; what should be kept in mind is that their goals in creating dashboards are often fundamentally different (i.e., they are looking for statistical insights such as change points and anomalies). Their expertise will be needed to confirm the quality of the data, clean, and access it. And above all, rather than simply recruiting them for individual studies, it is advisable to include them as participatory design members for a truly democratic design process [8].

(2) Design iteratively based on verified requirements. There are many steps to be done here: first, we need to acknowledge that it needs an HCI understanding of investigating what is really needed rather than straight up implementing the concrete desires as hurriedly expressed by the stakeholders (as described above, the complexity, lack of contextualisation, and use of synthetic data in dashboard prototypes can make it difficult for users to articulate what they need). With this in place, feedback should be gathered early and repeatedly (we found that many people will be hesitant to share their thoughts—or it may be politically sensitive to do so; or sometimes it is a vocal minority who engages in that process), and then implemented properly. As is the case in many HCI projects, an iterative design cycle is required as the first prototype is rarely the best solution.

(3) Create subsets of real and credible data. To avoid the distraction or even deception of users by synthetic data, real data should be used whenever possible. As many organisation will not have complete sets of accurate and cleaned data available for all time periods, we advise selecting relevant subsets that fulfil those criteria. This might mean to focus on shorter time frames, verified energy meters, or to identify periods of (seemingly) fault-free meters for prototyping. This way, researchers can have confidence in the data they display and the users' attention can be directed towards the tools/dashboards they are evaluating rather than the errors in the data. If the objective of the study is to identify outliers, patterns etc., we would recommend to add them afterwards and start with a test run with more regular data to ensure that the design of the tool or dashboard is based on data points that are not outliers, rather than on an incoherent dataset. Afterwards, additional subsets could be included that contains anomalies or outliers if searching for those in the visualisations is a desirable part of the user testing. This requires having a subset of the data ready to pipe into other tools, such as via CSV downloads or sample API calls, to be able to quickly get data for a prototyping session. The idea is, similar to hackathons, to have a working subset of data available that researchers can flexibly deploy (c.f., GreenHackathon[4]); once they are done prototyping or have confirmed that their tools work as intended, they can always add more real data and start dealing with gaps and errors.

(4) Use easy-access tools. A key concept behind prototyping is to create versions that are easy to change and adapt, and a strategy to achieve this is to use tools that already exist and try to amend them. For example, if an organisation uses power BI, Grafana, or a Jupyter notebook based dashboard, it can be helpful to make changes to those and start prototyping from there. This has the benefit that researchers do not get caught up in the process of establishing a new data pipeline or trying to fit data into a shape that is not compatible with current working practices and tools. At the same time, it has the benefit that it lets energy managers or other stakeholders relate to parts of the digital infrastructure that they are familiar with, to then focus on the changes and overcome the novelty effect. Using easy to access tools allows researchers to gather relevant feedback from working with real data. In other words, researchers should make use of existing tools as much as possible and innovate only the part they really want to test on, i.e., avoiding the temptation to try to change everything for users!

[4]https://www.greenhackathon.com/

(5) Empower users and foster critical thinking. Interpreting energy data accurately requires critical thinking skills, and while it can be difficult to look beyond the objective appearance of quantitative dashboards, users need to understand that consumption data does not stand alone. Here, researchers should encourage users to not blindly trust what technology presents them with, but to question it. Similarly, users should be encouraged (and enabled) to share and record contextual information and informal data and observations, so this can be integrated into the dashboards and shared within the organisation. There should be a central location where it can be stored, updated, and accessed, to avoid misinterpretation of consumption data and a situation in which staff members take relevant knowledge with them when they leave the organisation. Good design can help to encourage sustainable practices like these and simplify the process of exchanging contextual information. We are in progress to work on such a contextual storage solution, for example using ontologies [33].

6 CONCLUSION

Dashboards are a useful utility for viewing summary data and energy management analytics; they can offer forecasts and help signal about the quality of the energy system. However, their clean, factual appearance can easily conceal the complexity and uncertainty that are inherent to energy data, systems, and consumption practices, conveying a sense of objectivity and robustness that does not in fact match the reality. If we have learned one lesson from our research then it is that working with energy data is as important as it is messy. At each step of the process—including data gathering, storage, recall, visualisation, and interpretation—a number of challenges await. This is not to say that researchers should avoid working with energy data and dashboards—understanding and reducing energy demand is a critical challenge, but that we need to reduce and communicate complexities to maximise useful engagement. There are clear limits to the degree of complexity that can be meaningfully handled at any given time and it has proven easier to gather data than it is to handle it meaningfully, so we must find the right balance of reducing it while not oversimplifying things. In order to achieve this, we have found that limiting the scope of datasets and novel technologies while deepening the engagement with users can pave the way towards usable, useful, and contextualised energy dashboards.

ACKNOWLEDGMENTS

We thank our funders, EPSRC (Grant ref. EP/T025964/1) and the Leverhulme Centre for Material Social Futures Research (DS-2017-036).

REFERENCES

[1] Mohammed Alhamadi, Omar Alghamdi, Sarah Clinch, and Markel Vigo. 2022. Data Quality, Mismatched Expectations, and Moving Requirements: The Challenges of User-Centred Dashboard Design. In *Nordic Human-Computer Interaction Conference*. ACM, New York, 1–14.

[2] Severin Beucker and Simon Hinterholzer. 2019. Building Energy Management Systems and their Role in the Energy Transition. In *International Conference on ICT for Sustainability (ICT4S)*. ACM, New York, 8.

[3] Severin Beucker and Simon Hinterholzer. 2020. Effects of ICT-Enabled Flexible Energy Consumption on the Reduction of CO2 Emissions in Buildings: Findings and outlook from an explorative case study in Germany. In *International Conference on ICT for Sustainability (ICT4S)*. ACM, New York, 1–10.

[4] Severin Beucker and Simon Hinterholzer. 2022. Decarbonization Potential of Building Automation: Results of a German Study and Outlook. In *International Conference on ICT for Sustainability (ICT4S)*. IEEE, New York, 119–127.
[5] Eshta Bhardwaj, Han Qiao, and Christoph Becker. 2023. Pathways to urban sustainability: Design perspectives on a data curation and visualization platform. In *Proceedings of the Ninth Workshop on Computing within Limits*. LIMITS, New York, 13. https://limits.pubpub.org/pub/iw8v1y2m.
[6] Ane Blázquez-García, Angel Conde, Usue Mori, and Jose A. Lozano. 2021. A Review on Outlier/Anomaly Detection in Time Series Data. *ACM Comput. Surv.* 54, 3, Article 56 (2021), 33 pages.
[7] Eli Blevis, Chris Preist, Daniel Schien, and Priscilla Ho. 2017. Further Connecting Sustainable Interaction Design with Sustainable Digital Infrastructure Design. In *Proceedings of the 2017 Workshop on Computing Within Limits* (Santa Barbara, California, USA). Association for Computing Machinery, New York, NY, USA, 71–83.
[8] Susanne Bødker and Morten Kyng. 2018. Participatory design that matters—Facing the big issues. *ACM Transactions on Computer-Human Interaction (TOCHI)* 25, 1 (2018), 1–31.
[9] Stewart Brand. 1995. *How buildings learn: What happens after they're built.* Penguin.
[10] Christina Bremer, Oliver Bates, Christian Remy, Alexandra Gormally-Sutton, Bran Knowles, and Adrian Friday. 2023. COVID-19 as an Energy Intervention: Lockdown Insights for HCI. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, New York, 1–7.
[11] Adrian K Clear, Samantha Mitchell Finnigan, Patrick Olivier, and Rob Comber. 2018. ThermoKiosk: Investigating roles for digital surveys of thermal experience in workplace comfort management. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. ACM, New York, 1–12.
[12] Marie Denward, Annelise De Jong, and Rebekah Olsen. 2015. Energy matters in buildings: Individual and collective issues. In *EnviroInfo & International Conference on ICT for Sustainability (ICT4S)*. Atlantis Press, Paris, 172–179.
[13] Helen Freeman and Lorna Christie. 2021. Reducing the whole life carbon impact of buildings UK Parliament POST brief POST-PB-0044.
[14] Alexandra M. Gormally, Kirstie O'Neill, Michael D. Hazas, Oliver E.G. Bates, and Adrian J. Friday. 2019. 'Doing good science': The impact of invisible energy policies on laboratory energy demand in higher education. *Energy Research & Social Science* 52 (2019), 123–131.
[15] Murray Goulden and Alexa Spence. 2015. Caught in the middle: The role of the Facilities Manager in organisational energy use. *Energy Policy* 85 (2015), 280–287.
[16] Jessica Granderson. 2011. Energy information handbook: Applications for energy-efficient building operations.
[17] Melissa Gregg and Yolande Strengers. 2024. Getting beyond Net Zero dashboards in the information technology sector. *Energy Research & Social Science* 108 (2024), 103397.
[18] Andrea Hamm. 2020. Particles Matter: A Case Study on How Civic IoT Can Contribute to Sustainable Communities. In *International Conference on ICT for Sustainability (ICT4S)* (Bristol, United Kingdom). Association for Computing Machinery, New York, NY, USA, 305–313.
[19] Gillian R Hayes. 2011. The relationship of action research to human-computer interaction. *ACM Transactions on Computer-Human Interaction (TOCHI)* 18, 3 (2011), 1–20.
[20] Björn Hedin, Victor Larsson, and Henrik Artman. 2018. Indoor temperature awareness using an Ambient Information Display-a semi-longitudinal study of one household.. In *International Conference on ICT for Sustainability (ICT4S)*. EasyChair, online, 112–124.
[21] Christian Hinrichs, Michael Sonnenschein, Adam Gray, and Curran Crawford. 2015. Self-organizing demand response with comfort-constrained heat pumps. In *EnviroInfo & International Conference on ICT for Sustainability (ICT4S)*. Atlantis Press, Paris, 353–360.
[22] Timo Jakobi and Gunnar Stevens. 2015. Energy saving at work-and when not working! Insights from a comparative study. In *EnviroInfo & International Conference on ICT for Sustainability (ICT4S)*. Atlantis Press, Paris, 180–189.
[23] Somya Joshi, Teresa Cerratto Pargman, Andreas Gazis, and Daniel Pargman. 2016. Whose future is it anyway? Limits within policy modeling. In *Proceedings of the Second Workshop on Computing within Limits* (Irvine, California). Association for Computing Machinery, New York, NY, USA, Article 5, 6 pages.
[24] James E Krier and Clayton P Gillette. 1985. The un-easy case for technological optimism. *Mich. L. Rev.* 84 (1985), 405.
[25] Cheng Li, Forrest Meggers, Mo Li, Jithendrian Sundaravaradan, Fei Xue, Hock-Beng Lim, and Arno Schlueter. 2013. Bubblesense: wireless sensor network based intelligent building monitoring. *International Conference on ICT for Sustainability (ICT4S)* 13 (2013), 159–166.
[26] Jakob Nielsen. 1993. *Usability Engineering*. Academic Press, Cambridge, MA.
[27] Faris Nizamic, Tuan Anh Nguyen, Alexander Lazovik, and Marco Aiello. 2014. GreenMind-An architecture and realization for energy smart buildings. In *International Conference on ICT for Sustainability (ICT4S)*. Atlantis Press, Paris, 20–29.
[28] Teresa Cerratto Pargman, Somya Joshi, and Uta Wehn. 2019. Experimenting with Novel Forms of Computing: The case of the Swedish Citizen Observatory for Water Quality Conservation. In *Proceedings of the Fifth Workshop on Computing within Limits* (Lappeenranta, Finland). Association for Computing Machinery, New York, NY, USA, Article 8, 10 pages.
[29] James Pierce, William Odom, and Eli Blevis. 2008. Energy Aware Dwelling: A Critical Survey of Interaction Design for Eco-Visualizations. In *OZCHI*. ACM, New York, 1–8.
[30] Trivellore E Raghunathan. 2021. Synthetic data. *Annual review of statistics and its application* 8 (2021), 129–140.
[31] Aina Rasoldier, Jacques Combaz, Alain Girault, Kevin Marquet, and Sophie Quinton. 2022. How realistic are claims about the benefits of using digital technologies for GHG emissions mitigation?. In *Proceedings of the Eighth Workshop on Computing within Limits*. LIMITS, online, 14.
[32] Christian Remy, Oliver Bates, Vanessa Thomas, and Matthew Broadbent. 2018. "Sustainability... it's just not important.": The Challenges of Academic Engagement with Diverse Stakeholders. In *International Conference on ICT for Sustainability (ICT4S)*. EasyChair, online, 16.
[33] Christian Remy, Adam Tyler, Paul Smith, Oliver Bates, and Adrian Friday. 2024. Wasted Energy? Illuminating Energy Data With Ontologies. *IEEE Pervasive Computing* Early Access Article (2024), 1–10.
[34] Tina Ringenson, Elina Eriksson, Miriam Börjesson Rivera, and Josefin Wangel. 2017. The Limits of the Smart Sustainable City. In *Proceedings of the 2017 Workshop on Computing Within Limits* (Santa Barbara, California, USA). Association for Computing Machinery, New York, NY, USA, 3–9.
[35] Sarah Royston, Jan Selby, and Elizabeth Shove. 2018. Invisible energy policies: A new agenda for energy demand reduction. *Energy Policy* 123 (2018), 127–135.
[36] Maha Salman, Steve Easterbrook, Samar Sabie, and Josie Abate. 2016. Sustainable and smart: Rethinking what a smart home is. In *International Conference on ICT for Sustainability (ICT4S)*. Atlantis Press, Paris, 184–193.
[37] Paul Shabajee, Chris Preist, Daniel Schien, John Brenton, and Chris Jones. 2016. Diagnostic visualisation of building management system energy data. In *International Conference on ICT for Sustainability (ICT4S)*. Atlantis Press, Paris, 233–233.
[38] M Six Silberman, Lisa Nathan, Bran Knowles, Roy Bendor, Adrian Clear, Maria Håkansson, Tawanna Dillahunt, and Jennifer Mankoff. 2014. Next steps for sustainable HCI. *Interactions* 21, 5 (2014), 66–69.
[39] David Stefan and Emmanuel Letier. 2014. Supporting sustainability decisions in large organisations. In *International Conference on ICT for Sustainability (ICT4S)*. Atlantis Press, Paris, 333–341.
[40] Seyed Amin Tabatabaei. 2016. A data analysis approach for diagnosing malfunctioning in domestic space heating. In *International Conference on ICT for Sustainability (ICT4S)*. Atlantis Press, Paris, 11–18.
[41] Seyed Amin Tabatabaei, Dilhan J Thilakarathne, and Jan Treur. 2014. Agent-Based Analysis of Annual Energy Usages for Domestic Heating based on a Heat Pump. In *International Conference on ICT for Sustainability (ICT4S)*. Atlantis Press, Paris, 104–111.
[42] Huaxin Wei, Jeffrey C. F. Ho, Kenny K. N. Chow, Shunying An Blevis, and Eli Blevis. 2019. Should Do, Can Do, Can Know: Sustainability and Other Reflections on One Hundred and One Interaction Design Projects. In *Proceedings of the Fifth Workshop on Computing within Limits* (Lappeenranta, Finland). Association for Computing Machinery, New York, NY, USA, Article 6, 18 pages.