

Review

Carbon-Responsive Computing: Changing the Nexus between Energy and Computing

Dawn Nafus ^{1,2,*}, Eve M. Schooler ³ and Karly Ann Burch ² 

¹ Intel Labs, Intel Corporation, Santa Clara, CA 95054, USA

² Centre for Sustainability, University of Otago, Dunedin 9016, New Zealand; karly.burch@otago.ac.nz

³ Next Generation and Standards, Intel Corporation, Santa Clara, CA 95054, USA; eve.m.schooler@intel.com

* Correspondence: dawn.nafus@intel.com

Abstract: While extensive research has gone into demand response techniques in data centers, the energy consumed in edge computing systems and in network data transmission remains a significant part of the computing industry's carbon footprint. The industry also has not fully leveraged the parallel trend of decentralized renewable energy generation, which creates new areas of opportunity for innovation in combined energy and computing systems. Through an interdisciplinary sociotechnical discussion of current energy, computer science and social studies of science and technology (STS) literature, we argue that a more comprehensive set of carbon response techniques needs to be developed that span the continuum of data centers, from the back-end cloud to the network edge. Such techniques need to address the combined needs of decentralized energy and computing systems, alongside the social power dynamics those combinations entail. We call this more comprehensive range “carbon-responsive computing,” and underscore that this continuum constitutes the beginnings of an interconnected infrastructure, elements of which are data-intensive and require the integration of social science disciplines to adequately address problems of inequality, governance, transparency, and definitions of “necessary” tasks in a climate crisis.

Keywords: smart grid; demand response; social aspects of energy; carbon footprint; carbon intensity; edge computing; data center; big data; linked data; mixed qualitative and quantitative methods



Citation: Nafus, D.; Schooler, E.M.; Burch, K.A. Carbon-Responsive Computing: Changing the Nexus between Energy and Computing. *Energies* **2021**, *14*, 6917. <https://doi.org/10.3390/en14216917>

Academic Editors: Ben Anderson, AbuBakr S. Bahaj, Thomas Rushby, Patrick James and Carlos Henggeler Antunes

Received: 30 July 2021

Accepted: 6 October 2021

Published: 21 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With growing digitization of human activities, greenhouse gas emissions (GHG) from information and communication technologies (ICTs) are increasingly a matter of concern. While the energy efficiency of hardware continues to improve, and a partial turn to renewables for powering computing systems ensues, these trends might not be able to keep pace with the explosive growth in demand for digital services [1,2]. Estimates of ICTs' share of global electricity consumption vary widely and transform rapidly with changing technology and changing demand for digital products [3]. However, there is a narrower band of estimated greenhouse gas emissions. A review of peer-reviewed estimates shows most researchers place ICTs at 1.8–2.8% of current global emissions, putting it roughly on par with the aviation industry [4]. Belkhir and Eleligi [5] predict ICTs' relative GHG contribution could rise to 14% by 2040, approximately half of the total GHG emissions of the entire transportation sector.

In this context, computing systems are increasingly looking to reduce their emissions impact through increased energy efficiency, whether in data centers (especially given their extensive cooling needs) [6,7], in data transmissions [8,9] and the use of local or co-located renewable energy sources [10]. However, a series of more radical responses are emerging (i.e., delaying, redistributing, stretching or shrinking computing workloads) [11], as a function of real time electricity carbon intensity (technically defined as the amount of carbon by weight emitted per unit of energy consumed, but used here to connote the percentage of energy from non-renewable sources). As a form of demand response,

temporal and spatial shifting of workloads are optimizations that could: (1) reduce the overall carbon footprint of ICT devices; (2) further incentivize the transition to renewables by creating demand for “spare” renewable energy; and (3) support grid providers’ needs for manageability, therefore reducing the need to expand energy generation capacity, whether renewable or not. We term this new approach carbon-responsive computing (CRC) and such systems could substantially increase the participation of the ICT industry in demand response.

Carbon-responsive computing is the umbrella term we use to describe a range of techniques that currently exist under various other names. For example, in a hyperscale data center (i.e., large warehouse-size data centers typically owned by the largest technology companies), carbon responsiveness has been referred to as geographic load balancing [12], smart grid integration [13], or time shifting [11]. However, as we show below, limiting CRC to data centers misses important opportunities in other areas of computing. Therefore, in carbon-responsive computing, as we define it, energy sources and computing elements of all kinds—PCs, networking devices, electric vehicles, Internet of Things (IoT) devices, large or small data centers—collaborate with the shared goal to prioritize the usage of energy with the least carbon intensity, regardless of whether the strategy is spatial, temporal, or a combination.

This umbrella term enables researchers to consider such techniques holistically, across a range of different types of computational infrastructures and energy generation infrastructures and in light of the sociotechnical implications that arise as these techniques mature. Without this holistic picture carbon-responsive computing could be built in siloed ways that prevent it from becoming a more substantial energy-computing infrastructure, and it could be designed in ways that exacerbate economic exclusion and invite greenwashing. We aim to prevent that outcome by supporting an interdisciplinary dialogue across computer science, energy research, and social studies of science and technology (STS). With this aim, this paper reviews relevant literature across energy research, computer science, and STS—disciplines that do not always communicate—to arrive at a more socially aware research agenda for carbon-responsive computing.

To summarize the argument, carbon-responsive techniques are fairly well understood in the context of hyperscale data centers, where they have been initially applied. There is great interest in data center applications because data centers have grown massively in recent years [6,12], and much of their workloads can tolerate delays to flexibly adapt to match the variability of renewables [13]. However, changes in both energy infrastructure and computing infrastructure make a data center-centric research agenda insufficient. Specifically, rapid growth in edge computing—that is, computing that happens more proximate to the very edges of the network infrastructure—means more computational power, storage, networking, control, and data management is situated outside big data centers [14,15]. As we discuss below, there is some engineering research on implementing carbon response within edge systems, though the approach is less commonplace. This constitutes a particularly problematic research gap, because renewable energy generation is also growing in distributed ways [16]. The proliferation of renewables at the edges of the electrical grid amplifies opportunities to align the energy use of computing tasks both locally and non-locally through energy–compute exchange systems.

The problem, however, goes far beyond simply developing optimization techniques in edge systems in ways analogous to those found in data centers. With growth outside data centers also taking place, and CRC techniques beginning to evolve in that direction, the stage is set for CRC to transform from a series of related but discrete technologies into a more cohesive and interdependent set of underlying social and technical concerns to address. That evolution requires that researchers acknowledge urgent social considerations often overlooked in efficiency-focused technical approaches. Non-sociotechnical approaches treat carbon responses as a matter of a single company conducting demand response with a single energy provider using a bespoke suite of algorithms. With a sociotechnical approach, CRC considers carbon responsiveness to be a matter of wider public

interest: it is taking place within heterogeneous sociotechnical systems with multiple actors involved and various technical and social dynamics that demand attention. Less powerful actors, such as owners of small-scale renewable energy systems, become newly implicated in CRC, and we cannot assume the interests of all actors are easily aligned.

We argue that this is likely to come to a head with respect to the underlying data that enables carbon responses, which serves as a shared resource for heterogeneous actors. Not only is more research required to better specify the technical requirements of shared data infrastructure to meet the needs of different ICT systems, but social scientific research is also necessary to address societal questions that arise in the design of that infrastructure. These questions concern social equity, inclusion, transparency, governance, the role of markets, environmental impact, and ask what constitutes the need for a computing task to proceed at all, or in a different form, in the context of a climate crisis. We find that many of these more fundamental questions are only answerable with a more holistic approach that includes integration with STS.

In this paper we qualitatively connect the dots across fields and subfields not traditionally considered together, to make the case for a broader sociotechnical research agenda that lays a foundation of equitable inclusion in this emerging way of allocating energy and computing power. Where available, we leverage pre-existing reviews of the state-of-the-art for this purpose. We proceed as follows. In Section 2, we present the trajectory of carbon-responsive computing from early experimentation in data centers through to new areas of application such as edge-based systems, and we identify the areas of innovation that become newly feasible given the increasingly distributed nature of renewable energy as well as computing infrastructure [17]. In Section 3, we employ the classic sociotechnical model of infrastructure growth from Hughes [17] to reason that data, particularly carbon intensity data, are likely to be a newly important enabler that arises from the early growth we trace. We then outline a synthesizing framework for CRC that provides high-level guidance for other researchers in identifying the implications of their activities in one area for another. In Section 4, we delve into the key challenges that must be addressed by this interdisciplinary community—researchers and industry—in a sociotechnical manner. Lastly, we conclude by identifying areas of research in CRC that must be expanded to address broader societal challenges.

2. State of the Art: Foundations for Carbon-Responsive Computing

One widely recognized problem in the transition to renewables is the variability in renewable energy production, as renewable energy is only produced when, for example, the sun is shining or the wind is blowing. Times of peak demand, such as winter evenings when heating is most needed, can also require energy generation capacity that in all other times of the year would constitute an overprovisioning. Even a small offset of peak demand can mitigate the need for such overprovisioning [18]. Computers do not have the same biophysically driven temporalities as wind and sun. They can schedule and distribute their operations with tremendous plasticity. Armed with the right data insights, computing's spatial and temporal plasticity can turn into a significant advantage for sustainability initiatives, as systems can be optimized to favor the availability of renewable energy.

Section 2.1 begins by describing the contributions that data center research has made to the development of CRC, highlighting their contributions to technology optimization and the building of relationships between carbon-intensive data centers and smart grids. In Section 2.2, we discuss the evolution of CRC beyond data centers, describing the drivers for this evolution, and what is at stake when CRC is used in edge systems. In Section 2.3, we describe the changing relationships between computing and energy which directly impact CRC practices. In Section 2.4, we end with a discussion on the state of the art enabling data that can support the expansion of CRC within heterogeneous sociotechnical systems.

2.1. Data Centers and Carbon-Responsive Computing

2.1.1. Rationale

Not surprisingly, data centers have been leaders in both the scholarship and the industry in adopting carbon-responsive techniques. Energy is a significant percentage of the cost of operation, and a portion of a data center's workload can tolerate delays that can be programmatically controlled in response to price or carbon signals [19,20]. What is not acknowledged in the engineering literature, but central in the STS scholarship on data centers, is the public contestation over data centers' impacts on energy and water systems [21]. Deeper concerns over "energy gentrification" [22] have also developed—that is, concerns about energy providers and policymakers prioritizing the energy needs of global firms at the expense of local citizens, raising questions of democracy and what constitutes the public interest [23]. For these economic, societal, and environmental reasons, there are enormous incentives to more tightly manage energy use.

2.1.2. Optimization Technologies

Wierman et al. [19] offer an extensive survey of the data center optimization literature, noting that the relevant scholarship is scattered across disciplines. Nevertheless, it is worth raising a few salient aspects that have been foundational. Initial work focused on schedulers, estimating available near-term solar energy, and balancing this with grid energy prices and jobs' deadlines [24], maximizing the use of on-site renewables [25] while mitigating impacts on availability for tasks inappropriate for scheduling [26]. Goiri et al. [27] note the challenges of integrating on-site renewables, not only in simulation, but also in the real world, including underappreciated topics such as dealing with contractors unfamiliar with demand response and the temptation to overprovision. Other optimizers specifically aim to balance both direct energy costs and monetary incentives provided by a regulator to participate in demand response to reduce overall costs while meeting service level agreements [20]. Because the largest data centers are owned by firms with an expansive geographic presence, there has been an emphasis on geographic load balancing where data center placement of workloads follows patterns of sun and wind around the globe rather than seek to locally optimize. Dou et al. [28] observe that energy costs to cool are related to temperature, and therefore cooling needs, and not just on-site renewables production, must be a factor in decisions about where to run workloads. Liu et al. [29] conducted a numerical experiment using workload data from Hewlett Packard Labs and corresponding energy mix data and found that to successfully space-shift workloads across different data centers, significantly more wind production (80% of the renewables mix) was required, and small-scale on-site energy storage remains useful. They note an important tradeoff in attempting to "follow the renewables" [29] (p. 234). In routing workloads to a remote data center with cheaper energy, the transmission delay and energy usage might increase, and thus the data center might need to complete the job faster and thus use more energy overall. Indeed, one challenge for CRC is right-sizing both on-site electricity generation and storage.

More recently, at least one major cloud provider (Google) began time-shifting non-urgent workloads in its data centers for maximum renewables availability, claiming in 2020 that, "carbon-aware load shifting works" [30]. Their method involves matching externally provided day-ahead hourly forecasts of carbon intensity with internal forecasts of computing task demand over the same time period. In the public sector, the DC4Cities project was able to shift 5–30% of data center workloads participating in their trial of a federated system of carbon-response across multiple data centers in a city [31]. These examples demonstrate that it is possible to bring working systems into production, but how commonplace they are in practice is unclear from publicly available information.

2.1.3. The Relationship with Smart Grids

Where there are still unknowns is how these hyperscale data center methods integrate with the evolving definition of the smart grid. Whereas one of the first steps toward a smart grid was to deploy smart meters that periodically monitored utility customers' whole-house energy usage, over time the term smart grid has come to mean an electric grid infrastructure imbued with the ability to measure and monitor its operation along many dimensions, at many different places in the topology, and to perform sophisticated AI algorithms to predict, manage, and optimize its performance.

Thus, where do data centers and smart grids intersect? Data centers are enormous customers of electricity and as cited earlier have introduced CRC techniques to offset their carbon footprint with renewable energy, where possible. Of course, this is only a useful technique when the carbon intensity of the electricity consumed by Internet cloud service providers is low, which is a function of where they are sited [32,33]. In both [13,19], the authors note that the relationship with smart grids is ambiguous with respect to pricing mechanisms, the lack of appropriate regulatory frameworks, and the potential for data center grid customers to use their purchasing power to game markets, market complexity, and control points. For example, if the grid supplier makes energy supply available based on the carbon intensity of that energy, that is a very different social and technical proposition from supplying infinite access to energy and allowing devices to choose how much they want to use, depending if the energy is renewable or not. Lin et al. [12] examine the impact on the grid in terms of renewables absorption and power prices under three conditions: (a) data centers running at constant load regardless of the status of the grid; (b) data centers conducting local time-shifting; and (c) grid-wide optimization, where the grid controllers set power levels of multiple data centers across a grid, and those data centers adapt to the power available. They show that the grid-wide optimization can increase a grid's renewables fraction and reduce the data center's energy costs, but could create large changes in data center capacity that could be unacceptable to data center operators.

Since the Wierman et al. survey [19], proposals for pricing and exchange mechanisms have emerged in various quarters, primarily from simulations. Some examples include online auctions to incentivize tenants to delay batches in a co-located data center [34], a bargaining approach specifically designed to overcome the information asymmetries between utilities and data centers [35], and an energy marketplace for exchanging both energy flexibility and thermal waste heat [13]. The latter work productively incorporates the synergistic effects of cooperating multiple actors and networks across a smart city, including scenarios where data centers can exchange workloads between themselves to optimize energy costs, and scenarios where smaller data centers collaborate to maximize the amount of green energy available. This is an excellent example of the higher-level perspective that we argue will be required as carbon responsiveness grows from one-off implementations to broader adopted practice. Nevertheless, these areas of inquiry only scratch the surface of the social dynamics of adoption, which we address in Section 4.

2.2. Carbon-Responsiveness beyond Data Centers

2.2.1. Drivers for Expanding CRC

However, while data centers receive the bulk of attention, they do not necessarily account for the majority of ICT energy consumption [36]. Various types of networking equipment (from the core backbone, to on-premise consumer-facing access points, to the xG wireless access networks) constitute approximately 45% of total energy consumption, while consumer data centers account for 36%, equipment, electricity associated with equipment production makes up another 8%, and device usage 2% (Figure 1).

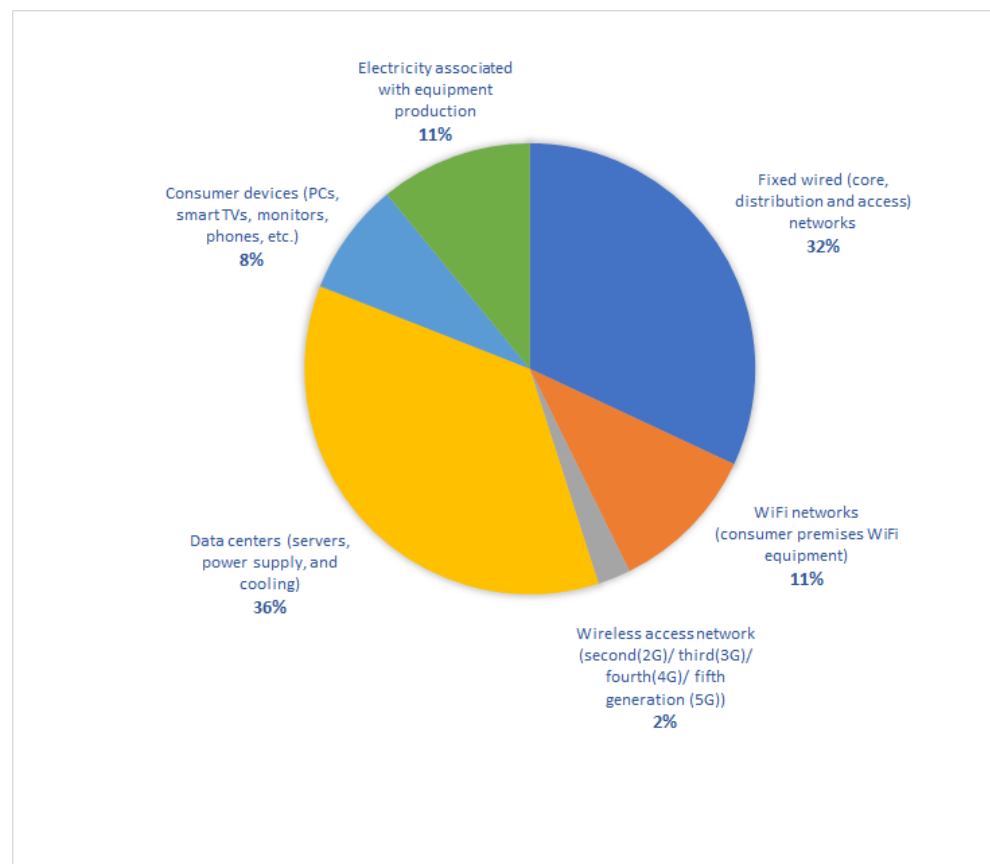


Figure 1. Projected 2030 distribution of ICT energy consumption (percent of total terawatt hours (TWh)). Created from data published in: Lorincz, Capone, and Wu [36].

Dubbed the “*overlooked environmental footprint of increasing Internet use*” [37], the various Internet networking equipment that transmit data around the world, onto phones, televisions, and computers, stand to have as much if not more of an impact on carbon footprint than data centers. How energy consumption translates to carbon footprint varies by country, due once again to where the network is drawing electricity; Obringer et al. [37] estimate a range of 28–63 CO₂e per GB of data, depending on the country. Estimates in ICT energy overhead typically involve large uncertainties regarding data transmission [3], and Aslan et al. [38] found estimates of data transmissions energy consumption spanned five orders of magnitude due to variation in system boundaries, age of data, and other assumptions. Even with these uncertainties, Figure 1 nevertheless illustrates an important underlying disjuncture between where research effort has been most prolific for carbon responsive computing (data centers), and where additional energy consumption might be found and harnessed to balance renewables integration. One unknown is the amount of CO₂ emissions savings that are possible in different scenarios from carbon responses. For now, we can reason that as long as even the lowest estimates of ICT emissions are non-zero [3,4,36,37], and as long as demand for data continues to outpace equipment efficiency gains [3], there is ample reason to continue to apply carbon responsiveness across the spectrum of ICT systems, given the urgent need for climate action.

Although Figure 1 identifies that networks rival data centers in electricity usage, it does not capture new developments in edge computing, which introduce a proliferation of resources outside the back-end cloud data center. In other words, the “elsewhere” beyond the data center is growing. Indeed, Gartner predicts that by 2025, 75% of all data will be created and processed in edge computing systems [14]. Estimates predict a 24% compound annual growth rate (CAGR) for edge data centers by 2025, as upwards of 100,000 edge-based data centers will be deployed [15]. While it would be unwise to take

any prediction of an emerging technology at face value, such claims do reflect a changing sensibility within the industry about the technologies worth investigating and building. Well-designed interactions between the local edge and remote back-end data center have the potential to reduce energy consumption for a given system [39], but whether those will be implemented in real-world systems remains to be seen. Edge growth in aggregate could mean more energy-consuming devices overall if they ultimately add to, not displace, data center growth. With such growth there is a need to control edge-related carbon emissions as rigorously as data centers often do.

While devices that live at the network edge might include cameras, phones, Internet of Things (IoT) sensing devices, PCs, cars, the edge itself is often a reference to a collection of resources that do not reside in a back-end hyperscale data center cloud, but instead reside locally or proximately. These resources may or may not rise to the level of a data center, but they are not a remote cloud in the sense of being a large scale, far away hyperscale provider offering outsourced computation cycles and storage as a service [40]. Some might nevertheless be “cloud-like” in the sense that they might be built as a shared resource or service. For example, a telecom company’s or other provider’s points of presence will typically have server storage alongside routers and other networking equipment, and these often cache content for quicker delivery to end users [41]. The beauty of the edge is that there are so many potential places where an edge might reside and where its placement can offer efficiency benefits for certain classes of data-centric applications.

With the tremendous growth in the numbers of devices connecting at the very edges of the Internet, and the torrent of data that they create, edge computing has growing appeal for the technology industry [42]. Tasks that formerly could only take place in a data center, such as training AI models, can now take place elsewhere, as computational resources are distributed widely and new techniques such as federated learning (a way of distributing the training process of AI models) make edge computing a viable alternative for those tasks [43]. Data processing, including AI, increasingly even happens in the network itself, not just in servers and on clients [44]. Many tasks once outsourced to a back-end remote data center cloud have requirements that render that choice no longer appropriate. The main reasons for this shift [45,46] tend to be data volume (it is impossible to fit all the data over the network pipe, at least in its original form), latency (the cloud is too far away), and privacy-sensitivity or regulatory/jurisdictional considerations that limit data movement (such as with HIPAA in the US, or GDPR in the EU). For example, smart city data might need cleansing at the network edge to ensure no personally identifiable information is stored or leaked, and that pre-processing takes computing power.

2.2.2. Carbon-Responsiveness at the Edge

Given the energy consumption of both data transmission and data center cooling, we expect sustainability, or at least energy cost, to be added to the rationale for building edge systems. Although there is a recent survey of the state of the art of energy considerations in edge computing [47], tellingly, its focus is not on renewables, carbon, or real-time carbon responses, but instead the traditional energy efficiencies of specific edge hardware, architecture, software, and networking design. Workload scheduling is addressed in terms of overall efficiency, not in terms of capturing the least carbon intensive energy. There is some discussion of power capping, where firm limits are placed on the energy consumed by a piece of hardware regardless of workload, and more discussion of techniques for task offloading to the cloud in order to save battery and other local resources. The authors note that an inefficient offloading strategy can consume more energy and have high associated carbon footprints. They also discuss that controllers that offload computational tasks typically either aim to maximize total throughput or to minimize the total energy consumption, and they call for a hierarchical approach where “*energy-intensive tasks should be offloaded to cloud servers, computing-intensive (CI) tasks should be offloaded to edge servers, and data-intensive (DI) tasks should be offloaded to servers that are close to the data source*” [47] (p. 575). We can extrapolate from this that the offloading controllers that orchestrate this

joint optimization might be in a position to include in their decision-making processes consideration of carbon intensity.

There are, however, a few studies within the edge computing literature that explore carbon responsiveness. Li et al. [26] offer a framework and a prototype for carbon-responsive computing at the edge, where the main energy grid is treated as a supplemental non-renewable energy source to locally produced renewables. Forecasting of solar and wind availability 24 h in advance enabled battery charging optimizations and a classifier distinguished between delayable and time-sensitive tasks. Another study [39] explored the energy impacts of different configurations of an IoT sensing system with an edge gateway (i.e., data processing capacity larger than a PC but smaller than a data center) powered by a microgrid (a decentralized group of electricity resources that is able to function independently from the traditional grid). They included different ways of relating these systems to the cloud, with one scenario where cloud processing was used only when local renewable energy was unavailable. Another [48] proposed the close integration of edge computing systems and local microgrids, to service time-elastic computations, whilst considering IoT workloads of a time-sensitive nature. A fourth study compared the full energy consumption, including network data transmission consumption, across four types of architectures, spanning from fully centralized to fully decentralized [49]. It concluded that a completely distributed architecture might save 25% of energy consumption through avoiding intra-data center network traffic and reducing the need for large-scale cooling systems.

Put together, this group of studies suggests that carbon responses at the edge are indeed nascent if underdeveloped. Although these systems are not in real-world production beyond academia, given the opportunity for carbon savings as shown in Figure 1, we expect them to emerge. One significant unknown is the extent to which algorithms used in data centers to schedule and distribute computing tasks can be repurposed for edge data centers, given the differences in scale, computing infrastructures, and types of tasks involved at the edge.

2.2.3. Edge Systems and Energy Markets

Another, perhaps deeper, unknown is whether the exchange and pricing mechanisms proposed in data center contexts are equally useful at the edge, or whether entirely different kinds of marketplaces or forms of exchange need to be created. Proposals for methods of exchanging renewable energy at the edge have emerged recently, without considering energy consumed by computing. For example, Mengelkamp et al. [50] proposes individual level peer-to-peer energy market trading for rooftop solar energy using blockchain. Vergados et al. [51] propose “prosumer clustering” of individual rooftop solar owners into virtual microgrids composed algorithmically, in order to participate as a single entity in the energy marketplace to reduce costs. Paladin et al. [52] showed through a simulation study that in high density, distributed solar generation with battery storage, a micro-energy market could reduce costs for all grid users by 4–20%. The next step for research, then, is to explore whether there are additional improvements that could be made by integrating carbon response techniques alongside other forms of demand response, although as we outline in Section 4, there are significant social factors at work that need to be taken into account beyond the current emphasis on efficiency, carbon savings, and cost. However, it is also quite promising that peer-to-peer relationships and federated groups of like-minded edges are emerging. This contrasts with the status quo, where large cloud service providers sequester data and services. Thus the impact of this shift to smaller regions of administrative control and to cooperation across administrative domains should be studied further.

2.3. *The Changing Relationship of Energy and Computing: What Is at Stake for CRC?*

When we think of the relationship between computing and energy, we tend to think about smart grids that have employed computing traditionally to measure, monitor, and ultimately optimize energy delivery. The growth of edge computing introduces a new dy-

namic into this relationship. Distributed renewable generation networks, whether rooftop solar, microgrids, or other types of configurations, are growing [53,54], leading to efforts to localize energy distribution and control perceived by practitioners as advantageous for resilience, efficiency, addressing social equity and justice, and for responding to local resource needs [53]. Edge computing, by definition, is similarly decentralized. Because there are energy and compute costs in processing data remotely, the decentralization of both energy and computing raises an important question: how should these two trends come together, both in engineering terms and social terms? What new types of combined energy–compute systems are possible? To underscore the significance of such combinations, we use the term “edge-ification” [55] which refers to moving resources away from large production centers and toward the edges of the infrastructure, whether energy creation, storage, or computational resources.

Data centers have begun to edge-ify, either in their siting choices that have incorporated considerations of access to cheap renewable energy, such as Facebook and Google’s presence in eastern Oregon, which offers plentiful hydropower, or in choices about building supplementary on-site renewable power [25]. More problematically, cryptocurrency miners have been relocating to areas of cheap renewable grid energy, creating problems for local electricity companies, who are forced to invest in last mile infrastructure improvements for high-demand users who can easily move their equipment to the next cheap electricity hotspot, saddling the local electricity provider with useless investments [56]. Data centers’ size means they benefit from being sited near energy generation megaprojects such as hydroelectric dams or gigantic wind farms. Yet for edge systems, the scale of necessary renewable energy is more modest.

Translating the techniques used by hyperscalers to the edge requires acknowledging that there are many kinds of edges, ranging in size and distance either from the devices themselves or from back-end warehouse sized data centers, forming an edge-cloud continuum [57–59]. For example, roadside units (RSUs) can couple data storage, compute processing, and data transmission with on-site solar or wind energy and energy storage of some kind [59]. Edges also are being sited at base stations to support cellular networks, and in POPs (points of presence) or COs (central offices) of the Telco infrastructure. In the former case, they are being opportunistic about the availability of a place to site services [60]. In the latter case, workload consolidation and virtualization have led to many fixed function machines being replaced by fewer high-end servers, leaving the spaces virtually empty and ripe for additional compute/storage to site a new edge [61]. Mobile phone infrastructures sometimes have on-site power generation, whether diesel or renewable, particularly in the Global South [62]. In these examples, public rights of way create physical opportunities for edge equipment installation. Larger industrial sites, city properties, and universities could offer similar opportunities for co-located, co-evolving systems, as these are places that are more likely to have either edge systems in place, and/or on-site renewable energy generation. These types of edges, and their particular requirements, have not yet been explored as sites of carbon-responsive computing.

Edge-ification also creates possibilities for data centers to resemble, or interact with, edges in new ways. For example, Yang et al. [63] propose siting modular data center-like units at each turbine of a wind farm to take advantage of reduced energy transmission losses and “free cooling.” This is a telling phrase used in data center contexts, which betrays just how far away from the natural world computing systems have come to be. In any other field, “free cooling” is simply called “air”! While Yang et al. [63] propose making the data center similar to the edge, Nurminen et al. [64] propose an interaction with the edge by redistributing the computational tasks that would normally take place in a data center and running them in homes where there are rooftop solar panels. Based on solar production data in Finland, and the associated spot price for cloud computing services during that same time, the authors’ economic viability analysis showed such an approach was feasible, and could introduce local economic benefits that could help overcome objections to “energy gentrification” [22] and the gentrification of low-carbon infrastructure more broadly [65].

The authors [64] argue that moving data is always more efficient than moving electricity and, therefore, data centers should be willing to pay rooftop solar owners for executing a computing task. This is all the more reason to expand research into carbon responses at the edge.

2.4. State of the Art of Enabling Data to Support CRC

Section 3 offers a rationale for why enabling data is likely to play a bigger role in the future of carbon responsive computing. Some elements of those datasets, and systems that enable use of those datasets, have emerged already for reasons unrelated to CRC. Here we identify two key areas of enabling data, and in the following section we elaborate on the changing role they play.

2.4.1. Available Carbon Intensity Data

All carbon responses, whether in the data center or elsewhere, rely on data about carbon intensity to make decisions about where and when to perform tasks (e.g., execute workloads, route data across a path in the network, postpone the movement of stored data). It stands to reason that carbon intensity data are likely to become an important shared resource for heterogeneous actors and therefore need to be properly maintained. It is important not to mistake energy prices for a sufficient proxy for carbon intensity. Prices are also affected by many factors unrelated to carbon intensity, making them an unreliable source of information for conducting carbon-responsive actions. Liu et al. [29] show that the ability to reduce fossil fuel use through geographic load balancing strongly depends on whether the data center's price directly reflects the instantaneous fraction of energy that is from fossil fuels. The extent to which incentive systems will, in practice, reflect carbon intensity specifically is an area of uncertainty.

On the national grid side, there is some amount of public data provision in many jurisdictions, but not all. Real time data, as opposed to downloadable historical data, can be even less readily available, as can fine-grain predictions and machine-readable data via application programming interfaces (APIs). An example of a carbon intensity data service specifically designed for enabling appliance demand response is the UK National Grid ESO's Carbon Intensity API [66], which provides half-hourly carbon intensity prediction and actuals for 14 UK regions in both human-readable and machine-readable formats.

To take advantage of differences in weather across the globe, a multi-jurisdictional carbon intensity data system is necessary. An example of an effort to integrate available data internationally, available in both human- and machine-readable formats, is Electricitymap.org [67]. This map is a useful indicator of the state of the art of publicly accessible carbon intensity data. It makes visible the coverage gaps across jurisdictions, which at the time of writing appears to be centered in Central America, Africa, and South, Central, and East Asia. South America, Europe, and Australasia appear to have largely consistent coverage. It also shows inconsistencies in maintaining live links between systems, which appear to be the greatest issue in North America. It also reveals a pattern of reportage by country for small and mid-sized countries, and by region for larger countries. Finally, it also maps energy import/export relations, which gives a better indication of the carbon intensity of consuming electricity in that region than raw energy generation figures.

Temporal granularity requirements for CRC have some associated research. Carbon-responsiveness decisions made in data centers on the hour consume significantly more non-renewable energy than those made on ten minute intervals [29]. Dou et al. [28] note that wind data shows an obvious variation at ten minute intervals, suggesting that ten minutes is also likely the smallest window of opportunity to respond. This suggests that ten minute windows are a potentially useful unit to report and to base predictions around. Many, but not all, grid providers appear to provide roughly compatible temporal granularity, though there is significant variation. Even within the US, "live" readings range from five minute to hourly intervals. In terms of forecasting, many of the systems cited above use day-ahead predictions, including Google's.

While there are significant gaps that create challenges for adapting currently available data to meet the needs of carbon responsive systems in light of edge-ification (see Section 2.3), these two examples demonstrate there is a good basis upon which to build.

2.4.2. Energy Consumption Measurement

When practices of time- and space-shifting workloads begin to affect energy providers, whether privately owned microgrids or public utilities, it will become vital to be able to predict, measure, and where necessary, disclose the energy consumption of shifted computing tasks. Peer-to-peer energy-compute exchanges across microgrids (i.e., you run my workload when you have energy and I run yours when I have energy) will need a way of predicting how much energy will be consumed by the exchanged task so the microgrid can decide whether or not to accommodate it. They will also need a way to confirm energy used. At an individual level, people cannot change their computing behavior without a sense of how much a difference it makes to delay or forego computing tasks.

There is a good deal of research on energy consumption of individual devices (see [47] for edge devices), and there is growing interest in measuring the energy consumption of AI models both as a matter of academic research [68,69] and in benchmarking [70]. However, software tasks executed across a system depend on different hardware and software configurations that change over time. In a real-world setting, a model is but one step in a larger process to complete a task; how it is integrated into that task has an effect on energy consumption. These complex dependencies mean that even measuring the difference between conducting inferencing with an AI model on a client such as a PC or in the cloud is remarkably complicated in a real-world setting. Energy measurement is not a routine practice of software developers, which makes matters even more difficult. The emergence of new categories of computing such as edge and the even more granular fog computing, which are intrinsically more distributed, make the gaps in knowledge even greater. For these reasons, there is more work to do to translate current practices into the capacity to predict how much of an energy “envelope” a given task will require.

However, there does appear to be more interest within the ICT industry in greater transparency about energy consumption. For example, calls have emerged for assessments of energy consumption as a regular part of machine learning model development and disclosing that information via model cards (a standard documentation format) [71]. New open source tools such as Power API [72] and SmartWatts [73] are also emerging. These tools allow developers a programmatic means to flexibly measure power consumption of algorithms or workloads at runtime across varied machines and contexts. As concerns over AI’s carbon footprint continue to grow, and carbon response techniques mature, there is more incentive to measure and disclose in appropriate ways, and more incentive to develop end-to-end toolsets that will allow real-world workloads to be measured and predicted.

3. Framework for Carbon-Responsive Computing in Diverse Contexts

In the previous section we described the increasing interdependence between energy and computing, the broadening of CRC into new types of ICT systems, and the nascent foundational groundwork laid for an enabling data infrastructure. These trends show no sign of abating. From a sociotechnical perspective, then, what do these seemingly disparate foundational elements of CRC amount to? How might we consider questions of social dynamics and social power as CRC matures? In this paper, we use social power to refer to a person or group’s ability to influence and shape the development of sociotechnical systems which, in turn, influences how power is distributed and experienced by different people in society. Because social power shapes technology and vice versa, we refer to technology systems as sociotechnical systems. Social power is always an element in technology development and use, whether it is explicitly recognized or not.

Thomas Hughes’ [17] classic history of how electricity became ubiquitous provides a useful sociotechnical model of how infrastructures form and evolve. Hughes describes five phases of infrastructural development: an invention and development phase, a tech-

nology transfer phase, a system growth phase, a system momentum phase, and a mature phase. Following this five-phase approach, it appears that, when we bring together the foundational elements of CRC, we have the makings of the early stages of an infrastructure. The relative maturity of carbon-response techniques in data centers, and the movement of the concept into nascent prototypes and experiments in edge systems, suggests that we have at least partially moved away from the initial invention and development phase, and into the technology transfer phase. In the technology transfer phase, the core concept of an infrastructure is adapted and elaborated, and local “styles” of how to instantiate it emerge. In CRC, the transfer is not from one type of society to another (as in Hughes’ original formulation), but a transfer between different computational settings. In Hughes’ account, the connections between systems and subsystems become more important as disparate point solutions become an integrated infrastructure. In the technology transfer phase, points of interoperability and shared, facilitating components become conceivable, though are not necessarily well-defined and not yet locked in. In the later system growth phase, “reverse salients” [17] (p. 14) become a factor, which is Hughes’ term for social or technical elements that hold back larger elements, potentially inhibiting an infrastructure from developing at all. These are put in place during the invention and technology transfer phases but become problematic during the system growth phase and beyond. For the time being, the final system momentum and mature phases appear to remain distant for CRC.

Thinking within this five-phased approach, we notice that, in the long term, the particularities of a scheduling algorithms in data centers or at the edge matters less than sociotechnical interdependencies. That is, if CRC is to become an infrastructure of some kind, it needs its own underlying connective tissue. What underlying capabilities do data centers, networks, edge systems, and energy grids large and small all have to rely on? When more ICT systems gain similar CRC capacities, what shared fate will befall them as interaction effects between different elements take hold at a more systemic level? When the priority is the development of working prototypes of individual systems, the infrastructural issue of shared fate—and its sociotechnical nature—has been largely neglected.

To address this missing facet, we offer a sociotechnically oriented framework to guide CRC development and practices (Figure 2). This three-staged framework clarifies in simple terms the cross-domain capabilities necessary to transition from domain-specific point solutions to a fully interdependent, interoperable infrastructure. These capabilities are described as three “jobs” to be done (column 2), deliberately leaving open to further debate how to do those jobs. The staged nature of the framework articulates the interdependencies between real-time carbon responses and the slower evolution of enabling infrastructures. This enables designers and engineers to map technical considerations onto social considerations, in order to anticipate the social and technical consequences of their designs for other parts of the interdependent system, and to mitigate potential problems before they arise. These stages are conceptually sequential, in that it is impossible to contemplate actions in the second stage without sufficient elements from the first stage. In practice, the cycle is iterative as experiments succeed or fail, and feedback loops support the maturation of systems.

3.1. Carbon-Aware Computing

As shown in Figure 2, carbon-aware is the foundational property that describes when ICT systems understand the carbon intensity of the energy consumed by a given device, element, or task. This stage describes a system that is able to characterize and predict the carbon intensity on the energy supply side, and measure energy consumption on the demand side. It also requires engagement with anthropological and sociological questions about the values and systems of social power that inform definitions of carbon intensity and shiftable tasks. Short-term response and long-term resilience planning (the next two stages) rely on knowing how many of which tasks can be delayed or redistributed to fit the “excess” energy available. Reliable carbon awareness becomes even more important in energy–compute exchanges across organizations. It is one thing to poorly estimate energy

consumption and accidentally pay more energy costs within one's own system, but it is quite another to pay someone else to take on your workload, only for them to discover it consumes much more energy than anticipated, or to incentivize someone to consume your excess electricity, only to discover the electricity generated is not as abundant as anticipated. For this reason, telemetry is needed to provide ongoing measurement and monitoring.

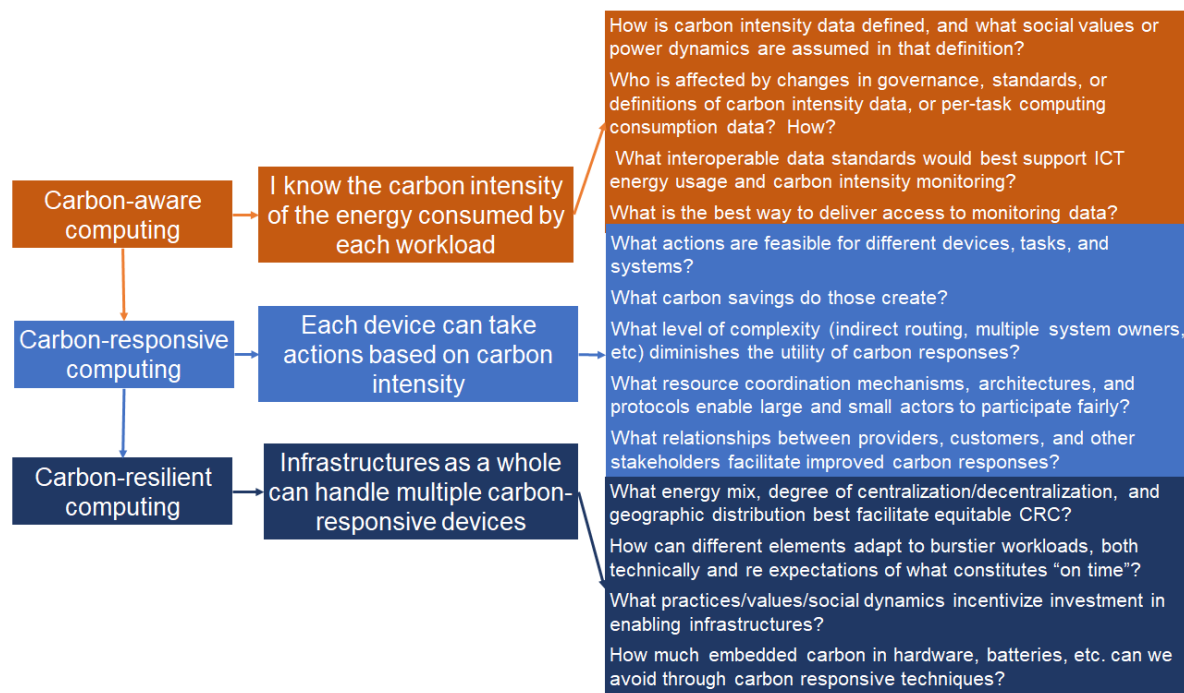


Figure 2. Three stages of carbon-responsive computing. Each phase, represented as a different shade of color, highlights key questions that need to be addressed when working at that phase. By examining the other phases, one can better anticipate the consequence for other stages.

3.2. Carbon-Responsive Computing

Carbon-responsive describes computing systems and their actions that are enabled by carbon awareness. Here, the research task for devices or systems is to identify and test which actions are feasible and desirable, based on data derived from the carbon-aware phase and additional social scientific and design research. This stage includes examining incentive structures (pricing, contractual arrangements, market position, etc.), and further engagement with sociotechnical research questions, particularly regarding the relationships necessary to facilitate improved carbon responses. Depending on context, actions might include time- and space-shifting, opportunistic sleep or other low-power states, or foregoing some or the entire task, such as postponing transmission. Another approach, used regularly in sleep and performance management algorithms, is to expand the time to execute a task, by slowing or speeding up a task in order to fill the time, given the energy available. It also includes any data systems necessary for matching workload providers and energy suppliers. Some researchers and early adopters in this field (e.g., Google) include both awareness and actions in their definition of "carbon-aware computing." However, keeping awareness and actions separate emphasizes that carbon-awareness is insufficient on its own—a system can be aware of its emissions and continue with business-as-usual nevertheless. In our framework, a carbon-aware system can serve many actors independent of the type of responses they pursue.

3.3. Carbon-Resilient Computing

Carbon-resilient systems explore the types of design, evaluation, and infrastructural changes on both the ICT and the energy side that will become necessary to manage carbon responsiveness at a more systemic level. These changes go beyond building a telemetry and data management infrastructure for carbon awareness and focus on what can be learned about a system as a whole from the carbon responses that emerge, and the monitoring available through ongoing carbon awareness. For example, a grid or microgrid might plan for different types of energy generation or levels of battery storage if it is clear that ICT and other energy demands can use otherwise stranded energy. Similarly, networking engineers might see that data transmissions become burstier upon renewable energy availability, and design new methods for prioritizing, scheduling, and routing time-elastic workloads toward edge data center locations with (excess) available renewables. A carbon-resilient computing system must also be able to use carbon-awareness data and response techniques to behave differently in emergency situations such as brownouts, denial-of-service attacks, or natural or climate-related disasters.

3.4. Using the Framework

Because the stages are interdependent, a designer, developer, or engineer can use the guiding, high-level questions outlined in Figure 2 to reflect on possible consequences in other areas of a CRC-enabling infrastructure. For example, if a networking engineer wanted to create software for routing packets in the least carbon-intensive way, she is working in the “carbon responsive” stage. However, if she examines the questions in the carbon awareness phase, she might not just look for the most convenient data, but realize that carbon intensity data that she is planning to use might have been defined in a way that does not match her values concerning different energy sources. She might advocate for changing how those data are collected, or choose another source so as not to give further legitimacy to those data. Conversely, the questions in the carbon-resilient stage might encourage her to probe on whether societal expectations of “on time” can change if there is a significant carbon gain, or if those expectations are indeed immutable and the better strategy is to encourage more on-site renewables at key nodes of the network.

This framework also allows researchers and developers to begin to anticipate where the reverse salients will be, by evaluating where the current state of the art does not support the goals listed in the second column. For example, in the carbon-aware stage, any data gap that inhibits the ability for a system to both understand carbon intensity, and energy consumption on a per-task basis prevents responses from happening at all. In the carbon-response stage, energy–compute exchange mechanisms, or a lack of methods to prevent gaming of the system, could inhibit participation from all relevant actors, and therefore system growth. In the carbon-resilient stage, reverse salients might come from an inappropriate energy mix whose bursts of energy supply cannot be made compatible with the temporalities of computing, or networking bottlenecks created by bursts of collective computing activity timed with the sun or wind.

The questions in the third column are not exhaustive, but point in an overall direction to facilitate further discussion outside any researchers’ or developers’ “natural” domain. Without critically reflective discussions, there is a risk of repeating what Hughes [17] found for the early days of electricity, namely, that those with the strongest social power can succeed in locking in their interests, not through a grand overarching architecture, but over the course of many seemingly small design decisions. As lock-in usually happens in the system momentum phase (see [74,75]), sociotechnical research and socially aware CRC practices will be essential for ensuring the social and technical elements enabling CRC promote equity, inclusion, and the flexibility to adapt to transforming social and environmental dynamics—something that will be key in times of climate crisis. In the next section, we will describe some of the practical challenges to conducting such essential sociotechnical research.

4. Practical Challenges and Areas for Sociotechnical Research

Sections 1–3 demonstrated the relevance of carbon responsive computing to an ever broadening suite of actors, from data centers managers, to edge developers, to microgrid owners, smart grid utilities, and energy–compute exchange enablers. Such heterogeneity requires change in research approaches. Not only is it necessary to build a body of engineering work that adapts these techniques to the edge and to the networks that connect them and builds the relevant energy–compute exchange mechanisms and underlying facilitating datasets, but it is also necessary to address the societal considerations that arise as this nascent infrastructure coalesces into an interconnected, cross-domain system. In this final section, we put the framework into action by examining two potential reverse salients in the carbon-aware stage (carbon intensity data and task energy consumption measurement) and one in the carbon-response stage (workload-for-energy exchanges) We begin by defining what it means to robustly address societal questions in this area.

4.1. The Democracy Gap

In the literature we examined in Section 2, proposals for new technologies or new exchange mechanisms were presented as if they were straightforward technical matters. STS scholars have raised concerns that this leaves insufficient room for democratic deliberation. For example, Ransan-Cooper et al. [76] have expressed this concern about algorithms that manage neighborhood-scale batteries, noting that this results in a bias towards designing only for the easily quantifiable, and problems of explainability and local autonomy. Without intervention, there is a high risk of replicating these problems in carbon-responsiveness, and creating what Miller et al. describe as “*stunted energy debates that provide limited opportunities for people other than energy engineers, bureaucrats, and economists to make influential contributions to energy policy deliberations*” [77] (p. 135). This is a longstanding issue, particularly in the case of technologies directly related to centralized geopolitical power (see [78,79] for the case of oil, see [80,81], for the case of nuclear, and [77] for US energy policymaking more broadly). However, efforts to promote democratic deliberation about energy technologies are increasing. For example, there are public deliberations on national nuclear energy policies in South Korea [82,83] and Finland [84]; participatory and deliberative public engagement in renewable energy systems development in Canada and Denmark [85]; participatory approaches to radioactive waste management in Belgium, Slovenia, Sweden, and the United Kingdom [86]; among others.

Deliberative and participatory processes are not linear or straightforward: e.g., they require navigating the uneven politics of participation [85] and developing innovative methods to elicit people’s values and interests [87], or to contend with dynamic social and material relations [88]. Nonetheless, they remain important to the success of energy design projects and therefore involve more actors (see [89,90] for examples of community engagement and opposition in wind energy and [88] for the role of social sciences and humanities scholars in the development of marine renewable energy). Ransan-Cooper et al. [76] provide a working example in the context of neighborhood batteries, where they showed that “upstream” engagement with involved communities leads to substantially different algorithmic control system design. They also show that there is no “optimal” algorithm that can successfully balance all needs; instead, public processes are necessary to come to consensus about priorities.

The need for social engagement and deliberation in energy technology and systems development is not only because these practices align with democratic principles, but because the knowledge shared by the public in these deliberations can greatly enrich the knowledge of experts and, thus the outcomes of technology development [80]. As noted by [80], energy technologies emerging out of democratic deliberations would necessarily be different from technologies developed through mere nominally public debate (Makhijani and Saleska [81] provide examples). Such insights fit into wider discussions on energy politics, which specifically attend to the social relations that result from adopting different energy sources [91]. These authors advocate for the need to acknowledge the social and

political aspects of energy systems since “*modes of fuelling and modes of governing society can no longer be easily separated*” [91] (p. 95).

Adding a social perspective, then, does not mean merely adding a collaborating sociologist. It also means inviting broader public deliberation to take place in order to shape the future direction of these technologies. In the “transfer stage,” (see Section 3), the enabling, data-intensive infrastructures are important sites where public deliberation and STS expertise can make a difference. We outline three enabling areas below to demonstrate how matters of public importance do in fact arise.

4.2. *Linked Carbon Intensity Data for Carbon Awareness*

In Section 2.4 we showed a good baseline for public availability of carbon intensity data. Because social power is always produced in data systems, whether consciously or not, (e.g., [92]) it is important to articulate whose perspective is reflected in those data. The current offerings we discussed in that section provide a grid-centered view of energy. Data from microgrids or other distributed sources appear less readily available in a similar systematic way. Yet to include multiple actors equitably, the grid-centered view needs extension, and this introduces problems of granularity, centralization vs. decentralization, standardization, and the social realities produced through data, which we discuss below.

4.2.1. Data Granularity

While there is emerging research on temporal granularity (see Section 2.4.1), spatial granularity has distinct social implications. In a world where energy generation is also being decentralized, and only some of that energy makes its way onto a national grid, a grid-centered view systematically underestimates edge-produced solar and wind energy. Some jurisdictions such as California require microgrids over a certain size to report measurements to public utilities, but not all jurisdictions do so. Associating carbon intensity with a geographic region to represent location, as in [67], belie the challenges and opportunities microgrids pose. For example, it might be unwise to send a workload to a region currently running at a high carbon intensity, but perfectly reasonable to send it to a microgrid where renewables are located within that region.

Improving spatial resolution beyond averages across a region could make a difference when it comes to choosing how to route data through a network. That might not mean reporting at a finer granularity than what currently exists (by county, for example) systematically, but it could mean reporting where there are pockets of renewables outside a national grid. This could support choices that take into consideration where renewables are located and in turn potentially alternate less carbon-intensive computing choices. However, any notion of spatial resolution needs to be sensitive to associated ownership of and access to the available energy as well as the associated data.

4.2.2. Data Centralization vs. Decentralization

Should energy-related data be compiled into a centralized or decentralized system, or a combined approach? A grid-centric view of the world would expand currently available datasets about grids to include information from microgrids and make them more interoperable across jurisdictions. A microgrid-centered view would do the opposite and propose a platform for microgrid owners to declare their excess energy production to one another with the intention of peer-to-peer exchange, where the electricity grid plays a more supportive backup role. A centralized approach provides generality, scale, and public interest that has utility. However, a centralized approach could force other smaller entities into unduly high barriers to participation. Large organizations tend to “see like a state” [93], meaning that they tend to require data from people, systems, and land that serves the interests of a large-scale bureaucracy, but makes less sense locally. For example, a large power utility has the resources to conduct detailed audits that appropriately validate claims about carbon intensity. However, that same costly scrutiny is unlikely to be necessary for a 100% renewable microgrid, where there is no risk of fossil fuels being present. Would

microgrid owners wish to contribute or receive data as a matter of inclusion, or would inclusion be experienced as coercive, or make them vulnerable to outsiders, as was the case in Scott's [93] classic study of centralization? Although one centralized system might marginalize smaller actors, a system specifically built for microgrid-to-microgrid exchanges or for edge-to-microgrid cooperation might miss cross-system opportunities, such as a data center sending workloads to households with rooftop solar [64], which requires information from both the microgrid and the central grid to understand if migrating the computation is less carbon-intensive.

4.2.3. Data Governance

Both distributed and centralized systems require governance, and as carbon intensity data becomes increasingly available as the basis for decision making, equitably designed governance practices begin to matter more. What kind of entities should provide governance structure(s), set the terms of access, conduct quality assurance, participate and fund development, and be responsible for maintenance? What resources are necessary for governance? For example, even just maintaining stable connections to APIs to retrieve carbon-intensity data can require multiple full-time developers. Who should be responsible for developing and maintaining the algorithms necessary to make carbon intensity predictions, and what happens when those predictions are inaccurate?

4.2.4. Standardization

What kind of standardization is desirable to support carbon responsive computing? Both software and people will need to be able to interpret carbon intensity data, though machines may be less forgiving about unexpected, incomplete, or incompatible data. For the purposes of interoperability, solution scalability, and ecosystem development, there is a strong need for data standards, to promote consistent representation and formats for data exchange. Participants will also desire standard methods to attest to data trustworthiness, particularly if other systems integrate and come to rely on these data, for example, to defer time-elastic workloads until such time as renewables are available, or to select one network path over another based on the relative carbon footprint. However, even if sustainability metrics for greener path selection [94] were standardized, there are sociotechnical issues to be investigated that will impact the process of wide-scale adoption and deployment in an infrastructure such as the Internet.

With the deployment of more edge data centers [14,15], and edge CRC, how can the industry support the federation of edges and/or their association with multiple cloud data centers, through a standard control protocol and/or a standard set of APIs? This federation would enable, for example, dynamic or semi-permanent regions where renewable energy availability and ICT energy consumption are naturally balanced. Will the emergence of edge service providers bring the larger cloud service providers to the table to open up the CRC algorithms and interfaces currently in use but proprietary?

At the same time, much social scientific scholarship [95–98] make clear that it is a fallacy to assume that, separate from consensus on data representation and formats, any data system can accurately arrive at the one true measurement of emissions. For example, [38] shows that, using carbon intensity to calculate the transmission of data on fixed line networks, there will exist a diversity of legitimate approaches that nonetheless arrive at a range of estimates.

We would expect to see multiple organizations and research teams continue to develop and evolve competing approaches to predicting carbon intensity or energy consumption needs, but industry is likely to prefer consistency in order to mitigate perceived risk. Will the institutionally-driven need for consistency hamper the adoption of new approaches and predictions? Is strict agreement on what counts as "low" carbon intensity necessary, or is it a sufficient goal simply to achieve lower carbon intensity than at present?

4.2.5. Data's Power to Create Social Realities

Those who design data systems have the power to create their own social realities [92,98]. When designing more carbon-responsive systems, whose version of what constitutes “carbon intensity” will become dominant? For example, on 17 June 2021, at 4:35 p.m. Carbonintensity.org.uk [66] reported the UK grid to be at 188 gCO₂/kWh, while Electricitymap.org [67] reported 285 carbon dioxide equivalent per kWh (gCO₂e/kWh), which includes non-carbon greenhouse gasses. The latter also incorporates a lifecycle analysis (LCA) of embedded emissions in energy equipment. Whose reality will become what computing systems take to be carbon intensity? Matters are likely to get even more heated when it comes to handling controversial energy generation technologies, such as nuclear. While nuclear energy itself emits no carbon dioxide, building, operating, mining for rare minerals, decommissioning, and waste storage certainly does, and estimates range from 1.4 g of CO₂e/kWh to 288 g CO₂e/kWh, with a mean of 66 g CO₂e/kWh [99]. What carbon intensity a system reports for nuclear (or any energy source) says much about systems' developers' beliefs about that energy source, and about the power that comes from defining the scope of data.

Datasets also create their own social realities through exclusions. For example, a literalistic approach might define any impacts on the environment beyond greenhouse gasses to be out of scope. Any negative impacts of hydroelectric power, nuclear, solar, wind, or battery production would be defined as irrelevant. This exclusion, however, replicates a core problem of sustainability: the externalization of problems deemed inconvenient. To the extent that most people equate “low carbon” with “good,” this exclusion greenwashes energy systems. Similarly, scholars have documented human rights abuses committed in the production of renewable energy technologies [100], and in the installation of renewable energy equipment (see [101]). It would be empirically valid to define these matters as out of scope, but including carbon intensity data from human rights violators, as if there were no issues in using that energy, potentially violating the UN Guiding Principles on Business and Human Rights. We are not arguing that carbon intensity data need also to be a repository of all social responsibility information. We are, however, suggesting that choices do need to be made about the entities to include, and where and how to acknowledge these problems in the datasets themselves. These cannot be treated as out-of-scope matters.

A third and final way datasets create their own social realities is through the ways they facilitate actors' abilities to pursue additional, unanticipated purposes, some of which could amount to gaming the system. Energy providers, governments, and even ICT actors tied to specific sources of energy have stakes in claiming that their systems are less carbon intensive than other systems. A likely secondary use of carbon intensity data will be in the compilation of emissions estimates by corporate social responsibility professionals, an arena where multiple carbon realities proliferate, depending on the goal and scoping of the measurement [97]. Could that use put pressure on the system to paint a certain type of picture that favors one actor or another? What can be put in place to mitigate that pressure?

Another use is likely to be in policymaking. If governments decide that interconnecting ICT and energy infrastructures is an important strategy, carbon intensity data could have significant impacts on planning decisions. For example, in Watts' [102] study of renewable energy generation on Orkney, locals were deeply frustrated by the UK National Grid's decision not to invest in an energy transmission cable to the mainland. One reason given was that there are few people on Orkney to serve, even if there is a great deal of wind. Does the calculus change, then, if the users are not people but data? Whose values are prioritized in that change? Would governments invest in coupled energy–compute infrastructures as a way of attracting private sector investment? Conversely, if much of the Global South remains heavily reliant on fossil fuels, carbon-responsive computing could re-entrench this longstanding path dependence by driving investment elsewhere. To prevent this, it is important to include research and development on carbon-response techniques by researchers situated in the Global South, where intermittent grid availability constitutes a research advantage.

4.3. Challenges in Facilitating Energy–Compute Exchange

4.3.1. Market Questions

We saw in Section 2 how various market mechanisms for enabling demand response and even exchanging workloads are beginning to emerge, but they did not address many of the questions raised in the carbon-response stage of the framework in Section 3. Future research that builds on Cioara et al. [13], taking into account a fuller range of ICT infrastructure, such as edge data centers and networks, can go a long way toward defining appropriate mechanisms for exchange given a broader heterogeneity of actors. Ethnographic and case-study research similar to [76] is also required to understand which systems or actors are currently leveraging machine-readable carbon intensity data, with a view to understanding emerging practices of workload-for-energy exchanges.

What has gone unsaid in the aforementioned designs and simulations (other than [76]), however, is the more fundamental question about the role of markets in the delivery of an essential resource. Addressing people as primarily or exclusively market actors might not motivate participation. Smart meter feedback systems are supposed to encourage market behavior by making energy savings visible, but they are not effective due to indifference to the relatively small monetary savings possible [103]. Conversely, a “peer-to-peer” energy trading system that does seem to work can be found in Watt’s ethnography of Orkney, UK [102], where non-market factors created strong incentives to pool excess energy and use it to alleviate on-island energy poverty—a factor not taken into consideration in the various trading algorithms we examined. Similarly, [53] report that those building local smart energy systems cite local resilience, justice, and equity as motivating values, underscoring the need to make these social aspirations a reality.

The Orkney example suggests that it is possible, and in fact might improve adoption, to organize exchanges as an embodied commons rather than an abstract commodity market. Giotitsas et al. [54] propose an energy commons predicated on the adoption of informatics systems, which might be adaptable for carbon responses. For example, one can imagine a system of delayed reciprocity (a routine principle of exchange observed by decades of economic anthropology since Mauss [104]) between two microgrids where if one entity is faced with the choice of turning to a backup diesel generator or shipping some workloads for all its IoT devices to its trading partner, it might choose to ship the workload, or pause it temporarily, if their trading partner is facing similar constraints. The receiving microgrid could also have its own optimizer that might scan the forecast of renewables supply and decide that even though there is not much energy flowing now, it is worth taking the risk of discharging some battery for their partners’ needs. An account could be kept concerning total energy used to ensure long-term imbalances do not arise, without the need to maximize profit in the short term. In a sense, this mimics what vertically integrated companies with multiple hyperscale data centers do: it makes no sense to internally charge a data center for services that the company also owns. Conversely, Nurminen et al. [64] work on distributing data center workloads to those with solar rooftops does require a price mechanism to ensure exchange. However, those with rooftop solar might not, in practice, have the opportunity to earn money if data centers lack a way to find willing solar owners, or if the necessary computing equipment is not in place. This barrier could easily lead to what geographers call “splintered urbanism” [105], where the infrastructures available to elites and non-elites are kept separate and unequal.

4.3.2. Complexity and Inclusion

Complex bidding and price clearing algorithms also raise barriers for democracy and ultimately social equity because they are inscrutable by non-experts who may be impacted by them. Özden-Schilling observes that contemporary energy markets have grown so complex that they “*transfer the burden to organize information onto market actors, making them into data processors and the markets themselves into infrastructures of information*” [106] (p. 68). Hyperscale data centers have the resources to manage the burdens of complex economic optimizations; a rooftop solar owner or rural microgrid cooperative might not. Both future

research and practical experimentation should explore ways of reducing the complexity of information signals concerning what workloads can be executed when, to ensure incentives can be realized by a heterogeneous set of actors, and that the signals are in fact directly aligned with carbon intensity. As noted in [76] and [68], this issue is also a problem of responsible AI more broadly.

4.4. Real-Time Energy Consumption Data

4.4.1. Improving Transparency

To build a better picture of ICT energy consumption on a task basis, the divide between in-house private sector research and public sector research needs to be crossed. Without access to telemetry, energy researchers are largely excluded from engineering conversations. As argued in [3], the ICT industry needs to begin making energy consumption telemetry a standard feature of telemetry offerings and make some of those measurements and inferences available to researchers and other members of the ecosystem for analysis. We recognize that measuring energy usage on a per computational task basis is significantly more complex than “adding telemetry”, and that there needs to be significant research efforts in-house to understand how appropriate measurement should appear. Companies will need to examine the granularity that they can reasonably make available for energy research such that it does not inadvertently disclose commercially sensitive matters. Nevertheless, this appears to be a worthwhile matter to examine—as single company, in-house research efforts are not likely to be sufficient for the task at hand. Technology developers cannot see the complex dynamics of energy delivery in ways that energy researchers can. Nor are developers going to have visibility into the energy consumption of the networks, servers, gateways, or clients they use without participation from the companies that own at least some of them, in order to derive measurements indirectly.

There are, of course, ways of implementing telemetry that could in effect significantly increase energy draw, and this also will need to be managed. Still, the body of telemetry data that contemporary software collects is already rich. The time is now to bring more creativity to the problem of collaborative approaches to measuring energy consumption. If public-facing bounties can be offered for finding security holes, why not offer them for finding pockets of unnecessary energy draw? Where are the grand challenges for finding opportunities for all of ICT to draw exclusively on renewable energy?

4.4.2. Defining Shiftable Tasks

Who gets to define what tasks are delayable, or even necessary? Smale et al. [107] argue, “*If these smart grid aspects are to be integrated into the household, this implies major changes in domestic everyday life*”. What is the extent of changes that are indeed required, versus the changes that can be re-designed to be pushed into the background, or run at lower CPU clock frequency, or altogether deferred until later? When delaying a task affects end users, qualitative research and broader public debate becomes vital. Following the practice-based approaches to energy consumption of [107,108], in the long term, how do individuals’ computing practices evolve with the temporalities of renewable energy availability?

To take a video streaming example, a company could choose to develop an algorithm to predict the next piece of content that will be watched, and pre-emptively download it onto the viewing device at times of renewables availability. A different company might indicate times of renewables availability on its platform and require the user to decide whether now is a good time to watch a film, or whether it is acceptable to watch in high definition given the current carbon intensity. These two strategies suggest very different understandings of what sustainability is about, who is responsible for it. One treats consumer “need” as immutable, and demand always “satisfied”, while the other assumes companies have a role in shaping what constitutes “necessary”. Conversely, the first company treats sustainability as its responsibility, while the other treats it as consumers’ responsibility. Indeed, Obringer et al. [37] argue that “over the top” service providers such as content delivery companies and remote meetings software providers are key stakeholders in reducing the Internet’s

carbon footprint through their design choices, yet users can also choose to simply turn off their video in meetings, which they suggest could save for an individual user 9 kg of CO₂e monthly. Because Internet video streaming also tends to coincide with times of overall peak energy consumption [108], an electricity company might also want to identify the best incentive to time-shift the network transmission portion of the content.

One promising research direction might be to instrument a computing task and explore how often people do it under different design conditions similar to the ones outlined above. Would people be willing to host video meetings if they could see directly that the bulk of the Internet traffic it created was using nonrenewable energy? Might they use it as a reason to keep the meeting brief, or reschedule, or turn off video? Qualitative interviewing would yield a perspective on why they took the actions they did and deliver a broader perspective on publics' views on who or what should shape "consumer need", while instrumentation could enable an evaluation of emissions savings of different carbon responses. Both of these will be vital to engineering research on algorithmic classifiers that distinguish tasks available for carbon responses from those that are not.

5. Conclusions

This paper demonstrated that CRC is on a trajectory of expansion, beyond data centers and into new areas of computing such as edge computing and networking, where many engineering unknowns remain. It also offered a three-part framework for considering this maturation beyond domain-specific components and articulates the sociotechnical interdependencies of that expansion. In that broadening, which coincides with the decentralization of energy production as well as the edge-ification of the Internet, the research agendas that take on urgency encounter problems of governance, transparency, inclusion, and what constitutes computing "need" in a climate crisis. These are not research agendas that can safely remain in the confines of social scientific journals unfamiliar to readers of *Energies*. They need to be acknowledged, understood, and acted upon by those who design and build carbon-responsive systems.

Discipline-specific expertise nevertheless informs interdisciplinary charters. To conclude, we list some examples of the kinds of questions researchers should begin to ask within more discipline-specific framing.

- For quantitative energy researchers: What could the impacts be of coupling edge computing with microgrids in terms of carbon savings? What do emerging qualitative and mixed methods findings about appropriate tasks to time and space shift tell us about the amount of aggregate energy demand that can be decarbonized? Can improved routes for data traffic be predicted based on carbon intensity data currently available, and what does this tell us about requirements for carbon intensity data for the future? What are the geographic and infrastructural factors that influence the ability to make savings?
- For designers and human–computer interaction (HCI) researchers: What constitutes computing "need" and "urgency" in what social situations, with what design strategies? What can be learned in environments that are already constrained by renewable energy availability? What are the best design strategies, of either data systems or ICT products, when matters of controversy arise (what constitutes carbon intensity, the acceptability or not of foregoing a computing task, etc.)? What lessons can be learned from parts of the world with intermittent electricity?
- For computer scientists and the computing industry: In what ways are the tradeoffs in doing carbon responses in edge systems and networks different or similar to those in hyperscale data centers? Under what conditions do the energy-saving strategies in large data centers require adaptation or abandonment in more proximate edge data centers and networks? What algorithms can best recognize and service the time-sensitivity of tasks, or recommend what type of response is best? How much of ICT equipment and its users' behavior can be predicted and effectively mapped to microgrid renewables surplus? What kind of energy consumption telemetry setup

most efficiently characterizes energy on a task basis? What is the most appropriate way to disclose energy consumption data in a standardized manner to the ecosystem and researchers, and to incentivize its use, whilst being mindful of privacy concerns?

- For STS researchers: If “*modes of fuelling and modes of governing society can no longer be easily separated*” [91] (p. 95), what kinds of energy politeness are forming as ICTs and energy increasingly converge? What actors gain, maintain, or lose social power in that coupling? What sociotechnical interventions are necessary to ensure carbon responses do not merely reproduce business as usual? What kinds of collaborations are necessary to ensure that ethical questions are indeed acknowledged, such as environmental social justice? What actors make for good energy-computing trading partners and poor ones?
- For economists/management studies scholars: Where markets do play a role, what types of markets could be designed, and who are the actors and stakeholders involved? What is the role of pricing, contract structures and business models in maximizing carbon emissions reductions? What algorithms can best match workloads with energy sources, and how might they scale to accommodate federation and peering of resources within different size regions? How will different regulatory conditions affect the geopolitical dynamics of investment in coupled energy-computing systems?
- For policy researchers: What disincentives are most effective for preventing actors to game or manipulate carbon intensity data, or energy usage data? What policies ensure the best provisioning of electricity as a public good in light of new energy-computing combinations, and across public-private relationships? What role might combined energy-computing systems play in national and international geopolitics, across varied cultural attitudes towards data privacy, and varied histories or electricity provisioning?

These questions are intentionally broad and incomplete. However, a systems-level problem requires a systems-level view. For too long, computing systems have been designed assuming an infinite supply of natural resources without consequence. Redesigning computing systems to fit the temporalities and spatialities of the energy available, at scale commensurate with the global scale of the climate problem, and commensurate with its urgency, is one way we can profoundly change the relationship between technology and natural systems. Doing that is an enormous multidisciplinary challenge for both research and practice, but likely to be a rewarding one.

Author Contributions: Conceptualization: D.N. and E.M.S.; investigation: D.N., E.M.S., and K.A.B.; writing—original draft preparation, D.N., E.M.S., and K.A.B.; writing—review and editing, D.N., E.M.S., and K.A.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors would like to thank colleagues at Intel: Jesse Bearce, Beppe Raffa, Mel Gregg, John Miranda, John Sherry, Michelle Chauprasert, Ryan Gorostiza and Sylvia Downing for feedback on early conceptual work that informed this paper. Ben Anderson greatly informed our thinking throughout. At Otago, David Eyers, Michael Jack, and the participants of the “ICTs and Sustainability” Workshop at the Centre for Sustainability gave helpful early feedback. Indira Venkatraman supported background bibliographical work.

Conflicts of Interest: Intel gave input on the topic selection and reviewed the manuscript to ensure trade secrets and intellectual property were not disclosed. It played no role in the analysis or writing.

References

1. Jones, N. How to stop data centres from gobbling up the world’s electricity. *Nature* **2018**, *561*, 163–166. [[CrossRef](#)]
2. Lange, S.; Pohl, J.; Santarius, T. Digitalization and energy consumption. Does ICT reduce energy demand? *Ecol. Econ.* **2020**, *176*, 106760. [[CrossRef](#)]
3. Koomey, J.; Masanet, E. Does not compute: Avoiding pitfalls assessing the Internet’s energy and carbon impacts. *Joule* **2021**, *5*, 1625–1628. [[CrossRef](#)]

4. Freitag, C.; Berners-Lee, M.; Widdicks, K.; Knowles, B.; Blair, G.; Friday, A. The Climate Impact of ICT: A Review of Estimates, Trends and Regulations. 2020. Available online: <https://eprints.lancs.ac.uk/id/eprint/158061/> (accessed on 31 August 2021).
5. Belkhir, L.; Elmeligi, A. Assessing ICT global emissions footprint: Trends to 2040 & recommendations. *J. Clean. Prod.* **2018**, *177*, 448–463. [[CrossRef](#)]
6. Masanet, E.; Shehabi, A.; Lei, N.; Smith, S.; Koomey, J. Recalibrating global data center energy-use estimates. *Science* **2020**, *367*, 984–986. [[CrossRef](#)] [[PubMed](#)]
7. Manganelli, M.; Soldati, A.; Martirano, L.; Ramakrishna, S. Strategies for improving the sustainability of data centers via energy mix, energy conservation, and circular energy. *Sustainability* **2021**, *13*, 6114. [[CrossRef](#)]
8. Usama, M.; Erol-Kantarci, M. A survey on recent trends and open issues in energy efficiency of 5G. *Sensors* **2019**, *19*, 3126. [[CrossRef](#)]
9. Chiaraviglio, L.; Mellia, M.; Neri, F. Energy-Aware Backbone Networks: A Case Study. In Proceedings of the 2009 IEEE International Conference on Communications Workshops, Dresden, Germany, 14–18 June 2009; pp. 1–5. [[CrossRef](#)]
10. Akoush, S.; Sohan, R.; Rice, A.; Moore, A.W.; Hopper, A. Free lunch: Exploiting renewable energy for computing. In *HotOS*; University of Cambridge: Cambridge, UK, 2011; pp. 1–5.
11. Radovanovic, A. Our Data Centers Now Work Harder When the Sun Shines and Wind Blows. Google. 2020. Available online: <https://blog.google/inside-google/infrastructure/data-centers-work-harder-sun-shines-wind-blows/> (accessed on 26 August 2021).
12. Lin, L.; Zavala, V.M.; Chien, A.A. Evaluating Coupling Models for Cloud Datacenters and Power Grids. In Proceedings of the Tenth ACM International Conference on Future Energy Systems, Phoenix, AZ, USA, 28 June 2021; pp. 171–184. [[CrossRef](#)]
13. Cioara, T.; Anghel, I.; Salomie, I.; Antal, M.; Pop, C.; Bertocini, M.; Arnone, D.; Pop, F. Exploiting data centres energy flexibility in smart cities: Business scenarios. *Inf. Sci.* **2018**, *476*, 392–412. [[CrossRef](#)]
14. Van der Meulen, R. What Edge Computing Means for Infrastructure and Operations Leaders. 2018. Available online: <https://www.gartner.com/smarterwithgartner/what-edge-computing-means-for-infrastructure-and-operations-leaders/> (accessed on 14 July 2021).
15. Verified Market Research (VMR) Edge Data Center Market Size, Share, Trends, Opportunities & Forecast. Available online: <https://www.verifiedmarketresearch.com/product/edge-data-center-market/> (accessed on 26 July 2021).
16. Blaabjerg, F.; Yang, Y.; Wang, X. Distributed Power-Generation Systems and Protection. *Proc. IEEE* **2017**, *105*, 1311–1331. [[CrossRef](#)]
17. Hughes, T.P. *Networks of Power: Electrification in Western Society, 1880–1930*; Johns Hopkins University Press: Baltimore, MD, USA, 1983; ISBN 0801828732.
18. Purdie, J. Climate Change Impacts on the New Zealand Energy System. In *Proceedings of the Seminar, Centre for Sustainability*; University of Otago: Dunedin, New Zealand, 2021.
19. Wierman, A.; Liu, Z.; Liu, I.; Mohsenian-Rad, H. Opportunities and Challenges for Data Center Demand Response. In Proceedings of the International Green Computing Conference, Dallas, TX, USA, 3–5 November 2015. [[CrossRef](#)]
20. Ghasemi-Gol, M.; Wang, Y.; Pedram, M. An Optimization Framework for Data Centers to Minimize Electric Bill Under Day-Ahead Dynamic Energy Prices While Providing Regulation Services. In Proceedings of the International Green Computing Conference, Dallas, TX, USA, 3–5 November 2014. [[CrossRef](#)]
21. Hogan, M. Data flows and water woes: The Utah Data Center. *Big Data Soc.* **2015**, *2*, 1–12. [[CrossRef](#)]
22. Libertson, F.; Velkova, J.; Palm, J. Data-center infrastructure and energy gentrification: Perspectives from Sweden. *Sustain. Sci. Pract. Policy* **2021**, *17*, 153–162. [[CrossRef](#)]
23. Brodie, P. Climate extraction and supply chains of data. *Media Cult. Soc.* **2020**, *42*, 1095–1114. [[CrossRef](#)]
24. Goiri, Í.; Beauchea, R.; Le, K.; Nguyen, T.D.; Haque, M.E.; Guitart, J.; Torres, J.; Bianchini, R. Greenslot: Scheduling Energy Consumption in Green Datacenters. In Proceedings of the 2011 International Conference for High Performance Computing, Networking, Storage and Analysis, Seattle, WA, USA, 12–18 November 2011; p. 1. [[CrossRef](#)]
25. Shuja, J.; Gani, A.; Shamshirband, S.; Ahmad, R.W.; Bilal, K. Sustainable Cloud Data Centers: A survey of enabling techniques and technologies. *Renew. Sustain. Energy Rev.* **2016**, *62*, 195–214. [[CrossRef](#)]
26. Li, D.; Qu, X.; Wan, J.; Wang, J.; Xia, Y.; Zhuang, X.; Xie, C. Workload Scheduling for Massive Storage Systems with Arbitrary Renewable Supply. *IEEE Trans. Parallel Distrib. Syst.* **2018**, *29*, 2373–2387. [[CrossRef](#)]
27. Goiri, Í.; Katsak, W.; Le, K.; Nguyen, T.D.; Bianchini, R. Designing and Managing Data Centers Powered by Renewable Energy. *IEEE Micro* **2014**, *34*, 8–16. [[CrossRef](#)]
28. Dou, H.; Qi, Y.; Wei, W.; Song, H. Minimizing Electricity Bills for Geographically Distributed Data Centers with Renewable and Cooling Aware Load Balancing. In Proceedings of the 2015 International Conference on Identification, Information, and Knowledge in the Internet of Things (IIKI), Beijing, China, 22–23 October 2016; pp. 210–214. [[CrossRef](#)]
29. Liu, Z.; Lin, M.; Wierman, A.; Low, S.; Andrew, L.L.H. Greening geographical load balancing. *ACM Sigmetrics Perform. Eval. Rev.* **2011**, *39*, 193–204. [[CrossRef](#)]
30. St. John, J. Google Tests Load-Shifting at Data Centers to Capture the Grid’s Peak Clean Energy Hours. Available online: <https://www.greentechmedia.com/articles/read/google-tests-shifting-data-center-loads-to-capture-grids-clean-energy> (accessed on 19 March 2021).

31. Klingert, S. DC4Cities—D2.4: Final Market Analysis. Available online: <http://www.dc4cities.eu/en/wp-content/uploads/%202016/05/D2.4-Final-Market-Analysis.pdf> (accessed on 18 March 2021).
32. Cook, G.; Lee, J.; Tsai, T.; Kong, A.; Deans, J.; Johnson, B.; Jardim, E. Clicking Clean: Who Is Winning the Race to Build a Green Internet? 2017. Available online: <http://www.clickclean.org/international/en/> (accessed on 17 September 2021).
33. Cook, G.; Jardim, E. Clicking Clean Virginia, the Dirty Energy Powering Data Center Alley. 2019. Available online: https://greenpeace.org/usa/wp-content/uploads/2019/02/Greenpeace-Click-Clean-Virginia-2019.pdf?_ga=2.229195804.1861093522.1628839278-976262951.1628839278 (accessed on 17 September 2021).
34. Sun, Q.; Ren, S.; Wu, C.; Li, Z. An Online Incentive Mechanism for Emergency Demand Response in Geo-Distributed Colocation Data Centers. In Proceedings of the 17th International Conference on Future Energy Systems, Waterloo, ON, Canada, 21–24 June 2016; Association for Computing Machinery: New York, NY, USA, 2016.
35. Bahrami, S.; Wong, V.W.S.; Huang, J. Data Center Demand Response in Deregulated Electricity Markets. *IEEE Trans. Smart Grid* **2019**, *10*, 2820–2832. [[CrossRef](#)]
36. Lorincz, J.; Capone, A.; Wu, J. Greener, energy-efficient and sustainable networks: State-of-the-art and new trends. *Sensors* **2019**, *19*, 4864. [[CrossRef](#)]
37. Obringer, R.; Rachunok, B.; Maia-Silva, D.; Arbabzadeh, M.; Nateghi, R.; Madani, K. The overlooked environmental footprint of increasing Internet use. *Resour. Conserv. Recycl.* **2021**, *167*, 105389. [[CrossRef](#)]
38. Aslan, J.; Mayers, K.; Koomey, J.G.; France, C. Electricity intensity of internet data transmission untangling the estimates. *J. Ind. Ecol.* **2018**, *22*, 785–798. [[CrossRef](#)]
39. Jalali, F.; Hinton, K.; Ayre, R.; Alpcan, T.; Tucker, R.S. Fog computing may help to save energy in cloud computing. *IEEE J. Sel. Areas Commun.* **2016**, *34*, 1728–1739. [[CrossRef](#)]
40. Gigis, P.; Calder, M.; Manassakis, L.; Nomikos, G.; Kotronis, V.; Dimitropoulos, X.; Katz-Bassett, E.; Smaragdakis, G. Seven Years in the Life of Hypergiants’ off-Nets. In Proceedings of the 2021 ACM SIGCOMM 2021 Conference, Virtual Event USA, 23–27 August 2021; Association for Computing Machinery: New York, NY, USA, 2021; pp. 516–533.
41. Wang, S.; Zhang, X.; Zhang, Y.; Wang, L.; Yang, J.; Wang, W. A Survey on Mobile Edge Networks: Convergence of Computing, Caching and Communications. *IEEE Access* **2017**, *5*, 6757–6779. [[CrossRef](#)]
42. Shi, W.; Cao, J.; Zhang, Q.; Li, Y.; Xu, L. Edge Computing: Vision and Challenges. *IEEE Internet Things J.* **2016**, *3*, 637–646. [[CrossRef](#)]
43. Wang, S.; Tuor, T.; Salonidis, T.; Leung, K.K.; Makaya, C.; He, T.; Chan, K. Adaptive Federated Learning in Resource Constrained Edge Computing Systems. *IEEE J. Sel. Areas Commun.* **2019**, *37*, 1205–1221. [[CrossRef](#)]
44. Sapio, A.; Abdelaziz, I.; Aldilajjan, A.; Canini, M.; Kalnis, P. In-Network Computation Is A Dumb Idea Whose Time Has Come. In Proceedings of the 16th ACM Workshop on Hot Topics in Networks, Palo Alto, CA, USA, 30 November–1 December 2017; pp. 150–156. [[CrossRef](#)]
45. Yu, W.; Liang, F.; He, X.; Hatcher, W.G.; Lu, C.; Lin, J.; Yang, X. A Survey on the Edge Computing for the Internet of Things. *IEEE Access* **2017**, *6*, 6900–6919. [[CrossRef](#)]
46. Lim, W.Y.B.; Luong, N.C.; Hoang, D.T.; Jiao, Y.; Liang, Y.C.; Yang, Q.; Niyato, D.; Miao, C. Federated Learning in Mobile Edge Networks: A Comprehensive Survey. *IEEE Commun. Surv. Tutorials* **2020**, *22*, 2031–2063. [[CrossRef](#)]
47. Jiang, C.; Fan, T.; Gao, H.; Shi, W.; Liu, L.; Cérin, C.; Wan, J. Energy aware edge computing: A survey. *Comput. Commun.* **2020**, *151*, 556–580. [[CrossRef](#)]
48. Li, W.; Yang, T.; Delicato, F.C.; Pires, P.F.; Tari, Z.; Khan, S.U.; Zomaya, A.Y. On Enabling Sustainable Edge Computing with Renewable Energy Resources. *IEEE Commun. Mag.* **2018**, *56*, 94–101. [[CrossRef](#)]
49. Ahvar, E.; Orgerie, A.-C.C.; Lebre, A. Estimating Energy Consumption of Cloud, Fog and Edge Computing Infrastructures. *IEEE Trans. Sustain. Comput.* **2019**, *1*, 1–12. [[CrossRef](#)]
50. Mengelkamp, E.; Gärttner, J.; Rock, K.; Kessler, S.; Orsini, L.; Weinhardt, C. Designing microgrid energy markets: A case study: The Brooklyn Microgrid. *Appl. Energy* **2018**, *210*, 870–880. [[CrossRef](#)]
51. Vergados, D.J.; Mamounakis, I.; Makris, P.; Varvarigos, E. Prosumer clustering into virtual microgrids for cost reduction in renewable energy trading markets. *Sustain. Energy Grids Netw.* **2016**, *7*, 90–103. [[CrossRef](#)]
52. Paladin, A.; Das, R.; Wang, Y.; Ali, Z.; Kotter, R.; Putrus, G.; Turri, R. Micro market based optimisation framework for decentralised management of distributed flexibility assets. *Renew. Energy* **2021**, *163*, 1595–1611. [[CrossRef](#)]
53. Ford, R.; Maidment, C.; Vigurs, C.; Fell, M.J.; Morris, M. Smart local energy systems (SLES): A framework for exploring transition, context, and impacts. *Technol. Forecast. Soc. Change* **2021**, *166*, 120612. [[CrossRef](#)]
54. Giotitsas, C.; Pazaitis, A.; Kostakis, V. A peer-to-peer approach to energy production. *Technol. Soc.* **2015**, *42*, 28–38. [[CrossRef](#)]
55. Schooler, E.M. The Edge-ification of the Internet: Implications for the Wireless Edge. In Proceedings of the 21st International Workshop on Mobile Computing Systems and Applications, Austin, TX, USA, 3 March 2020; p. 2. [[CrossRef](#)]
56. Lally, N.; Kay, K.; Thatcher, J. Computational parasites and hydropower: A political ecology of Bitcoin mining on the Columbia River. *Environ. Plan. E Nat. Sp.* **2019**. [[CrossRef](#)]
57. Mohan, N.; Kangasharju, J. Edge-Fog Cloud: A Distributed Cloud for Internet of Things Computations. In Proceedings of the 2016 Cloudification of the Internet of Things (CIoT), Paris, France, 23–26 November 2017; pp. 1–6. [[CrossRef](#)]
58. Shah-Mansouri, H.; Wong, V.W.S. Hierarchical fog-cloud computing for IoT systems: A computation offloading game. *IEEE Internet Things J.* **2018**, *5*, 3246–3257. [[CrossRef](#)]

59. Ku, Y.J.; Sapra, S.; Baidya, S.; Dey, S. State of Energy Prediction in Renewable Energy-Driven Mobile Edge Computing Using CNN-LSTM Networks. In Proceedings of the 2020 IEEE Green Energy and Smart Systems Conference (IGESSC), Long Beach, CA, USA, 2–3 November 2020. [CrossRef]
60. GSMA. Infrastructure Sharing: An Overview. 2019. Available online: <https://www.gsma.com/futurenetworks/wiki/infrastructure-sharing-an-overview> (accessed on 31 August 2021).
61. Carlini, S. Telco Central Offices are Being Transformed into the Edge to Power the Next Generation of Telco the Early Dilemma with Central Office Transformation—How to Repurpose Transforming the Telco Edge: NGCOs Telco Edge Technology: A Cheat Sheet. 2019. Available online: <https://blog.se.com/co-location/2019/02/14/telco-central-offices-transformed-into-edge-power-next-generation-telco/> (accessed on 31 August 2021).
62. Tweed, K. Why Cellular Towers in Developing Nations are Making the Move to Solar Power. 2013. Available online: <https://www.scientificamerican.com/article/cellular-towers-moving-to-solar-power/> (accessed on 31 August 2021).
63. Yang, F.; Chien, A.A. ZCCloud: Exploring Wasted Green Power for High-Performance Computing. In Proceedings of the 2016 IEEE International Parallel and Distributed Processing Symposium (IPDPS), Chicago, IL, USA, 23–27 May 2016; pp. 1051–1060. [CrossRef]
64. Nurminen, J.K.; Niemi, T.; Strandman, J.; Ruokosuo, K. Sunburn—Using excess energy of small-scale production for distributed computing. *Energy Effic.* **2018**, *11*, 97–119. [CrossRef]
65. Bouzarovski, S.; Frankowski, J.; Tirado Herrero, S. Low-Carbon Gentrification: When Climate Change Encounters Residential Displacement. *Int. J. Urban Reg. Res.* **2018**, *42*, 845–863. [CrossRef]
66. National Grid ESO Carbon Intensity API. Available online: <https://carbonintensity.org.uk/> (accessed on 27 August 2021).
67. Tomorrow Company ElectricityMap. Available online: www.electricitymap.org (accessed on 27 August 2021).
68. Bender, E.M.; Gebru, T.; McMillan-Major, A.; Shmitchell, S. *On the Dangers of Stochastic Parrots: Can Language Models be Too Big?* Association for Computing Machinery: New York, NY, USA, 2021; Volume 1, ISBN 9781450383097.
69. Strubell, E.; Ganesh, A.; & McCallum, A. Energy and policy considerations for deep learning in NLP. *arXiv* **2019**, arXiv:1906.02243. preprint. (accessed on 4 October 2021).
70. Ray, T. AI Industry’s Performance Benchmark, Mlperf, for the First Time Also Measures the Energy That Machine Learning Consumes. 2021. Available online: <https://www.zdnet.com/article/ai-industrys-performance-benchmark-mlperf-for-the-first-time-also-measures-the-energy-that-machine-learning-consumes/> (accessed on 27 August 2021).
71. Schwartz, R.; Dodge, J.; Smith, N.A.; Etzioni, O. Green AI. *Commun. ACM* **2020**, *63*, 54–63. [CrossRef]
72. Power API. Available online: <http://Powerapi.org> (accessed on 27 August 2021).
73. Fieni, G.; Rouvoy, R.; Seinturier, L. SmartWatts: Self-Calibrating Software-Defined Power Meter for Containers. In Proceedings of the 2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID), Melbourne, VIC, Australia, 11–14 May 2020; pp. 479–488. [CrossRef]
74. Nye, D. Electricity and Culture: Conceptualizing the American Case. *Ann. Hist. L’électricité* **2004**, *2*, 125. [CrossRef]
75. Hughes, T.P. Technological Momentum. In *Does Technology Drive History: The Dilemma of Technological Determinism*; Smith, M.R., Marx, L., Eds.; MIT Press: Cambridge, MA, USA, 1994.
76. Ransan-Cooper, H.; Sturmberg, B.C.P.; Shaw, M.E.; Blackhall, L. Applying responsible algorithm design to neighbourhood-scale batteries in Australia. *Nat. Energy* **2021**, *6*, 815–823. [CrossRef]
77. Miller, C.A.; Iles, A.; Jones, C.F. The Social Dimensions of Energy Transitions. *Sci. Cult.* **2013**, *22*, 135–148. [CrossRef]
78. Bromley, S. *American Hegemony and World Oil: The Industry, the State System and the World Economy*; Pennsylvania State University Press: University Park, PA, USA, 1991; ISBN 027100746X 9780271007465.
79. Barry, A. *Material Politics: Disputes along the Pipeline*; Wiley-Blackwell: New York, NY, USA, 2013.
80. Callon, M.; Lascoumes, P.; Barthe, Y. *Acting in an Uncertain World: An Essay on Technical Democracy*; MIT Press: Cambridge, MA, USA, 2009; ISBN 9780262033824.
81. Makhijani, A.; Saleska, S. *The Nuclear Power Deception: US Nuclear Mythology from Electricity “Too Cheap to Meter” to “Inherently Safe” Reactors*; The Apex Press: New York, NY, USA, 1999; ISBN 0945257759/9780945257752.
82. Chung, J.B. Public deliberation on the national nuclear energy policy in Korea—Small successes but bigger challenges. *Energy Policy* **2020**, *145*, 111724. [CrossRef]
83. Kim, P.; Kim, J.; Yim, M.S. How deliberation changes public opinions on nuclear energy: South Korea’s deliberation on closing nuclear reactors. *Appl. Energy* **2020**, *270*, 115094. [CrossRef]
84. Setälä, M.; Grönlund, K.; Herne, K. Citizen deliberation on nuclear power: A comparison of two decision-making methods. *Polit. Stud.* **2010**, *58*, 688–714. [CrossRef]
85. MacArthur, J.L. Challenging public engagement: Participation, deliberation and power in renewable energy policy. *J. Environ. Stud. Sci.* **2016**, *6*, 631–640. [CrossRef]
86. Bergmans, A.; Sundqvist, G.; Kos, D.; Simmons, P. The participatory turn in radioactive waste management: Deliberation and the social-technical divide. *J. Risk Res.* **2015**, *18*, 347–363. [CrossRef]
87. Pschetz, L.; Pothong, K.; Speed, C. Autonomous Distributed Energy Systems: Problematizing the Invisible Through Design, Drama And Deliberation. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK, 4–9 May 2019; pp. 1–14. [CrossRef]

88. Kerr, S.; Watts, L.; Brennan, R.; Howell, R.; Graziano, M.; O'Hagan, A.M.; van der Horst, D.; Weir, S.; Wright, G.; Wynne, B. Shaping blue growth: Social sciences at the nexus between marine renewables and energy policy. In *Advancing Energy Policy. Lessons on the Integration of Social Sciences and Humanities*; Ptashnyk, I., Robison, R., Eds.; Palgrave Macmillan: Cham, Switzerland, 2018; pp. 31–46.
89. Hall, N.L.; Hicks, J.; Lane, T.; Wood, E. Planning to engage the community on renewables: Insights from community engagement plans of the Australian wind industry. *Australas. J. Environ. Manag.* **2020**, *27*, 123–136. [[CrossRef](#)]
90. Martinez, N. Resisting renewables: The energy epistemics of social opposition in Mexico. *Energy Res. Soc. Sci.* **2020**, *70*, 101632. [[CrossRef](#)]
91. Winthereik, B.R.; Helmreich, S.; O'Doherty, D.; Amador-Jiménez, M.; Marres, N. Five theses on energy politics. In *Energy Worlds in Experiment*; Maguire, J., Watts, L., Winthereik, B.R., Eds.; Mattering Press: Manchester, UK, 2021; pp. 95–111.
92. Beer, D. The social power of algorithms. *Inf. Commun. Soc.* **2017**, *20*, 1–13. [[CrossRef](#)]
93. Scott, J.C. *Seeing Like a State: How Certain Schemes to Improve the Human Condition have Failed*; Yale University Press: New Haven, CT, USA, 1998; ISBN 0300070160.
94. Hossain, M.M.; Georges, J.P.; Rondeau, E.; Divoux, T. Energy, carbon and renewable energy: Candidate metrics for green-aware routing? *Sensors (Switzerland)* **2019**, *19*. [[CrossRef](#)]
95. D'Ignazio, C.; Klein, L.F. *Data Feminism*; MIT Press: Cambridge, MA, USA, 2020; ISBN 9780262358521.
96. Loukissas, Y.A.; Bowker, G.C. *All Data Are Local: Thinking Critically in a Data-Driven Society*; MIT Press: Cambridge, MA, USA, 2019; ISBN 9780262039666 0262039664.
97. Lippert, I. Environment as datascape: Enacting emission realities in corporate carbon accounting. *Geoforum* **2015**, *66*, 126–135. [[CrossRef](#)]
98. Power, M. *The Audit Society: Rituals of Verification*; Oxford University Press: Oxford, UK, 1997; ISBN 0198289472.
99. Sovacool, B.K. Valuing the greenhouse gas emissions from nuclear power: A critical survey. *Energy Policy* **2008**, *36*, 2950–2963. [[CrossRef](#)]
100. Riofrancos, T.N. *Resource Radicals: From Petro-Nationalism to Post-Extractivism in Ecuador*; Duke University Press: Durham, UK, 2020; ISBN 9781478012122.
101. Boyer, D. *Energopolitics: Wind and Power in the Anthropocene*; Duke University Press: Durham, UK, 2019; ISBN 1-4780-0439-8.
102. Watts, L. *Energy at the End of the World: An Orkney Islands Saga*; MIT Press: Cambridge, MA, USA, 2018; ISBN 9780262038898.
103. Buchanan, K.; Russo, R.; Anderson, B. The question of energy reduction: The problem(s) with feedback. *Energy Policy* **2015**, *77*, 89–96. [[CrossRef](#)]
104. Mauss, M. *The Gift: Forms and Functions of Exchange in Archaic Societies*; Cohen & West: London, UK, 1969; ISBN 0710018088.
105. Graham, S.; Marvin, S. *Splintering Urbanism: Networked Infrastructures, Technological Mobilities and the Urban Condition*; Routledge: London, UK, 2001; ISBN 0415189640.
106. Özden-Schilling, C. The infrastructure of markets: From electric power to electronic data. *Econ. Anthropol.* **2016**, *3*, 68–80. [[CrossRef](#)]
107. Smale, R.; van Vliet, B.; Spaargaren, G. When social practices meet smart grids: Flexibility, grid management, and domestic consumption in The Netherlands. *Energy Res. Soc. Sci.* **2017**, *34*, 132–140. [[CrossRef](#)]
108. Morley, J.; Widdicks, K.; Hazas, M. Digitalisation, energy and data demand: The impact of Internet traffic on overall and peak electricity consumption. *Energy Res. Soc. Sci.* **2018**, *38*, 128–137. [[CrossRef](#)]