Towards resource-efficient and service-oriented integrated infrastructure operation

Katy Roelich, Christof Knoer, Julia K. Steinberger, Liz Varga, Phil T. Blythe, David Butler, Rajat Gupta, Gareth P. Harrison, Chris Martin, Phil Purnell

ABSTRACT

Infrastructure is a means to an end: it is built, maintained and expanded in order to enable the functioning of society. Present infrastructure operation is characterised by: governance based on unmanaged growing demand, which is both inefficient and ultimately unsustainable; lack of integration of the end-users, in terms of the variety of their wants, needs and behaviours; separate and parallel delivery of different infrastructure streams prohibiting joint solutions. To achieve long-term sustainability, infrastructure needs to be designed and operated to provide essential service delivery at radically decreased levels of resource use. This new approach will need to: (1) incorporate the end-user, in terms of their wants and behaviours; (2) focus on the service provided; (3) use Information and Communication Technologies more effectively; (4) integrate the operation of different infrastructure systems; (5) be governed in a manner that recognises the complexity and interconnectedness of infrastructure systems; and (6) rethink current infrastructure valuation. Possible configurations incorporating these aspects with the explicit goal of contributing to long-term sustainability could be Multi-Utility Service Companies or “MUSCos”. This article presents new insights and ideas generated by considering the challenge of the transition towards a MUSCo infrastructure.

1. Introduction

Infrastructure is a means to an end: it is built, maintained and expanded in order to enable the functioning of society. In turn, however, the technical building blocks of infrastructure and its geographic layout determine, to a large extent, the level and composition of a society's resource demand, leading to long-term locking in of certain types of resource dependency and uses (Unruh, 2000). For example, on average the UK replaces its infrastructure at a rate of 1–2% per year, meaning that the decisions we make today regarding infrastructure will be with us for 50–100 years. Perhaps more surprisingly, physical infrastructure also shapes the institutional and social organisation of a society, through a historical process of change and evolution described as “co-evolution” (Foxon, 2011). This implies that changing infrastructure operation necessarily
involves larger social and institutional shifts as well as technical improvements.

Infrastructure is commonly referred to as the physical networks of water and energy supply, communication, transportation, and waste removal and treatment (e.g. Infrastructure UK, 2011), but increasingly incorporates the built environment as well. The present form of infrastructure operation consists of separate supply systems provisioning unconstrained demand: the demand is viewed unquestioningly as the “needs” of society which infrastructure must, somehow, reliably supply. This mode of operation has served industrialised societies well so far. However, defining the mission of infrastructure as the unerring supply of a growing demand is risky and ultimately unsustainable. Unlimited growth in demand means unlimited pressures on ecosystems and natural resources at a time when we are already well beyond our planetary safe operating space (Rockstrom et al., 2009).

From the perspective of societal resilience and security of supply, a system which understands and manages demand is arguably more robust than one of unlimited dependence on external, most often imported, inputs (Foresight, 2008). Achieving this outcome with the current configuration of the infrastructure system presents a number of challenges:

1. Governance based on unmanaged growing demand is both inefficient and unsustainable (e.g. increasing capacity of road transport networks leading to higher road usage and similar congestion known as “induced traffic” (Hills, 1996; Goodwin, 1996; Noland, 2001; Metz, 2008));

2. Current design and operation do not integrate the end-users, in terms of the variety of their wants and needs, and behaviours (e.g. car ownership as unique transport mode) and their crucial role in selecting and using technological options (e.g. selection and appropriate operation of energy efficient technologies); and

3. Separate and parallel delivery of different infrastructure streams prohibits the development of potential joint solutions (e.g. co-treatment of waste and wastewater), or even substitutions (e.g. substitution of electricity with gas through micro-combined heat and power (CHP)), between infrastructure systems.

These characteristics of current infrastructure operation act as obstacles to technical innovation and longer term sustainability. To achieve long-term sustainability infrastructure needs to be designed and operated with the goal of providing essential service delivery at radically decreased levels of resource. This requires a new approach to research that goes beyond analysis of individual failures to take a more systemic view of the purpose of infrastructure. This new approach will need to:

- Incorporate the end-user, in terms of their wants, behaviours and technological choices;
- Be focused on the service provided (e.g. thermal comfort) rather than supply of the vector (e.g. gas);
- Use Information and Communication Technologies (ICTs) and data more effectively to connect end-users to infrastructure systems;
- Integrate the operation of different infrastructure systems;
- Be governed in a manner that recognises the complexity and interconnectedness of infrastructure systems; and
- Rethink current infrastructure valuation.

We termed configurations that incorporate most of these aspects with the explicit goal of contributing to long-term sustainability, as Multi-Utility Service Companies or “MUSCos”. A MUSCo is an entity which delivers services to end-users, as opposed to electricity, gas, petrol or water. Since the payment to the MUSCo is on the basis of service, the costs of energy, water and material resources required for realising that service are internalised by the MUSCo: the MUSCo doesn’t profit from selling energy or other resources, it profits most by saving them, by providing the highest level of service at the lowest level of resource used (Steinberger et al., 2009; Stahel, 2010). The relation between a MUSCo and its customers would be established through performance-based contracts on agreed levels of service delivery, rather than by metering and billing physical resource use. A MUSCo could be a single organisation, or more likely a coalition of several different entities (including local authorities, technology providers and maintenance providers, for instance). A MUSCo would address the interconnectedness of infrastructure, since a single service required by the end-user could potentially be delivered through different technological options, using different mixes and qualities of physical resources. This definition of a MUSCo is clearly inspired by the niche market of Energy Service Companies (Marino et al., 2011; Hannon, 2012), but extends it beyond energy and beyond the usual business-to-business arena with the goal of contributing to massive, systematic economy-wide resource savings of the scale required by climate change mitigation, for example.

This article discusses this novel approach, and presents new insights and ideas generated by considering the challenge of the transition towards MUSCo-like infrastructure configurations. The first part of the article is concerned with the micro-level, starting with an analysis of the infrastructure as though the end-user mattered, including infrastructure services, ownership and control, and service quality, followed by a section on data requirements for enabling action on the user–infrastructure interface. The macro-level is covered in the second part, where the complexity and governance of integrated infrastructure operation are addressed, and the valuation of infrastructure is discussed. We conclude with a discussion of the findings and describe research required to accelerate the transition towards a more resource-efficient and service-oriented infrastructure bridging the gap between the micro- and macro-scale.

2. Infrastructure as though the end-user mattered

2.1. Does the end-user matter?

Society’s ultimate requirements of water, energy, communication, transportation and waste removal (in terms of overall volume and peaks of provision) are dictated by the aggregate demand of end-users. Measuring efficiency using the volume of utility products delivered (measured as kWh or passenger kilometre travelled, for instance) as the quantity to optimise, whilst excluding the end-users who set the level of demand, prevents the consideration of some of the most effective measures. Some would argue that demand management
includes the end-user in this system, and indeed there are increasing examples of demand management measures across infrastructure sectors to shift peak demand (e.g. energy (Carley, 2012), water (Russell and Fielding, 2010), and transport (Eriksson et al., 2006)). However a narrow focus on peak demand neglects the crucial role of the end-user in terms of their wants, behaviours and technical choices, which are critical to infrastructure operation in general, and to resource-efficient infrastructure operation in particular. We elaborate on these additional roles below.

The most promising cost-effective technologies for efficiency improvements are often found at the end-user level (McKinsey and Co, 2009; WB, 2009). Involving multiple and diverse actors and measures (as is the case with end-user centred measures) might be more difficult to implement than a small number of large scale supply side initiatives. This should not prevent action at this scale, since such distribution of measures across the actor landscape allows for faster technology learning and better risk distribution (environmental and economic) in case of failure (Grubler, 2010; Grubler et al., 2012).

End-user attitudes, beliefs, habits or routines, personal capabilities, and contextual factors have been identified as barriers to the adoption of cost-effective technologies in studies across different infrastructure streams (e.g. transportation (Eriksson et al., 2006), energy (Faiers et al., 2007), water (Russell and Fielding, 2010)). Lack of information about emerging technologies and/or personal resource performance (e.g. energy and water consumption (Faruqui et al., 2010)), as well as the economic ownership and benefit conflicts (e.g. principal agent problem (Murthiash and Sathaye, 2006)) appear frequently among the contextual factors limiting the exploitation of efficient technologies' full potential (Steinberger et al., 2012; Jaffe et al., 1999).

In fact, new technologies co-evolve with end-users' behaviour and organisational networks (Janssen and Jager, 2002; Rycroft and Kash, 2002) potentially leading to lock-in situations (Unruh, 2000). Furthermore, the end-user provides a point of integration for infrastructure streams since certain services (e.g. hygiene, mobility) require a combination of infrastructure networks. Such bundling of different infrastructure streams at the end-user level allows for new cross-sectorial efficiency solutions through novel integrated contractual agreements. Therefore, the end-user is essential for adoption of cost effective technologies and provides a key role in integrated infrastructure systems.

2.2 Delivering infrastructure services

In the last decade, researchers from different fields have put forward ideas for overcoming end-users’ efficiency barriers under the titles of performance, functional or service economy (Stahel, 2010; Mont and Tukker, 2006). In parallel, similar ideas developed from the business and marketing perspective (Gronroos, 2011; Vargo and Lusch, 2008) promoting the co-creation of value for both customers and providers, and from measuring ecosystem services (Loring et al., 2008). In general these ideas require a fundamental shift: away from selling products or metered quantities of utility products (e.g. kWh of electricity, gas or litres of water), and towards selling “services” (such as thermal comfort, illumination and cleanliness), which can be defined as the ultimate goal of the product or utility product purchased.

Fig. 1 illustrates such a shift in the energy system by comparing the traditional throughput-based energy economy, where transactions are based on units of products delivered, with a performance-based energy economy, where profits are based on energy savings (Steinberger et al., 2009).

At the product level the term ‘product service systems’ (PSS) is often used to describe such new commercial arrangements (Mont, 2002). They are generally based on leasing, allowing more flexible ownership which might lead to extended product life-time and higher efficiency level in the supply stream (Mont and Tukker, 2006). However, PSS usually do not include contractual agreements on guaranteed resource savings, or reference to baseline end-user performance.

At the infrastructure level, the potential and diffusion of Energy Service Companies (ESCos) have been analysed and reviewed in many countries, and across a variety of sectors (Marino et al., 2011; Vine, 2005). The ESCo business model relies on a guaranteed provision of energy service provided at a lower level of energy consumption (Sorrell, 2007). Although previous studies agree on the beneficial nature of ESCo operation for the implementation of energy and cost-efficient technologies (IEA, 2012), they also point at the considerable obstacles related to mainstreaming ESCos. High transaction costs for small clients and high asset specificity for large customers are currently the main barriers for further ESCo mainstreaming, particularly in the domestic sector. Standardized contracts, monitoring and accreditation schemes for ESCos might be required to overcome the above mentioned barriers (Sorrell, 2007). This in turn might increase the adoption of service-oriented contracts and supports a transfer of the energy service model to other infrastructure systems.

It is anticipated that service-oriented models would not only increase adoption of existing resource efficient and cost-effective technologies but could also bring about a change in the operation of these technologies, through contractual clauses concerning accepted modes of operation. A prominent example of this is car sharing, where a change in operation mode drastically increases efficiency whilst remaining cost-effective for a large part of the population (Pretenthaler and Steininger, 1999). Such a shift from a product to a service economy provides a fundamentally different perspective of how end-users could satisfy their wants and needs at reduced levels of resource consumption.

2.3 Technology choice vs. behavioural change

The resource consumption of a service is defined by the end-users’ final demand as well as how efficiently utility products are converted into the service required. The efficiency of such conversion itself depends on the conversion technologies in place (i.e. active appliances converting utility products into the service required), their passive context (Cullen et al., 2011; Cullen and Allwood, 2010) and how end-users operate these technologies. Current efficiency initiatives either focus on changing end-users’ final demand (i.e. behaviour) or end-users’ conversion technologies.

The majority of the relevant literature is focused on energy and water. Energy receives the broadest coverage, where efficiency initiatives target a change of conversion technologies...
and behaviour separately as well as in combination (e.g. Carley, 2012; Faiers et al., 2007; Faruqui et al., 2010; Linden et al., 2006; Boardman, 2004; Peacock and Newborough, 2005). Besides some retrofitting and grey-water reuse, most initiatives in the water sector concentrate on end-users' behaviour (e.g. Russell and Fielding, 2010; Olmstead and Stavins, 2009; Arregui et al., 2012; Willis et al., 2011; Carragher et al., 2012). In the transport sector the behaviour related topics are teleworking and eco-driving (e.g. Mokhtarian and Salomon, 1997; Barkenbus, 2010), and changing commuter schedules to move away from times of peak demand (Ben-Elia and Ettema, 2009; Hensher and Puckett, 2007). Changing transportation modes corresponds to the choice of conversion technologies (e.g. Eriksson et al., 2006; Smith, 2008).

Although behavioural change plays an important part for a transition towards more resource-efficient infrastructure operation, the active appliances in place limit its impact. We define active appliances as appliances which consume considerable amounts of energy and/or resources during their use phase. With the current development of increasing efficiencies a high replacement rate is key but might lose its importance once the efficiencies are maximised (Fernandez, 2001; Kim et al., 2003; Young, 2008). Apart from their actual design life, their replacement rates largely depend on the financial model of ownership. The two generic types of ownership are the classic product purchase, and an operating lease; where the lessor retains ownership and certain maintenance tasks and the lessee has possession and use of the asset for a prefixed payment period (Clark, 1978). In the classic ownership model active appliances are usually replaced at the end of their lifetime, independent of efficiency improvements of new products. Replacement rates for end-user conversion technologies are significantly higher in leasing agreements and increase further if the product life span is optimised for energy consumption, and remanufacturing of the asset is included (Intlekofer et al., 2010). Offering attractive loans combined with rebates has been found most effective in influencing residential choices towards high-efficiency appliances (Train and Atherton, 1995).

Another key aspect of active appliances is their embedding in a passive context, such as the built environment. Cullen et al. (2011) show that in building systems the largest energy savings are possible through improvement of the passive heating/cooling system (98% of practical available saving potential untapped) although appliances themselves offer considerable improvements (67% available savings untapped). Such improvements are based on mature technological innovations such as the Passivhaus standard, but will only achieve their full potential when appropriately used and maintained, which requires changing end-users behaviour (Hinnells, 2008). Appropriate use and maintenance is more likely to occur in mixed ownership models, where end-users and utility providers share risk and revenues (Steinberger et al., 2009; Sorrell, 2007), see also Bertoldi et al. (2013) for a further discussion. Hence high performance and guaranteed efficiency delivery are most likely to occur when a change from ownership to leasing-with-savings model for active appliances is encouraged, and end-users are supported with expertise, financing, and maintenance. Efficiency initiatives should therefore target end-users’ service demand, as well as active appliances, their passive context and an appropriate use and maintenance of these conversion technologies.

2.4. Bespoke quality

Utilities are provided to end-users at quality standards that have been established principally as a result of historical events, rather than practical needs (often referred to as path dependency), and are enforced by strict regulatory measures. A clear example of this path dependency and lock-in is power generation and transmission, commonly referred to as “the war of currents” (e.g. McNichol, 2006). Historically power was
generated at large, centralised plants and transferred over increasingly large distances to end-users, requiring high voltages during transmission (to limit potential transfer losses) but low voltages at the end-user (for safety reasons). More efficient voltage transformers for alternating current (AC) compared to direct current (DC) transformers favoured AC for power transmission, and led to a physical lock-in to a highly standardized AC power system.

These standards are provided and regulated regardless of whether they are still required to meet the needs of end-users. If we start from the point of view of the actual service required, which is usually the output of an active appliance converting the utility product, not all services would need utilities of the same quality, and we could accept some flexibility in the quality and reliability of utilities provided. Instead of conventional and arbitrary standards, bespoke quality or multiple qualities would in fact match the characteristics of service-delivering technologies and appliances, and could therefore lead to a reduction in total resource consumption.

Examples of bespoke quality leading to more efficient service delivery can be found in each type of infrastructure. In the case of water, only 50% of the domestic supply needs to be potable (Butler et al., 2011; Environment Agency, 2007). There are a number of uses, such as garment washing, toilet flushing and plant watering, which could be supported by water that has been treated to a lower standard such as grey-water (i.e. wastewater from all sources except the toilet) or treated sewerage effluent. For electricity, this is even more dramatic, as shown in Fig. 2, where the majority of the current household communication and lightning appliances actually use DC and have to be supplied with rather inefficient (i.e. about 50%) AC/DC converters (Calwell and Reeder, 2002). The bulk consumption appliances for sustenance and hygiene (e.g. refrigerator, freezer, dishwasher, washing machine and tumble dryer) currently rely on AC. However, they could be replaced with appliances using more efficient and less maintenance intensive DC pumps and motors, which would offer a 60% efficiency gain (von Weizsacker et al., 1998). For transportation, the majority of car use is for recurring single person transport trips (e.g. commuting and education) (DFT, 2011), and thus replaceable with public transport, car-pooling or car sharing schemes. In the case of gas, high quality and pressure standards restrict the use of potentially more sustainable fuels such as biogas and hydrogen from waste. Another example is district heating with various qualities required for residential, commercial and industrial heating purposes. In the domestic context room temperature is rarely above 20–22 ºC provided through comparably small high temperature heaters, establishing a technical barrier for efficient low temperature central heating systems. In addition hot water is often provided over 50 ºC although the mixed used is only 37 ºC (Environment A, 2007).

An example of a well-accepted bespoke quality service delivery is waste removal infrastructure, where different quality waste streams are segregated (DEFRA, 2011). Separate waste storage containers and different collection vehicles are provided according to the waste type. As a consequence, recycling and higher resource efficiency are possible. When waste separation systems are managed appropriately, the resource to production cycle is more effective as a result of reduced processing requirements and contamination (Bovea et al., 2010).

Technically, it would be possible to provide utilities at multiple bespoke quality levels better adapted to the needs of the end-user: for example providing non-potable water or direct current power to support appropriate services. These bespoke quality utilities could be delivered through central distribution, or could be supported by local generation, for

![Fig. 2. Total 2011 UK domestic power consumption by different services and form of electricity [TJ/year] (i.e. alternating current (AC) red arrows and direct current (DC) in yellow arrows). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.) Data from DECC (2011).](image-url)
example through rainwater collection, or decentralised power or heat generation. Moreover, many of the technologies for providing bespoke service quality are readily available and implemented for off-grid applications at remote locations or in mobile homes.

Of course, despite the potential efficiency benefits, there are several challenges associated with an implementation of a bespoke quality service delivery.

2.4.1. Regulation

Service delivery, in particular for water and electricity is currently heavily regulated, in the former case led from the EU, making change a difficult and long-term endeavour. Bespoke quality would not only imply a change in service delivery regulation, but require new construction (e.g. plumbing and wiring) and appliance standards. Pilot-level exceptions from regulation may be a necessary intermediate step.

2.4.2. Safety/quality

Safety and quality issues are a core concern of regulatory development, shifting safety considerations from the end-user to the provider. Bespoke quality services, such as higher voltage DC or lower quality water, transfer some of the responsibility to the end-users, and forces them to take a more active role in their safety. For non-potable water delivery for example, quality issues have been found to be the main challenge (Dixon et al., 1999).

2.4.3. Reliability

Current infrastructure is understandably designed for high levels of reliability. Paradoxically, for most uses, with the exception of transportation, a much lower level of reliability would be acceptable (e.g. Sullivan et al., 1996). This could result in significant resource savings by reducing back-up needed to support reliability (e.g. gas power plant operation at most efficient loads), and would also help integration of renewables (e.g. biogas in gas networks).

2.4.4. Infrastructure supply network investments

Massive sunk investments in current physical infrastructure supply networks (e.g. power and water), represent a barrier to additional central generation of bespoke quality infrastructure products. Local generation of such infrastructure products (e.g. in-house central AC/DC conversion and DC distribution) could be one way to overcome this barrier and open a market for appliances using such bespoke quality directly (e.g. large domestic DC appliances).

2.4.5. End-user practice

High standards allow for a maximum flexibility of end-user behaviours, which would be necessarily limited by a bespoke quality delivery. Such intervention may lead to additional resistance to accepting new standards. More proactive consumers however could demonstrate benefits (e.g. reduced cost) and facilitate a transition towards resource efficient infrastructure operation, in accordance with smart appliance providing user-friendly feedback, and able to cope with safety and reliability issues.

3. From data to action

3.1. Challenges enabling action through feedback

Integrating the end-user into infrastructure operation and realising MUSCo models of infrastructure require an appropriate interface to be created between the user and the infrastructure. Such an interface must provide feedback to the user — in the form of data, information and knowledge captured and processed by interconnected Information and Communication Technology (ICT) systems (i.e. smart systems) spanning the different infrastructure streams. Providing feedback at the user–infrastructure interface has the potential to reduce resource demand, by supporting resource efficient decision-making and catalysing behaviour change. Understanding the nature of this interface, and the type of feedback it would provide, is therefore critical to achieving the goal of our work — to incentivise and ensure resource-efficient operation at the end-user level.

In the future, it is likely that the user–infrastructure interface will present feedback to the user from across interconnected smart systems — e.g. smart home and building management systems, and smart electricity, water, heat or transport networks. Such systems will be underpinned by the pervasive deployment of ICT across infrastructure — employing technological paradigms such as the Internet of Things (Atzori et al., 2010). For example, in the context of the electricity network, a smart electricity grid with innovative dynamic tariff structures is envisioned to play a key role in demand reduction and demand response, providing feedback to the user and promoting behaviour change (Rahimi and Ipakchi, 2010).

Many competing interests and perspectives (e.g. technological, ethical, and security) will need to be reconciled when designing an appropriate user–infrastructure interface. To date, there has been limited success in the design of applications at this user–infrastructure interface. For example, smart (electricity) meters providing feedback on consumption generally have a limited impact on informing decision-making or catalysing behaviour change (Darby, 2010; Hargreaves et al., 2010; AECOM, 2011). In the following we elaborate on the key challenges in designing the user–infrastructure interface.

3.1.1. The user-orientated, integrated infrastructure vision calls for an atypical use of feedback

Typical feedback, provided by ICT systems, is either employed in a business context to enable organisational objectives to be met; or presented in a domestic context as a by-product of fulfilling a specific user need. At the user–infrastructure interface, feedback would be presented to the user across domains (domestic, leisure, work and transit) to enable an overarching societal need to be met (i.e. resource efficiency). In this case the usual drivers of organisational control and specific user need are both absent, with the feedback provided seeking to make users aware of often unconscious and habitual consumption behaviours.

3.1.2. Interoperability in a complex system of actors with competing interests

Ensuring that feedback can be presented at the user–infrastructure interface will require the interoperability (i.e. the ability to interact directly to exchange and use data) of smart
systems across multiple infrastructures. These systems need to be able to communicate with each other to provide feedback to users in a consistent way, whilst appliances connected to the systems should not interfere with each other’s operation. A review of building automation interoperability (Ferreira et al., 2010) suggests that major interoperability challenges are likely where subsystems come from multiple manufacturers and are developed on heterogeneous technologies.

3.2. Guiding principles in designing the user–infrastructure interface

3.2.1. Big data

The large-scale adoption of smart systems and infrastructure (as discussed above) has the potential to create a deluge of ‘big data’ (Hey and Trefethen, 2003). Deriving value from such vast quantities of data presents major potential challenges in processing data to creatively distill and communicate the information and knowledge that will drive resource efficient decision-making (at individual user, community and organisation scales). Examples of the potential impact of ‘big data’ are emerging, for instance on a community or neighbourhood scale, through spatial mapping of energy consumption and related CO$_2$ emissions on a house-by-house level (Gupta, 2009).

3.2.2. Ensuring automated decision-making is used appropriately

Striking a balance between automation and allowing users to make consumption-related decisions (rather than seeking maximum degree of automation) offers the possibility of reducing demand in a way that integrates with the user’s lifestyle — e.g. in a smart home context (Rashidi and Cook, 2009), and enables learning from their choices to deliver bespoke decision support.

In order to drive and enable demand reduction it will be necessary to create a two-way flow of data, information and knowledge across the user–infrastructure interface. Such a flow creates both the potential for misuse of data, information and knowledge, and associated user concern about potential misuse, that must be addressed — e.g. the privacy concerns associated with smart electricity grids (McDaniel and McLaughlin, 2009).

There is little doubt about that pervasive ICT systems will underpin future infrastructure systems at the user–infrastructure interface due to their benefits on various levels. Without coordination however, they might provide inappropriate feedback, lead to over-automation, and open a range of concerns about data security. Coordination between ICT system and appliance manufacturers, service providers, and end-users on the other hand offers significant potential for a more resource efficient infrastructure operation. This has to be facilitated by considering the complex interrelations between infrastructure streams, and through innovative governance and infrastructure operation schemes designed to overcome existing market and regulatory failures. Cross-organisational e-contracting is one of the examples where complex interrelations and business processes are facilitated through more elaborate use of ICT (Rouached et al., 2005).

4. New thinking through infrastructure integration

4.1. Complexity

The transition to a new, user-centric approach to infrastructure operation requires a systemic approach to change. Infrastructure cannot be understood as a set of discrete technologies but must be seen as a complex, interconnected system of technology embedded in society and the environment, interacting with public and private institutions; it can be described as both a socio-technical system (Unruh, 2000) and an ecological-economic system (Costanza, 1989). Interconnections occur among different technical infrastructure systems (for example the use of energy in water and wastewater treatment) (CST, 2009; Rinaldi et al., 2001) and between technical and socio-economic systems (Foxon, 2011; Hall et al., 2012, 2014).

Interactions within this complex system mean that it is unstable and unpredictable: the outcome of intervention into the system is difficult to forecast; feedback loops make it difficult to distinguish cause from effect; there are significant time and spatial lags; and relationships are non-linear resulting in thresholds (transition points between alternative states) (Costanza et al., 1993; Liu et al., 2007). Small historic events can favour the initial adoption of a particular, often inferior, technology (such as the AC distribution system described in Section 2.4). Once a technology is established, systemic interactions between technologies, infrastructures, institutions and users can increase the returns to adoption of that technology and constrains the development and adoption of a superior one (Unruh, 2000; Arthur, 1989). This phenomenon of path dependency can lock us into inefficient technologies and behaviours and limit the potential of intervention to create disruptive change.
The process of change in complex systems can be explained using evolutionary concepts of variation, retention and selection (Norgaard, 1994; Nelson and Winter, 1982). Individuals and organisations can acquire (and drop) routines, skills, traits and assets and thereby rapidly evolve to match their environments. Mutual causal influences among technical systems and between technical and social systems mean that the activities in one system can influence selection criteria and replicative capacity in other systems (Foxon, 2011). The separate and parallel governance of technical infrastructure systems is not designed to respond to these influences, leading to financial and operational inefficiencies (CST, 2009).

Understanding change, or transitions, as a co-evolutionary process “highlights the uncertain, path-dependent and cumulative nature of systems change” and helps to identify interventions that recognise the interconnectedness of infrastructure, with the potential for overcoming system lock-in (Foxon, 2011); hence the proposal of radically new arrangements, such as MUSCos. MUSCos inherently cover a multitude of policy fields, including climate change, fuel poverty, energy security, water resource management and housing quality, engage a wide range of stakeholders and cross territorial boundaries. This presents significant challenges to the current siloed governance system and significant institutional reform is necessary to enable better inter-policy co-ordination and regional co-operation (Monstadt, 2007).

4.2. Market and governance failures

Infrastructure provides a public service; therefore, the services it delivers need to be acceptably reliable, of adequate quality and quantity, and offer value for money. Little (2005) and Frischmann (2005) go further and suggest that infrastructure is a common good which is non-excludable and requires continual replenishment to deliver reliability and quality.

Government (de)regulation is intended to ensure functioning of utility markets. State regulation is commonly applied to; ensure fair access and to prevent excessive pricing in natural monopolies, avert collusion and the building of anti-competitive market share, ensure right of access, consumer protection and the security of supply in retail markets (BIS, 2011). However, regulation has not always succeeded in achieving these aims. Regulations targeting specific market malfunctions often result in novel market problems, which in their turn demand a regulatory response, and so on. Some liberalisation policies, designed to increase competition in a former monopoly, can hinder long-term competition if they handicap some competitors on an ongoing basis (Armstrong and Sappington, 2006). We argue that a multi-utility service perspective would enable regulation aiming at the real benefits of public services, avoiding more short-sighted and over-targeted measures.

In the UK, for example, the privatisation of the electricity and water sectors, originally designed to overcome public sector inefficiencies, resulted in new problems and challenges. In a recent review of UK energy, Thomas (2006) notes that the British electricity reforms are viewed as a model worldwide, but that this reputation is undeserved. One of the most important issues is the failure of government intervention to develop efficient wholesale markets, which means that competition is only created if consumers switch suppliers regularly, which does not happen in practice. This is supported by a recent review by the UK energy regulator (OFGEM, 2011) which finds that consumers are at risk from market features that reduce the effectiveness of competition, including: complex pricing structures, which create an obstacle to switching suppliers; the dominant market shares of the incumbents, which acts as a barrier to new entrants; and, the disconnection between wholesale cost decreases and energy prices. In addition, the economic regulation system designed to enhance competition in the network via price controls has limited the investment of electricity network companies in transformative infrastructure. It also restricts their ability to make profit by any means other than increasing throughput, which disincentivises demand management and alternative models of delivery, such as Energy Service Companies (ESCos). These problems are also found in Canada and the USA that use similar privatization and regulatory models (Boland et al., 2011). Attempts have been made to move away from network price control to incentivise innovation, such as the UK’s new RIIO (revenue = incentives + innovation + outputs) model for gas and electricity distribution charges, which encourages investment in innovation, notably on the demand side (OFGEM, 2010).

Similarly, the privatization of water supply in England and Wales is regularly characterised as a success (Bakker, 2005), whereas in reality the government was unable to introduce true competition and cost-reflective pricing. This has reduced the effectiveness of market mechanisms and required a significant degree of subsequent regulation to improve water quality standards, encourage water conservation and control pricing. In some circumstances this has resulted in a reversal of privatization (Ohemeng and Grant, 2008). Indeed the narrow definition of economic efficiency used in regulation may actually deter water conservation (Bakker, 2005).

Existing infrastructure, market structures and governance structures have evolved in different jurisdictions in different ways (Stenzel and Frenzel, 2008). Major interventions by governments have often been motivated by ideological considerations, whilst regulations have been focused on addressing specific perceived problems in the context of the state of the infrastructure and governance regimes at a specific time. Commercial suppliers respond to government interventions in the context of their business strategies and perceived profit opportunities; and regulators respond in turn to perceived issues arising from those commercial responses. These activities tend to be narrowly focused on short-term issues such as consumer detriment, utility price or security of supply and relate primarily to individual supply streams. These mechanisms are not sufficient to address longer-term strategic issues such as climate change and the need to minimise the consumption of scarce natural resources. It is only by moving beyond the separate regulation of different infrastructure networks, and focusing on the actual services delivered by infrastructure, that environmentally-coherent and efficient regulation and markets can occur.

4.3. Innovative governance & operation

The complexity and interdependence of infrastructure systems, both technically and socio-economically, requires that any action to intervene must recognise both the historic
co-evolution of the system and the path dependency that could limit the potential of disruptive change. Overcoming lock-in requires systemic action that cuts across policy sectors, public and private institutional boundaries and state jurisdictions (Smith, 2007). This systemic action also needs to address both the cross-sectoral interactions during infrastructure operation and the integration of infrastructure systems at the end-user.

Precedents of the transformational change required stem from innovative governance and operation including; those where institutions have taken on a new role and intervened directly in infrastructure service provision, those where innovation and research is enabled or encouraged, those where alternative business models have driven transformative change in service delivery and resource consumption, and those that have exploited opportunities for infrastructure integration.

The key challenge is to identify the conditions under which these precedents can become mainstream, and in particular to identify appropriate governance arrangements. Many of the alternatives described below require the involvement of a broad range of actors in order to be innovative, successful and durable (Furlong and Bakker, 2010). Private companies are worse at engaging non-state actors and are less likely to deliver public benefits, which would suggest that an increasing role for public sector organisations in infrastructure delivery would be beneficial (Furlong and Bakker, 2010; da Cruz and Marques, 2012). It has been suggested that public–private partnerships (PPPs) might overcome the high resource costs of solely public sector ventures. However, PPPs are subject to a great deal of criticism, and face challenges in attaining a balance between public and private objectives, so arrangements must be appropriate to the particular project and capabilities of the actors involved (da Cruz and Marques, 2012; Sohail et al., 2005; Koppenjan and Enserink, 2009).

Direct intervention into service provision has been successfully implemented by Woking Borough Council, which recycles savings from energy efficiency into construction of low-carbon energy systems to provide affordable energy to residents. This initiative was made possible by a more integrated, outcome focused approach to problem framing, which aligned strategies from several different policy sectors (e.g. climate change mitigation and housing renewal). This resulted in more effective intervention that reduced negative unintended consequences (Smith, 2007; Leach et al., 2010). The barriers to this kind of approach include procurement rules, financing arrangements and an inability to prioritise whole-life cost over short-term funding cycles in the public sector (Hannon, 2012).

Investment in innovation and transformative research is essential to deliver disruptive technologies and practices (CST, 2009). This kind of research and investment can be encouraged if infrastructure is considered to be a strategic issue (Cooremans, 2007, 2012). There have been examples of infrastructure organisations investing in technologies, where a combination of drivers, such as climate change, energy prices and tight regulatory controls, increase the strategic importance of investment. Despite this, regulation could present a barrier, where it discourages investment through price regulation, as is the case in the UK water and energy industries. This is exacerbated in the UK water industry by short investment planning cycles which limit flexibility over payback periods.

In addition to innovating to deliver the same business model more efficiently, there is a growing trend in infrastructure companies, particularly energy companies, delivering different business models under the guise of Energy Service Companies (ESCOs) (Steinberger et al., 2009). These include Energy Supply Contracting (where the contractor is responsible for supplying useful energy) and Energy Performance Contracting (where the contractor takes responsibility for reducing throughput of energy needed to provide energy services) (Hannon, 2012). Drivers for adoption of these new business models include the increasing focus of regulation and stakeholders on sustainability and rising energy prices. However, regulation designed to disband monopolies and encourage cost efficiency could present barriers to wide-scale adoption (Hall et al., 2012).

Awareness of the opportunities presented through more integrated infrastructure operation is increasing. A recent review of water saving trials in the UK reported not just the water savings but also the energy and carbon dioxide savings that resulted from reductions in hot water consumption and energy required during water treatment (Waterwise et al., 2011). The report recommended a more integrated approach to the roll out of energy and water efficiency measures.

These alternative approaches provide promising examples of transformative changes at the niche scale; however, the scale of efficiency improvements required would necessitate a step change in the rate of adoption of these alternative approaches and support in application at a larger scale. We need a more creative and systemic approach to governance to encourage and enable this step change.

5. Infrastructure as though infrastructure mattered

There are many problems related to the ways in which infrastructure is valued, leading to challenges in investing in sustainable infrastructure, or operating existing infrastructure sustainably. For instance, widespread capital depreciation rates may not apply to long-lived infrastructure. Moreover, the value of infrastructure, as a public service, in enabling the functioning of societies, is notoriously difficult to quantify, since the positive (and negative) impacts of infrastructure are often indirect (through intermediary activities and actors) and diffuse in space and time.

In particular, the operational phase – i.e. post-installation and commissioning – of infrastructure is systematically undervalued. A cost is attached to the construction of infrastructure assets, such as road and bridges, and often to maintaining assets to ensure they continue to serve their original function, but not to the value generated by their utility in delivering a service to society (Little, 2005). Infrastructure undervaluation increases decay, since investors are reluctant to fund the necessary initial quality to ensure longevity, and/or maintenance programs get squeezed. This inhibits investment in more sustainable systems, since the returns, i.e. return on value, are unclear (CST, 2009; Infrastructure UK, 2010). A further challenge is the need for infrastructure to be operated in an openly accessible manner that allows all users to access infrastructure services on equal and non-discriminatory terms (Frischmann, 2012). This makes it difficult for us to determine how to ‘make money’ from the sustainable operation of infrastructure.
In most charging schemes, the end-user does not pay for the service and well-being provided, but for some indirect or even unconnected measure thereof (e.g., gas, not heat; a flat-rate for water supply and waste disposal; and road tax, not miles travelled). This focus on the physical infrastructure rather than the service it provides is unhelpful; however it arises because it is easier to measure cost/price than it is to measure value. Truly cost-reflective pricing is further hindered by regulated price controls deemed necessary to ensure non-discriminatory access (for example in the water industry) or indirect cost recovery (for example through road taxing) [NAO, 2013]. The narrow definition of economic efficiency used, particularly when setting pricing formulae, incentivises increasing demand as a means to increase profits (Bakker, 2005). It actively deters operating regimes which maximise resource efficiency.

An alternative approach to infrastructure valuation during operation would be to quantify the service delivered by the infrastructure, such as thermal comfort or access to employment, education and leisure. The valuation would also include indirect services — the value provided by the other services enabled by this asset i.e. as a result of interdependency/interconnections. It should be noted that the valuation of infrastructure services should not be equated to charging customers for this value: charging schemes should take into account considerations of equity, progressive pricing (affordable for basic services, more expensive for luxury uses), but also include environmental externalities. A full discussion of the valuation-charging topic is obviously beyond the scope of this paper, but would be interesting for future research.

There are further benefits to an infrastructure valuation approach centred on services: the focus of investment is likely to move from investment in new capacity to investment in demand management, since this is where the most significant cost savings can be found; some assets might be retired or converted to a new use because they are not central to delivering the service; and maintenance regimes are likely to become more active and consider the long-term value of the infrastructure asset in delivering the desired service.

6. Getting there from here

Long-term sustainability in infrastructure operation can only be achieved through a radical transition from the current mode of operation to one incentivising essential service delivery at the lowest possible level of resource use (Steinberger et al., 2009). This paper discusses a number of issues that need to be addressed to enable this transition away from supply of unmanaged demand towards resource-efficient service delivery. In the following we synthesise our findings from considering micro- and macro-level changes.

6.1. Micro-level

The importance of engaging end-users: Since the most promising resource efficiency technologies appear to be on the demand side, most of the identified barriers and enablers for more efficient infrastructure operation through diffusion of efficient technologies are end-user related (Grubler et al., 2012; Wilson et al., 2012).

The potential of the service/performance perspective: which offers an alternative view on how efficient, cost-effective technologies and alternative operation modes could be brought into the mainstream.

Technological vs. behavioural change: resource consumption of a service is ultimately determined by the end-users’ final demand level, choice of active appliance, its passive context (Cullen et al., 2011; Cullen and Allwood, 2010) and patterns of use and maintenance.

Bespoke quality instead of arbitrary standards: Bespoke quality, or multiple qualities, are better matches for the actual characteristics of services or resources needed, rather than conventional or arbitrary standards, and could lead to a reduction in resource consumption.

The interface between the user and infrastructure: by providing feedback, this will enable resource efficient decision-making and catalyse behaviour change, but only if provided in a user-friendly way.

6.2. Macro-level

Infrastructure systems are complex and interconnected: we need to recognise this interconnectedness and examine the co-evolution of technology, end-users, businesses and institutions and their causal interactions.

The challenge of problem shifting: privatisation of infrastructure, intended to overcome public sector inefficiencies, has obliged government to regulate to address new types of market failure. We need to recognise complexity and focus on system-wide goals to avoid shifting problems back and forth between markets and regulation.

More innovative governance and operation: this might include alternative problem framing and investment strategies as well as integrating infrastructure streams, including new actors at different scales.

Appropriate valuation of infrastructure services: infrastructure, in particular in its operational phase, is undervalued, which drives decay, misdirects investment, and incentivises demand.

6.3. Next steps

These challenges cannot be addressed in the current, throughput-based system of infrastructure operation. We need to accelerate the transition to alternative infrastructure operation configurations which incentivise resource efficiency. Multi-Utility Service Companies, or MUSCos, the new type of actor we described in the introduction could contribute to this transition since they are: centred on the end-user and their demand for services; concerned with implementing resource efficiency improvements; and take into account multiple utility streams simultaneously.

The widespread adoption of a MUSCo configuration is currently constrained by a number of barriers including lack of trust and experience, lock-in to mainstream technologies and operation modes, high transaction costs associated with creating and monitoring contracts and the fragmented and entrenched regulatory and policy framework (Roelich et al., 2013). There is a need for innovative research involving partnerships between academia, localities and industry, which investigates particular aspects of the challenge, but hopefully builds towards a system-wide transition. The academic side of the partnership should moreover be interdisciplinary, including
expertise from social science (sociology, psychology, political science), economics, and engineering, as well as computer science, for example, and also integrative, so that the outcomes are not reflective of one single field, but rather bring together the different disciplines to provide answers to core research questions. For example, insights from the micro- and macro-level considerations we outline above could be integrated into a socio-technical modelling framework for use in order to test theories or hypotheses of different factors of change in within this complex system.

Indeed, we recommend a multi-scale approach to this new research area. Micro-studies, focusing on a small number of end-users, or specific technologies, should be conducted with a high level of detail, in order to bring out unexpected outcomes, and generate results that are reflective of the diversity of end-user behaviour even within a specific neighbourhood or industry sector. Meso-studies, at the community or industrial sector level, are important because they will highlight the crucial role of diverse actors, in terms of their relations, expertise and incentives, including contradictory incentives, but also possibly unexpected new types of relations and collaborations which will be necessary for MUSCo infrastructure operation. Macro-studies, comparing infrastructure governance and policies (including procurement policies) are necessary to understand and change constraints on service-based infrastructure from high-level regulatory bodies. This research should be conducted across Europe, in different national contexts, in order to develop a European strategy on sustainable infrastructure operation which maximises resource efficiency and priorities service delivery over throughput volume.

Truly sustainable infrastructure operation needs to move away from the status quo based on provisioning unconstrained demand, and the research outlined above will chart and accelerate the way forward.

Acknowledgements

We are grateful for the comments of anonymous reviewers. This research was conducted as a part of the EPSRC funded project “Land of the MUSCos” (grant number: EP/J00555X/1).

References


Katy Roelich is a Senior Research Fellow at the Sustainability Research Institute and the Institute of Resilient Infrastructure at the University of Leeds. Prior to this she was a Senior Research Fellow at the Stockholm Environment Institute where she co-directed the Rethinking Development theme, and a consultant providing sustainability and resource management advice to engineering and planning projects and policy development. Her research focuses on the governance of sustainability transitions.

Christof Knoeri is a Research Fellow at the Sustainability Research Institute at the University of Leeds and Visiting Lecturer at the Chair for Human-Environment Relations at the University of Munich, Germany. Prior to this he was a Research Associate at the Life-Cycle Assessment and Modelling Group at Empa Materials Science & Technology, Switzerland. In addition, he worked as a Guest Lecturer at the Institute for Systems Science, Innovation & Sustainability Research at the University of Graz, Austria. His research interests are sustainable transitions in socio-technical systems in general and the application of modelling techniques like agent-based modelling in particular.

Julia K. Steinberger is an Associate Professor in Ecological Economics at the University of Leeds. She is the Principal Investigator of the Land of the MUSCos project. Her field of research focuses on measures to achieve human wellbeing at the lowest level of resource use and environmental degradation. Her research areas include applying novel econometric tools to international environmental, economic and social datasets, as well as modelling the resource use implications of infrastructure choices, in particular in urban environments. She was a Lead Author of the Global Energy Assessment chapter on Urbanisation and is on the Council of the International Society for Industrial Ecology.

Liz Varga is a Principal Research Fellow and the Director of the Complex Systems Research Centre at Cranfield University. Liz has a number of research grants located within infrastructure systems, each of which has a strong modelling component with a focus on demonstrating the application of complex systems models to practice. Recent projects include multi-modal freight logistics (ABIL) and smart grids (CASCADE). Current projects include utilities infrastructure conversion (TUCP), Multi-Utility Service Companies (MUSCos) and International Centre for Infrastructure Research (ICfI). Her work makes a general contribution to the co-evolution of infrastructure systems and to approaches that have historically rarely moved from the purely academic realm.

Phil T. Blythe is Director of the Transport Operations Research Group and Professor of Intelligent Transport Systems at Newcastle University. His research focuses on the interface between technology and policy and in particular how new and emerging technologies can offer new possibilities on how to deliver a policy more efficiently, effectively and smartly. Current key areas of research include: electric vehicle and recharging infrastructure; eco driving; travel demand management; infrastructure pricing; accessibility; smart future cities and resilience.

David Butler is Co-Director of the Centre for Water Systems and Professor of Water Engineering at the University of Exeter. He is currently an Engineering & Physical Sciences Research Council Established Career Fellow. He specialises in urban water management including sustainable and resilient water systems, integration and control, urban flooding and water–energy–carbon interactions. Much of his research focuses on the interaction between infrastructure, its users and the environment.

Rajat Gupta is Professor of Sustainable Architecture and Climate Change, Director of the Oxford Institute for Sustainable Development (OISD) and leader of the OISD: Low Carbon Building Group at Oxford Brookes University, UK. As recipient of the inaugural 2006 RIBA President’s award for outstanding research, Professor Gupta has nearly 90 research based publications on carbon counting, global common carbon metrics, building performance evaluation, low carbon communities and climate change adaptation of buildings. As Principal Investigator, he has won over £5 million in research grants from Research Councils, Government and Industry. He also advises government at senior level and is on the boards of several industry task groups.

Gareth P. Harrison is Bert Whittington Chair of Electrical Power Engineering and Deputy Head of the Institute for Energy Systems at the University of Edinburgh. He holds a Bachelor’s degree and a Doctorate from the same institution and was appointed to staff in 2000. He leads research activity across a wide area including grid integration of renewable energy, renewable resource assessment, climate change impacts on electricity systems; and carbon footprints of energy systems. Professor Harrison is a Chartered Engineer, a member of both the Institution of Engineering and Technology and the Institute of Electrical and Electronics Engineers and is an Affiliate of the Association of Chartered Certified Accountants.

Chris Martin is the Principal Investigator for a Management and Business Development Research Fellowship funded by the ESRC, Society for the Advancement of Management Studies (SAMS) and the UK Commission for Employment and Skills (UKCES). His research draws on diverse disciplinary perspectives to develop understanding of the interconnections between innovation (social and technological), open data and public management and policy. Prior to his current role he worked as a Researcher and Interdisciplinary Research Facilitator in the Centre for Integrated Energy Research at the University of Leeds and a Principal Scientist within the Research and Innovation Functions of the Environment Agency of England and Wales, where he led a portfolio of projects employing novel web technologies to catalyse environmental behaviours and support environmental management.

Phil Purnell is Professor of Materials and Structures at the University of Leeds, and Director of the Institute for Resilient Infrastructure, ensuring that the physical infrastructure systems underpinning our way of life can adapt to changes, both in the way we use them and in the social and physical environment in which they are created, designed, built and operated. He is engaged in research into concrete technology, resilience and critical materials, and acts as a consultant to the civil engineering, nuclear, utility and materials industries.