

Assemblage theory, data systems and data ecosystems: The data assemblages of the Irish planning system

Big Data & Society
July–September 1–12
© The Author(s) 2025
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20539517251352822
journals.sagepub.com/home/bds



Rob Kitchin¹ , Juliette Davret¹ , Carla M Kayanan¹ and Samuel Mutter¹

Abstract

The work of the state has been transformed by digitalisation and datafication, with paper systems increasingly complemented and replaced by digital data systems. While much research has been undertaken on the effects of these data systems on e-government and e-governance, relatively little critical attention has been paid to the ontological nature of the data systems themselves, how they come into being, operate, iterate, and work in conjunction with other systems. In this article, we build on work that understands the nature of data systems as data assemblages; that is, as being thoroughly socio-technical in nature, emerging and unfolding contingently, contextual and relationally through material and discursive practices and processes. We extend the notion of a data assemblage through an engagement with assemblage theory, and detail how data systems are articulated and scaffolded into being, and scaled into a functioning data ecosystem, which is conceptualised as an assemblage of data assemblages. We illustrate the utility of this reconceptualisation through an analysis of the development and control function of the Irish planning system, which is composed of a number of interlinked data systems used for managing the entire planning process from seeking planning permission, monitoring construction, to completed property.

Keywords

Data assemblage, data ecosystem, data system, digitalisation, assemblage theory, e-government

Introduction

Since the birth of modern digital computing in the 1950s, the state has been undergoing a process of digitalisation as bureaucratic and operational practices transfer from paper-based systems to digital systems. This process of digitalisation is still underway, but states are now digital enterprises, with systems of e-government and processes of e-governance core modes of operation (Dunleavy et al., 2006; Falk et al., 2017; Datta, 2023). Accompanying this digital turn has been a ‘data revolution’, wherein digital data have become a vital asset for public bodies in managing their internal affairs, performing their operations, engaging with clients/customers and coordinating their relationships with other organisations (Mayer-Schönberger and Cukier, 2013; Kitchin, 2014). These twin processes of digitalisation and datafication have attracted significant critical attention in recent years, with a range of new concepts formulated for making sense of data regimes and the work data do in the world (Kitchin, 2025). Two such concepts are data assemblage, used to refer to the amalgam of technical

and contextual factors required to produce and utilise data (Kitchin, 2014), and data ecosystem, used to refer to a set of related data systems that work together to enact a data regime in a particular domain (van Schalkwyk et al., 2016).

In this article, we seek to advance and entwine the concepts of data assemblages and data ecosystems to: (a) advance a new over-arching conceptualisation of the organisation and operation of data-driven state work within and between public sector bodies and other entities at various scales; (b) map out the socio-technical construction and operation of this data landscape and how it forms, is held together and functions in practice. In the first part of the

¹Maynooth University Social Sciences Institute, Maynooth University, Maynooth, Ireland

Corresponding author:

Rob Kitchin, Maynooth University Social Sciences Institute, Maynooth University, Iontas Building, North Campus, Maynooth, County Kildare, Ireland.
Email: Rob.Kitchin@mu.ie



article, we consider how to make sense of the data landscape of state work. We contend that data systems are socio-technical and emergent in nature, constituting data assemblages. Here, we use data system as a descriptive term for defining an IT system designed primarily to capture, manage, process and share data, with data assemblage being a conceptual term for explaining the produced nature of a data system. We outline Kitchin (2014) and Kitchin and Lauriault's (2014) initial conceptualisation of data assemblages and extend it: (a) through the application of assemblage theory as conceived by Nail's (2017) reading of Deleuze and Guattari (1987); (b) by examining how data systems become interconnected to form a data ecosystem that discursively and materially shapes how data work is undertaken within a domain. This data ecosystem is itself an assemblage composed of many inter-connected data systems (an assemblage of data assemblages).

We illustrate the utility of this assemblage thinking for making sense of state data work using a case study of the planning system in Ireland, with a particular focus on the development and control pipeline: making and processing planning applications, undertaking stakeholder feedback and public consultation on applications, appealing decisions, monitoring construction and building control, and creating official statistics. We detail how processes of articulation and scaffolding are employed to initiate, assemble and implement data systems, and how they continue to iterate in response to prevalent conditions, actors and objects, by examining the establishment and on-going operation of the Irish Building Control Management System (BCMS). Finally, we chart how several data systems are topologically arranged and inter-connected through multiple seams and data mobilities to form a planning data ecosystem that sustains various forms of data-driven work along the entire development and control pipeline.

Understanding data systems as data assemblages

Kitchin (2014) defines a data assemblage as a complex socio-technical arrangement composed of many apparatuses, actors and practices whose central concern is the production, management, analysis and sharing of data. In Kitchin's (2014) formulation, a data system consists of the thorough intermeshing of a technical and contextual stack that frame and compose the data assemblage. The technical stack is, in effect, the embodiment of the data architecture – the component technologies (e.g. network, hardware, operating system, database, software, interface) that comprise the instrumental means by which data are generated, processed, stored, shared, analysed and experienced. The contextual stack consists of all the discursive and material elements that shape how a data system is built, operates and is maintained over time (e.g. systems of thought, forms of knowledge, finance, governmentalities,

individual actors and communities, marketplace). In other words, all kinds of material apparatus and discursive elements are assembled together within multiple, overlapping contexts to produce and maintain a data system, defining what is 'possible, desirable and expected of data' in relation to a task or domain (Kitchin, 2014: 24).

Kitchin's formulation of a data assemblage draws inspiration from Foucault's (1980: 194) concept of the *dispositif*, which refers to a 'thoroughly heterogeneous ensemble consisting of discourses, institutions, architectural forms, regulatory decisions, laws, administrative measures, scientific statements, philosophical, moral and philanthropic propositions' that enhance and maintain the exercise of power within society. The *dispositif* of a data system produces what Foucault terms 'power/knowledge,' that is, knowledge that fulfils a strategic function:

'the apparatus is thus always inscribed in a play of power, but it is also always linked to certain coordinates of knowledge which issue from it but, to an equal degree, condition it. This is what the apparatus consists in: strategies of relations of forces supporting, and supported by, types of knowledge' (Foucault, 1980: 196).

From this perspective, data systems are the product of data politics and data power (Kennedy et al., 2015; Ruppert et al., 2017), and they reinscribe and reproduce these relations. Data systems are expressions of power/knowledge; they shape what can be done, how it is performed, and for what ends. A data assemblage, through the organisation and operation of its elements and their relations, enables data power to be claimed and exerted: data are captured, processed and analysed; choices and decisions are made; and actions are undertaken, intended to produce certain outcomes.

Kitchin and Lauriault (2014) observe that data assemblages are not discretely bounded, fixed and stable, but rather are emergent, contingent on an ever-shifting context and the mutability of action across actors and actants (non-human components). This contingency is present with respect to their day-to-day form and operation, but also their longer term development, with data assemblages constantly evolving:

'as new ideas and knowledges emerge, technologies are invented, organisations change, business models are created, the political economy alters, regulations and laws are introduced and repealed, skill sets develop, debates take place, and markets grow or shrink' (Kitchin and Lauriault, 2014: 7).

Data assemblages then are 'never entirely coherent, and [are] always being re-made' (Allen and Vollmer, 2018: 23); they are not accidental in design, but neither are they preordained. In other words, data systems designed to perform the same roles and tasks vary in their form, conditions, actors and use across organisations and context. As a result, the data power of a data assemblage is contingent,

contextual and relational; its affordances never static and determined but always emergent depending on the configuration and operation of apparatus and relations and circumstance. Yet, while data power is emergent, it is not open and unstructured as the data assemblage seeks to produce intended and desired outcomes.

While Kitchin and Lauriault's formulation of a data assemblage is rooted in the ideas of Foucault, its core ideas and logics (the assembly of material and discursive elements and processes; its dynamic, contingent, and relational nature), aligns with assemblage theory as devised by Deleuze and Guattari (1987) and extended and reworked by DeLanda (2006, 2016) and others. Nail (2017) notes that Deleuze and Guattari's vision of assemblage theory considers how a heterogeneous set of elements – that can be loosely categorised into three basic types (conditions, objects and agents, or what they call 'abstract machine', 'concrete assemblage' and 'personae') – are entwined through a set of contingent, relational arrangements to (re)produce an assemblage. As Allen and Vollmer (2018: 27) note:

'Conditions' are the abstract, governing ideas and sets of relations that connect objects and agents in meaningful ways. 'Objects' are the concrete parts that get arranged in particular ways. 'Agents' are those people who arrange the objects according to the prevailing conditions'.

All three types of constituent elements are necessary for an assemblage to be produced, sustained, and to operate, and '[e]ach one presupposes and is immanent to the other' (Nail, 2017: 28). In contrast to actor-network theory:

'where each entity is defined by its interactions with other entities – by its capacity to affect and be affected by other entities – ... assemblages consist of parts that remain autonomous and independent of their network. ... Each part is capable of detaching to form new relations and therefore assemblages are characterized by 'relations of exteriority.' ... [W]hile the parts of an assemblage remain independent and are not defined by the whole, the nature of the whole assemblage is a function of the interrelations of its parts' (Ball, 2018: 242–43).

That is, the relations between independent objects, conditions and actors collectively produce emergent properties and affordances that enable a data system to perform work in the world.

Deleuze and Guattari (1987) identify four kinds of assemblage characterised by a dominant element, which shapes how the assemblage is constituted and operates: territorial (objects), state (conditions), capitalist (actors), and nomadic (can shift between the three elements). While an element might dominate in each type of assemblage, there is always a mix of elements in operation. As we illustrate below, the data assemblages of the planning system are predominately state in form. As Nail (2017: 30) explains,

'[s]tate assemblages are arranged in such a way that the conditioning relations attempt to unify or totalise all the concrete elements and agencies in the assemblage' as a means to manage populations and social and economic activity and coordinate and deliver services. These processes of management and coordination are stratified, organised vertically across scales and horizontally across organisations, with Deleuze and Guattari (1987) identifying three main forms of stratification: binary (two way relations that might be hierarchical), circular (multiple relations that are entangled), and linear (one way relations along a dominant path, with possible deviations and side-loops) (Nail, 2017). For Deleuze and Guattari (1987), assemblages continually reproduce or transform themselves through processes of deterritorialisation and reterritorialisation as the nature of constituent elements and stratification change character, or elements are added or removed and their relations reconfigured.

Following Allen and Vollmer (2018), Dalton (2020) and Boyd (2022), we think it is productive to rework Kitchin's (2014) data assemblage, reconfiguring the technical and contextual stacks and their component parts into the three elements of conditions, objects and agents (see Figure 1). In this framing, conditions (the abstract machine) are identified as broadly relating to knowledge (systems of thought, forms of knowledge), economics (finance, political economy, marketplace), governance (governmentalities and legalities), sites (places), and human action (practices, perceptions and assumptions), with their combination contextualising and shaping the production and functioning of a data assemblage. This formulation of a data assemblage aligns with Muller (2015) and Nail's (2017) observations that assemblages are relational (dependent on connections between entities to form new entities), productive (they produce new effects in the world), heterogeneous (composed of many apparatus and elements), dynamic (constantly mutating, transforming and breaking up), desired (they are designed to solve tasks), and they constitute a multiplicity being 'neither a part nor whole' (Nail, 2017: 23) and they can be disassembled and reassembled differently (unlike a body, which is a unity, an organic whole). In this sense, an assemblage is a 'fragmentary whole', since elements can be 'added, subtracted, and recombined with one another ad infinitum without ever creating or destroying an organic unity' (Nail, 2017: 23).

It is through the assembly of objects and agents within a field of conditions that data affordances and capacities are produced that enables data power to be realised. Data systems capture required data and their algorithms enable workers to process and analyse data to perform a prescribed task and make decisions. Automated, data transfer processes, such as an ETL (extract, transform, load) function, enable data to be extracted from one data system, be reconfigured and uploaded into another, allowing other actors to leverage insight and value from them. Linking together data systems enables a highly complex set of processes and

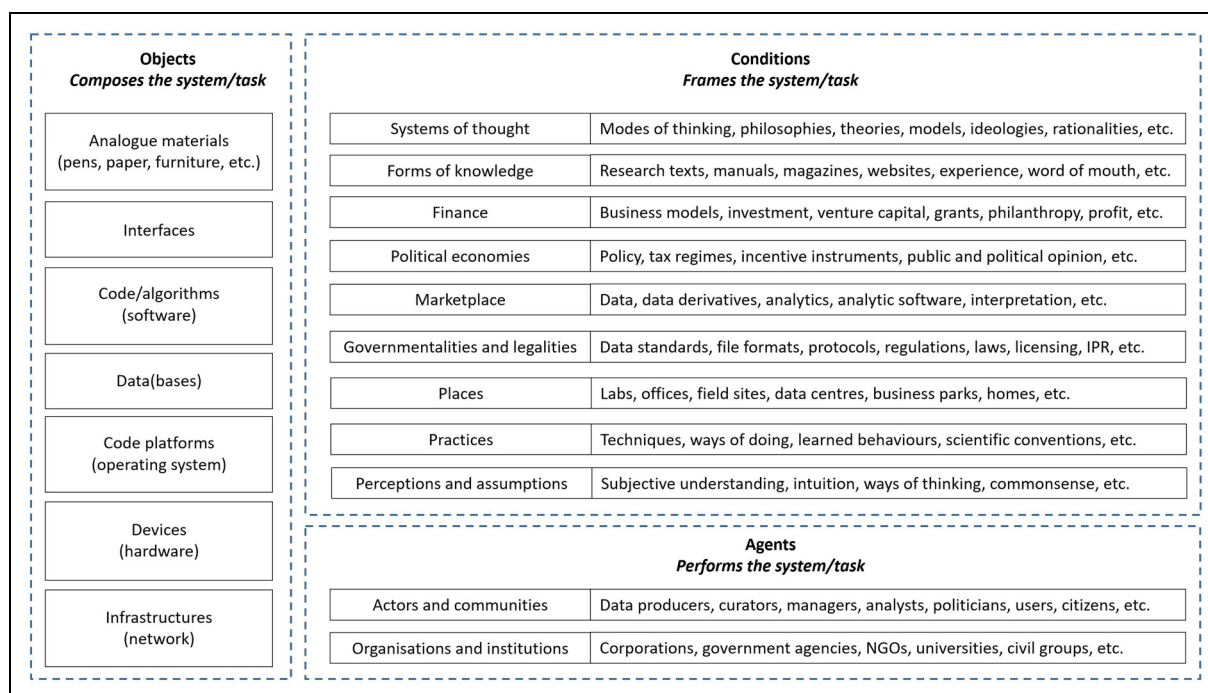


Figure 1. The conditions, objects and actors that make-up a data assemblage. Source: A re-organisation and extension of Kitchin (2014) and Kitchin and Lauriault (2014) to align with Deleuze and Guattari's assemblage theory.

actions to be coordinated across actors and institutions to solve tasks. Without certain inter-relationships between objects, conditions and agents particular desired processes and actions cannot be performed, or are performed inefficiently or ineffectively. At the same time, simply assembling a set of objects, conditions and agents does not lead automatically to a particular expression of data power: it is in their configuration and their interplay that affordances arise and data power emerges and is expressed. In other words, data power is not deterministic or causal in nature but contingent; though data assemblages are produced to enable certain kinds of work to be undertaken. As we note below, while the Irish planning data systems are meant to lead to uniform outcomes within and across the 31 local authorities (LAs), variances in their constitution and operation creates contingent decisions (the same planning application might receive a different verdict, and varying planning compliance conditions, across planning officers in the same LA, or across LAs).

For some, the relational, contextual, emergent understanding of social systems provided by assemblage theory reduces analysis to limited example-specific observations. If every action and outcome is contingent on objects, conditions and actors, then all analysis descends to relativism, with no over-arching structural or causal processes driving how social relations unfold. This is the critique levelled at urban assemblage theory by political economists, who contend that assemblage theory positions urban relations outside of the 'context of contexts' in which they are

embedded, including local historical specificity and the wider, contradictory, hierarchical and institutional forms of global capitalism (Brenner et al., 2011; Kinkaid, 2020). Assemblage theory is seen to privilege 'potential, transformation, and flux over ideas of identity, structure, and system' (Kinkaid, 2020: 480). The charge here is that by focusing on minutiae of how individual data systems and data ecosystems are assembled and work in everyday practice, the wider historical context and political-economic forces shaping their formation and operation largely disappear from view. For example, the digitalisation of the Irish planning system is taking place in a context of neoliberal reforms to the public sector, austerity measures arising from a financial crash in which planning failures were cited as a contributing factor, political lobbying for planning deregulation, and a housing crisis with chronic undersupply (Hearne, 2020; Kitchin et al., 2012). Moreover, assemblage theory can focus on describing what exists, largely failing to document modes of resistance and the visioning and production of alternatives (Dalton, 2020). Failing to contextualise the production of assemblages, critics argue, works to naturalise dominant regimes and structures of power, rather than challenging them (Kinkaid, 2020).

While such critiques have merit, they are not antithetical to assemblage theory, though they do require analysis to more adequately frame contextually the assembly of objects, conditions and actors, and the work that assemblages do in reproducing political-economic and social relations, while not a priori assuming the nature of such relations,

or ignoring attempts to resist and recast assemblages (Anderson et al., 2012; Dalton, 2020; Kinkaïd, 2020). Moreover, there is a need to consider the ways in which data assemblages are scaled into data ecosystems, to which we now turn our attention.

Data systems assembled as a data ecosystem

While the concept of data assemblages has utility for understanding the development, constitution and operation of data systems, it also potentially places a boundary/scalar limit on this endeavour. Indeed, data systems are often scaffolded together in order to ensure that a suite of tasks within a domain can be performed. In other words, data systems are themselves assembled together to form data ecosystems bound together by data mobilities (Bates et al., 2016; Kitchin et al., 2025a) and socio-technical arrangements (contracts, governance arrangements, regulations, standards). In DeLanda's (2006) terms, a data ecosystem is itself an assemblage, which in turn can be part of another data assemblage composed of multiple data ecosystems: society in his terms is composed of a multitude of interconnecting and nested assemblages that span scales in a flat ontology from local to global. Or to put it another way, the world is composed of 'assemblages all the way down', with '[e]ach component of an assemblage ... also an assemblage' (Ball, 2018). Boyd (2022: 1348) thus conceptualises data themselves as assemblage – 'a mutable, portable, sociotechnical arrangement of material and symbolic components that serves as evidence': a product of a data assemblage (a socio-technical infrastructure).

Much of the literature on data ecosystems (e.g. Oliveira and Loscio, 2018; Gelhaar et al., 2021) describe them on the same terms as Tarantino's (2019) notion of a datascape; that is, composed of several data systems, spread across a number of actors, that interact with each other in order to exchange, produce and consume data around a common endeavour. The constitution of the data ecosystem and how it operates is defined by a set of technical, social and governance/legal arrangements and the interdependencies between actors and data systems (Scassa, 2019). There is little notion of the biological roots of the ecosystem metaphor evident in this framing. However, others have used the metaphor more literally, casting the notion of an assemblage as ecologies of interdependent relations. For example, van Schalkwyk et al. (2016) draw directly on the language and ideas of ecosystems theory to discuss how data ecosystems involve co-determinate and symbiotic relationships between mutually interacting organisms (e.g. firms, institutions, customers, etc.), including 'keystone species' such as data intermediaries that create enabling conditions for successful collaborations by providing services that add value (e.g. research and training consultancies). Organisms are complexly arranged, with movements between them

cyclical and reinforcing, and interdependencies existing between organisms and resources, which together enables adaptation and creates resilience (van Schalkwyk et al., 2016). A literal application of the ecosystem metaphor might be appealing, but when applied in practice it soon becomes strained. Nonetheless, the ecosystem metaphor is useful for designating a set of interlinked data systems that collectively perform the data work that enables a domain, such as a planning system, to operate, but in our view is best conceived as an 'assemblage of assemblages' (DeLanda, 2016).

A core aspect of the functioning of a data ecosystem is the sharing and mobility of data within and between data systems, binding them together through co-dependent interconnections (Bates et al., 2016). In a separate paper, we provide a detailed analysis of the data mobilities of the Irish planning system revealing the complex set of data flows and exchanges that occur to enable the assembly of all the data required to make an informed planning decision (Kitchin et al., 2025a). Such data mobility is facilitated by seams (links and interfaces between systems) and shared metadata, standards, protocols (Inman and Ribes, 2018) and hindered by frictions and vulnerabilities (e.g. access controls; incompatible data formats and standards; mistakes and glitches; resistance by actors; costs and skills capacities; and regulatory and legal limitations) (Edwards, 2010; Bates, 2018). Without these data mobilities the tasks being performed within a data ecosystem cannot be realised as downstream processes are reliant on upstream outcomes and input data (Kitchin et al., 2025a). Data frictions that limit data mobility thus produce fragmented data ecosystems (Kitchin and Moore-Cherry, 2021). Data ecosystems are sustained through shared practices of maintenance and repair, and they adapt to prevalent conditions and evolve through the introduction of new actors, innovations and relationships. They are supported by legal arrangements (e.g. contracts, licensing), institutional imperatives (e.g. strategic partnerships), technical protocols and standards, and underpinned by shared values. At the same time, data ecosystems are full of endogenous institutional politics as parties seek to influence and control resources, capacities, management responsibilities and decision-making, and are open to exogenous forces (e.g. political lobbying, regulatory capture) that can render relations precarious or conflictual requiring mediation and active management, or threaten to destroy the ecosystem (van Schalkwyk et al., 2016).

In the following sections, we illustrate our extension to the notions of data assemblage and data ecosystem with respect to the Irish planning system based on empirical research that charts the various data systems and data infrastructures used by planning authorities and related actors to manage and fulfil their statutory functions and perform their various roles and responsibilities, documents the data generated and how they are used, and maps the relationships and data mobilities within and between each

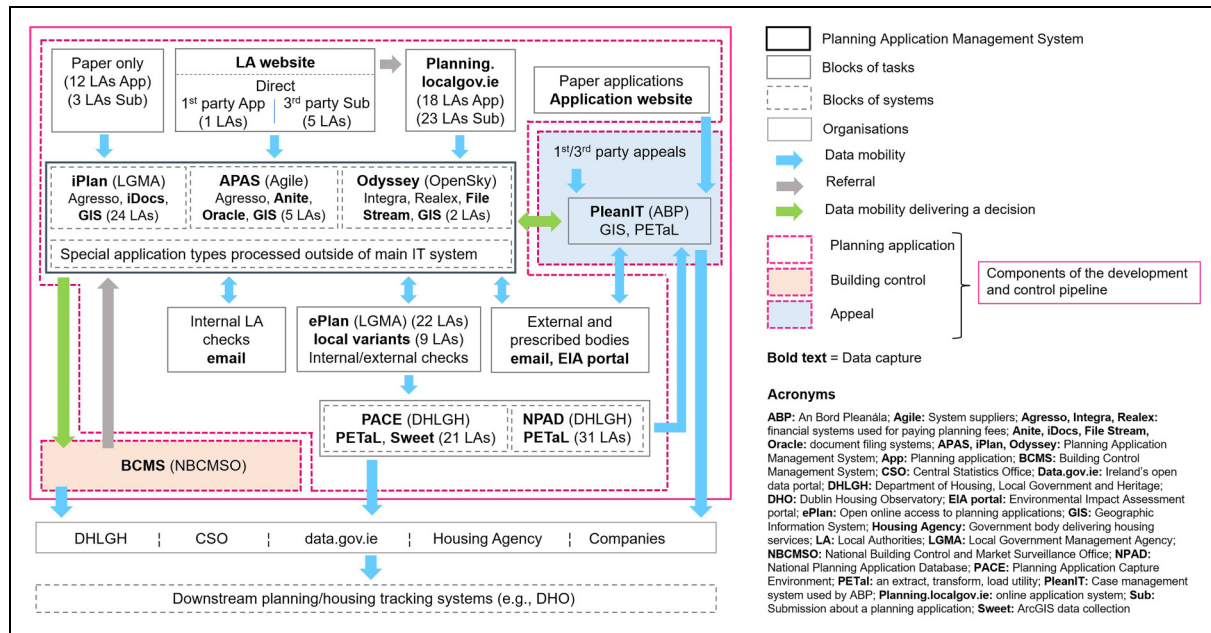


Figure 2. The data assemblages and data ecosystem for terrestrial planning development and control in Ireland, August 2023.

system and infrastructure across multi-tiered system of governance. The fieldwork was undertaken during the summer of 2023 using a multi-method approach. Twenty-nine public sector officials working in LAs, state agencies and government departments were interviewed about the data systems used in the development and control aspects of the planning system. A number of the interviews were of a walk-through nature, with the interviewee demonstrating and explaining the workflow related to the use of a data system. In addition, we examined the user manuals for the data systems and performed data audits on five data systems (iPlan, APAS, Odyssey, BCMS, and planning.localgov.ie), and identified the flow of data from the planning data systems into open data sites and interactive graphing and mapping utilities (e.g. Dublin Housing Observatory). This suite of methods enabled an understanding of the data architecture of each system, as well as how they were interconnected to form a wider data ecosystem.

The formation and on-going emergence of a data assemblage

Since the early 2000s the planning and control function of the Irish planning system, a role undertaken by the LAs, has been undergoing a process of digitalisation as the tasks and processes of planning assessment has increasingly become digitally mediated (see Kitchin et al., 2025b). In Dublin, the four LAs adopted the use of APAS, a commercial planning application management system, in 2000 to manage and process applications for planning permission. Other LAs in Ireland adopted iPlan from 2002 onwards, a management system developed by the Local Government

Computer Services Board (LGCSB) for the same purpose. Planning applications were submitted to a LA as paper documents and staff would then scan them and enter key data into APAS or iPlan, which were used to process applications, store data related to the application and third-party submissions about them, and record decisions. APAS and iPlan did not simply provide a digital version of a paper-based system, providing a more efficient means of storing and processing data, but also provided new tools, capacities and capabilities, such as shared data dictionaries, the inter-linking of data, the tracking of administrative work, and the coordination of data sharing. However, while these systems could be used to manage most application types, in a number of specialised cases applications had their own procedures due to limitations in data system functionality. These specialised cases were typically handled using spreadsheets and/or GIS. By 2004, ePlan (developed by LGCSB), and an equivalent system for LAs using APAS, provided an online means of viewing planning application documents, though submissions about proposed developments had to be submitted on paper, which were then scanned and logged. Subsequently, several other data systems have been adopted to complement APAS, iPlan, and e-planning websites, to manage, assess and monitor the development pipeline from application to construction to turn-key property to form a relatively complex planning and control data ecosystem (see Figure 2).

Each of these data systems comprise a data assemblage that consists of a unique collection of objects, agents and conditions. Each is composed of a constellation of infrastructure, hardware, databases, software, interfaces and other elements, assembled into a data architecture designed

to undertake designated tasks. Each has been assembled by a set of actors (software engineers, IT specialists, data stewards, policy makers, consultants, domain specialists) and stakeholders (LAs, government departments, state agencies, IT companies). This assembling has taken place within a context (conditions) that has shaped intention, design thinking and implementation, including forms of knowledge, established ways of undertaking work, political pressure, policy directives, financing, regulations, law, contracts, and so on, with these being negotiated within and across actors. A core conditioning factor is the relationship between data systems and on-going processes of digitalisation; for example, the extent to which a data system complements or replaces a paper-based system, or how it relates to, and is expected to interconnect with, other data systems. In the case of data systems such as iPlan, used by 24 LAs, while the data architecture is the same across LAs, workflows and work practices and cultures relating to its use can vary locally (it has its own constellations of local agents and conditions). With respect to APAS, the data architecture can also vary and in the Irish case the five APAS instances are quite different in their form.

As already noted, data systems do not come into the world ready-made (Bowker and Star, 1999), but ‘emerge through an incremental process of enacting, extending, standardising and embedding technical and social practices in specific contexts for unique needs’ (Aula, 2019: 2). The notions of articulation and scaffolding provide a useful means for describing how a data system is made-up. Articulation refers to the identification of all the tasks necessary to complete a job, breaking these down into a sequence of steps, and working out how to schedule, coordinate, and monitor them to ensure successful delivery (Kaltenbrunner, 2015; Tanweer et al., 2016). Articulation work includes creating a workplan, establishing which actors should be involved, and determining the resources needed to produce and run a data system (Nadim, 2016). Various stages in the development of a data system might be articulated in tandem or separately, progressing in an iterative sequence. A key component of articulation is working through a course of action to reach completion despite problems encountered, such as glitches and unanticipated events. Once a data system has been articulated, the means to achieve its realisation has to be scaffolded into place. Such scaffolding includes assembling all the necessary resources (e.g. people, finance, technology), putting in place institutional and governance arrangements, and creating the political and legislative conditions to encourage enactment (Halfmann, 2020). Over time, new rounds of articulation and scaffolding are employed as data systems and their operations are reviewed and revisions and upgrades are devised.

How objects, actors and conditions interact to articulate and scaffold a data assemblage is evident in the case of the BCMS, first introduced in 2014. The aim of the BCMS was to provide a nationwide data system for tracking the

pipeline of construction from commencement to completion, and to monitor fulfilment of building control requirements. Scaffolding together a new data system aimed to provide planning authorities with the data power to assess compliance and intervene in the development process if required, and enhance the ability to produce national-level construction statistics and data analytics. The prompt for developing the BCMS was the introduction of new regulations (S.I.9 of the Building Control Amendment Regulations 2014) that required the notice of construction commencement and completion, and the certification of compliance with building regulations, to be recorded on a new National Statutory Building Control Register (Dwyer-Bond et al., 2019). The new regulations were deemed necessary due to the failings of building control during the Celtic Tiger property boom (1994–2007), which led to thousands of buildings with poor build quality, pyrite contamination, and inadequate fire safety measures (Ahern et al., 2018).

Prior to 2014, building control was monitored separately by the 34 building control authorities (which were the LAs; since reduced to 31 due to local government reform). Each LA had its own system that differed in its forms and processes (e.g. paper and/or spreadsheet based), monitoring was a time and resource intensive task, and it was troublesome to try to produce harmonised data and a national picture of construction activity (Dwyer-Bond et al., 2019). In addition, it was possible for unqualified staff to certify compliance. BCMS would digitalise the monitoring of building control and be the new statutory national register, providing a standardised process and harmonised data across all building control authorities that could only be uploaded by accredited professionals. A new organisation, the National Building Control and Market Surveillance Office (NBCMSO), hosted by Dublin City Council, was established to deliver the service (O’Dowd, 2016). The creation of the BCMS involved a number of partners, including the Local Government Management Agency (LGMA), the Department of the Environment, Community and Local Government (DECLG), An Bord Pleanála (ABP, national planning appeals agency), and the 31 LAs.

On initiation a relatively complex project governance structure was put in place to help guide the articulation and scaffolding of the system, including oversight (strategic direction) and steering (project management) groups (see Figure 3). Such a governance arrangement illustrates the tangle of institutional structures required to create a national-level data system that is to be used by multiple organisations and hints at the institutional politics at work within these committees and groups that can delay, stymie and redirect new data systems. As such, the articulation work undertaken within this governance structure was contested and negotiated, with the data system design shifting through various iterations in order to try to gain consensus on the vision, functions and data architecture of the BCMS. In addition, the views of stakeholders had to be taken into

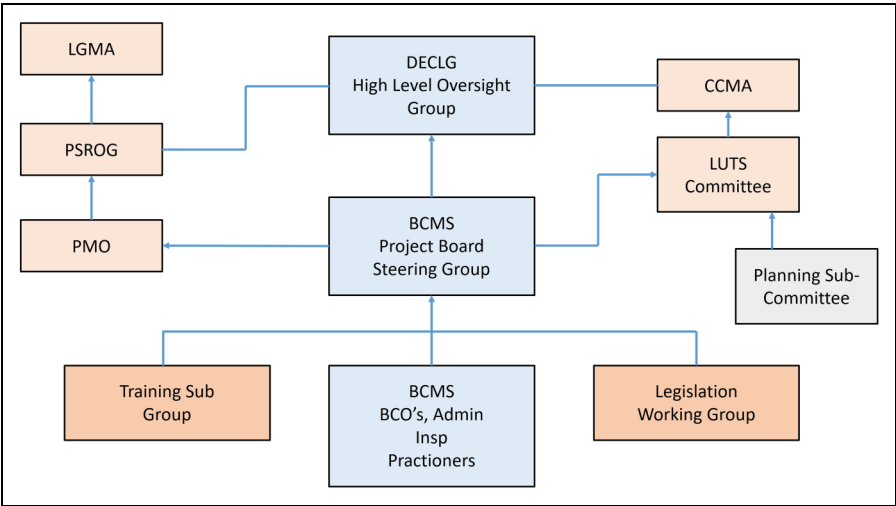


Figure 3. BCMS project governance. Source: Redrawn from O’Dowd (2016). CCMA : City and County Management Association (a non-statutory body composed of the chief executives of the local authorities); DECLG : Department of Environment, Community and Local Government; PSROG : Public Sector Reform Oversight Group (a sub-committee of the LGMA board composed of senior representatives of the CCMA, LGMA, DECLG and the private sector); PMO : Local Government Programme Management Office; LUTS : Land Use and Transportation.

Table 1. Timeline of projected project development for 2016.

December 2015:	2nd Quarter 2016
<ul style="list-style-type: none">• Tender Review• Review and Update of FAQ’s	<ul style="list-style-type: none">• Re-design of Register• Project Assessment• 7-Day Notice (7 DN) functionality ‘switched on’• Inspection/Enforcement Module• Data Analytics• Business Plan Review
January 2016:	3rd Quarter 2016
<ul style="list-style-type: none">• Successful tenderers informed• Campaign reminder email issued• Review of Framework for Building Control Authorities drafted• Results of Resources Survey drafted	<ul style="list-style-type: none">• Fire Safety Certificate Module• Disability Access Certificate Module• CN/7 DN Revised Information function
February 2016	4th Quarter 2016
<ul style="list-style-type: none">• CCC Module goes live	<ul style="list-style-type: none">• Links to Professional Registers• Companies Module• Dispensation/Relaxation Module• Document Management solution• System Architecture review and software updates
Launched 19th February 2016	
<ul style="list-style-type: none">• New front end interface• User Guides for front and back end• Commencement & Completion process fully online• Opt Out Module	

account, and there was active consultation with industry to try to ensure the proposed design was efficient and cost effective for those uploading data and documents (Dwyer-Bond et al., 2019).

This articulation work ultimately lead to project and infrastructure master plans (see O’Dowd, 2016). These master plans provided the roadmap for development, with project staff then working to scaffold these into operation by assembling relevant resources within conditioning constraints

(e.g. available finance, successful tendering, the tendered company having sufficient skilled staff, knowledge, etc.). The tendered company was expected to work to a project timetable and milestones (see Table 1). This process of articulation and scaffolding continued for a couple of years, with an initial launch followed by continued development and rollout of additional functionality. It also had to deal with changing conditions, such as the intervention of the then Minister for DECLG, Alan Kelly TD, who

subsequently altered the system specification by providing an exemption option for recording data in BCMS for those building one-off houses and housing extensions to avoid a 'cost burden' (Kelly, 2015).

In subsequent years, the BCMS has been tweaked through processes of maintenance and repair, upgraded to new software and migrated onto new hardware, and extended through additional development work in response to user feedback and changes in regulations and legislation. Its design and operation is presently under review to address data quality issues and improve the ability to track developments across data systems. In other words, BCMS has constantly undergone a process of being reterritorialised; it is an emergent socio-technical system, shaped by shifting prevalent conditions, the work of diverse actors, and new technical innovations. Its data power has thus also been reterritorialised, with new features and functionality providing building control officers with enhanced knowledge that can be used more effectively track and enforce compliance. Such data power is not universal, but arises in the context of specific planning applications, how BCMS is used in practice by planners, and how its utility might be subverted by weak data quality measures (e.g. building control data is self-reported and open text fields mean that it is easy to upload false data). An important aspect of the BCMS assemblage and its emergence is its relationship to, and interactions with, other data assemblages and its position and role in the wider planning development and control data ecosystem.

The data ecosystem of the Irish planning system

The delivery of the development and control function of the Irish planning system has, since its inception in 1963, consisted of an amalgam of inter-related and inter-connected data systems that form a data ecosystem. Data was initially captured in paper form, requests were sent to different internal and external units for feedback, the public's opinions were sought, decisions were made and shared. Data were stored in ledgers, folders, and filing cabinets of different officials and were passed between teams and institutions. The processes of digitalisation and evolution of digital technologies, the desire for institutional and regulatory control, and the ability to formulate and track evidence-informed policy, have radically increased the volume of data captured, the degree of interconnection between data assemblages and the scope of data sharing, and by extension the complexity and scale of the data ecosystem. This in turn has increased the data affordances and capacities available to planners and reshaped the nature of their work, and that of other actors (e.g. developers, policy makers, academics) and the public through e-planning and open data.

Figure 2 charts the development and control data ecosystem as of August 2023, its various data assemblages, the actors primarily responsible for their operation and

management, and their interconnections and data mobilities. The chart is loosely organised sequentially from top-to-bottom, with the exception of the appeal process, with different data assemblages being used to manage different stages of the development and control process from application, through assessment, appeal, construction and building control, and the production of open data. As the various arrows indicate, each system is dependent on data mobilities and data being shared with other data systems in order to perform key tasks and collectively deliver the development and control functions as statutorily required by legislation (Kitchin et al., 2025a). Interestingly, the data ecosystem operates without any one entity being in charge, or its actors necessarily knowing how it all fits together and operates. In fact, most of the actors we spoke with did not know the full extent of their own organisation's data work and certainly had little knowledge of the data ecosystem as a whole.

At the top of the chart is the process of applying for planning permission. In August 2023, the digitalisation process remained partial. Twelve LAs still required paper applications, and three only accepted third-party submissions as a written letter. On receipt of paper applications they are scanned by LA staff and added to a digital document store, and details are manually typed into a planning application management system. Likewise, ABP only accepted paper applications, though for large developments it requested the applicant to set up a website that hosted all relevant documents in digital form. For legal reasons, ABP printed out any digital material sent to them, such as emails, and added them to their paper case files. Eighteen LAs had adopted planning.localgov.ie as a portal for digital submission of applications, and one LA used their own portal. Data submitted via these portals are imported directly into the planning application management system. While the planning application management systems are used to process the vast majority of applications, a number of specialist applications – such as those made under Sections 5, 35, 42, 44, 57, 247 and Parts V, VII and XI of the Planning Act – are processed separately (particularly for LAs using APAS) due to system limitations.

At the application stage, a check is made with payment systems (Agresso for iPlan and APAS; Realex or Integra for Odyssey) regarding whether the planning fee had been paid. After initial checking of submitted materials, selected details of the application are passed into an ePlan system, which allows the public to inspect the proposed development. Feedback is sought on every application from internal LA units (e.g. transportation, environment, and archaeology and heritage departments) and selected external bodies (e.g. Office of the Planning Regulator) and prescribed bodies (e.g. Department of Housing, Local Government and Heritage, DHLGH; Transport Infrastructure Ireland; Health Service Executive) via email, ePlan or EIA (Environmental Impact Assessment) portal. PETaL, an automated ETL

process, is used to harvest up to 25 data fields from each LA's e-planning site and import them into NPAD (National Planning Applications Database) and PACE (Planning Application Capture Environment). NPAD provides an open, online, interactive mapping tool to view all planning applications made to each of the 31 LAs since 2012 and is entirely dependent on source data from ePlan sites to function. If a planning application is unsuccessful then an appeal can be made to ABP who will request the transfer of associated data and documents from the relevant LA and, once a decision has been made, pass back the outcome and its conditions. Once permission is given for a development to proceed, the building control phase is initiated by notifying the BCMS of commencement.

Selected data are also shared with third parties for the production of official statistics and use as open data. In the case of the Central Statistics Office (CSO), each LA and ABP submit, on a monthly basis, 14 data fields relating to each granted planning permission in order to comply with the statutory provision under the Short Term Statistics Regulation (EC) Number 1882/2003 to supply Eurostat with data necessary to compile variables 411 and 412 of Annex B (Construction). Other selected data are shared with DHLGH (Department of Housing, Local Government and Heritage) and the Housing Agency, some of which is made openly available on their data hubs and via data.gov.ie (the national open data site). These data are imported into a number of planning and housing tracking systems that provide open dashboard visualisations and interactive maps for monitoring key performance indicators. These include:

- Department of Housing, Local Government and Heritage Housing Delivery Tracker (<https://story-maps.arcgis.com/stories/ab12ed6d50a540e2891170c24955ff49>)
- Housing for All dashboard (https://public.tableau.com/app/profile/statistics.unit.housing/viz/HousingforAll/0_Overview)
- Office of the Planning Regulator Digital Planning Hub (<https://opr-hub-oprgis.hub.arcgis.com/>)
- Dublin Housing Observatory (<https://airomaps.geohive.ie/dho/>)
- Regional Development Monitor (<https://rdm.geohive.ie/>)
- Dublin Housing Task Force mapper (<https://housinggovie.maps.arcgis.com/apps/View/index.html?appid=3fa56a71ee774f9487d14a9e5336b00c>)

In addition, data held in ePlan and BCMS are scraped by private companies on a daily (Construction Information Services Ireland – CIS) and weekly (Building Information Ireland – BII) basis and converted into commercial data products. CIS and BII clean and wrangle the data into more useable forms, validate the data and link it to other datasets (such as procurement data scraped from e-tender portals),

and produce analysis tools that enable site development to be tracked from permission to completion. Through the provision of open data and commercial services the reach and utility of the data ecosystem is massively extended.

While this data ecosystem functions largely as intended, in that it enables informed decisions to be made regarding planning applications and for building control to be monitored, it is not optimal or fully comprehensive in its scope or functionality and several data frictions exist that hinder the interconnections between data assemblages (Kitchin et al., 2025b). The continued use of paper and manual data re-entry is inefficient and weakens data quality through mistyping and miscodings. It is difficult to track a proposal along the development pipeline as it is assigned a new ID at each stage in the process: at pre-planning consultation, planning application, the appeals process, and building control. Additionally, there is no standardised ID reference numbering system across LAs. As noted, a number of specialist applications cannot be processed by planning application management systems, making the compiling of reports and summary statistics time-consuming. Moreover, each data system functions in a different way, using different data standards, reducing data compatibility and interoperability and limiting the ability to produce harmonious national-scale datasets and official statistics. For example, open text fields are frequently used in iPlan and Odyssey, and Odyssey and APAS make extensive use of check boxes and drop-down menus. There are 65 mandatory fields in iPlan, but only 40 in Odyssey and 21 in APAS. These data frictions and variable data architectures produce variances within and across LAs and other government agencies in how the planning system functions.

Conclusion

The work of the state has been transformed by digitalisation and datafication. While paper remains an important media in state bureaucracy, it is rapidly being complemented and replaced by digital data systems. While much research has been undertaken on the effects of data systems on e-government and e-governance, relatively little critical attention has been paid to the ontological nature of the data systems themselves, how they come into being, operate, iterate, and work in conjunction with other systems. In this article, we have provided a socio-technical conceptualisation of the organisation and operation of data-driven state work by recasting Kitchin (2014) and Kitchin and Lauriault's (2014) notion of a data assemblage through an engagement with Nail's (2017) reading of Deleuze and Guattari's (1987) assemblage theory, and extending its framing and logics to the notion of a data ecosystem, which we conceptualise as an assemblage of data assemblages. In so doing, we have forwarded two, inter-related contentions, illustrating our argument with respect to the empirical case study – the assessment of planning

applications and appeals, the monitoring of building control and the production of open data in Ireland.

First, we have made the case that data systems are thoroughly socio-technical in nature, the product of objects, conditions and agents. Data systems are assembled and continually reterritorialised, emerging contingently, contextually and relationally through material and discursive practices and processes, the product of data politics and data power. The assembly of objects, conditions and agents in turn produces affordances and capacities that enables forms of data power and diverse, context-specific work to be performed in the world. We used the example of the BCMS to illustrate how a data system is assembled in practice through processes of articulation and scaffolding, and is maintained over time through on-going discursive and material processes that are enacted individually and institutionally.

Second, data systems are rarely constituted and work alone, but are inter-connected and inter-dependent with other data assemblages, meshed together into emergent data ecosystems. Data systems are bound together through shared goals, institutional arrangements, legislative mandates, infrastructure, and data mobilities. Data systems exchange data and work in partnership, with the nature of the connections between them evolving over time. In the case of development and control in the Irish planning system, the BCMS is one of several data systems that have been assembled into a functioning, emergent data ecosystem that incorporates a relatively large number of organisation across scales from the local to the national. As such, while it is feasible and legitimate to examine a single data system to understand its constitution and operation, this scale of analysis can obscure the position, relations and role of a data system in a wider data ecosystem and its co-dependencies.


We believe that our use of assemblage theory to make sense of, and frame research about, data assemblages and data ecosystems has wide utility for charting the processes of digitalisation, datafication and data-driven operations within domains. In our view, more research is required to further develop and extend the tenets and test the robustness of our conceptualisation of data assemblage and data ecosystem. This includes exploration and application with respect to other forms of assemblage (territorial, capitalist and nomadic) and other domains (e.g. health, transport, education, etc.), paying close attention to the specific processes of articulation and scaffolding at work and the political-economic context that frames their unfolding, while being mindful of resistance and attempts to construct alternative assemblages.


Acknowledgements


The research reported in this article was funded by the European Research Council (no. 101052998) and the Local Government Management Agency. The authors are grateful to Dr Bernie O'Donoghue Hynes, Tom Elders, Tara Kerrins and Holly Morrin at LGMA for their advice, all the interviewees for spending


time to explain and demonstrate their work and the systems they use. Thanks to Mairéad Phelan and Ronan Glynn of the National Building Control and Market Surveillance Office for insights into BCMS. Thanks to Danielle Hynes and Oliver Dawkins for feedback on an initial draft.

ORCID iDs

Rob Kitchin  <https://orcid.org/0000-0003-4458-7299>

Juliette Davret  <https://orcid.org/0009-0005-7778-6595>

Carla M Kayanan  <https://orcid.org/0000-0002-4359-3534>

Samuel Mutter  <https://orcid.org/0000-0002-9246-7147>

Funding

The authors disclosed receipt of the following financial support for the research, authorship and/or publication of this article: The European Research Council (grant number 101052998), Local Government Management Agency, Ireland.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

References

- Ahern E, Jimenez-Bescos C and Desai A (2018) Evaluating how Ireland has improved building regulations compliance and energy efficiency. In: Elsharkawy H, Zahiri S and Clough J (eds) *Proceedings of International Conference for Sustainable Design of the Built Environment (SDBE) 2018*, London: University of East London, 421–432. https://aru.figshare.com/articles/conference_contribution/Evaluating_how_Ireland_has_improved_Building_Regulations_Compliance_and_Energy_Efficiency/23777493/1.
- Allen WL and Vollmer BA (2018) Clean skins: Making the e-border security assemblage. *Environment and Planning D: Society and Space* 36(1): 23–39.
- Anderson B, Kearnes M, McFarlane C, et al. (2012) On assemblages and geography. *Dialogues in Human Geography* 2(2): 171–189.
- Aula V (2019) Institutions, infrastructures, and data friction – Reforming secondary use of health data in Finland. *Big Data & Society* 6(2): 1–13.
- Ball A (2018) Manuel DeLanda, assemblage theory. *Parrhesia* 29: 241–247.
- Bates J (2018) The politics of data friction. *Journal of Documentation* 74(2): 412–429.
- Bates J, Lin Y-W and Goodale P (2016) Data journeys: Capturing the socio-material constitution of data objects and flows. *Big Data & Society* 4(2): 1–12.
- Bowker G and Star L (1999) *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.
- Boyd C (2022) Data as assemblage. *Journal of Documentation* 78(6): 1338–1352.
- Brenner N, Madden DJ and Wachsmuth D (2011) Assemblage urbanism and the challenges of critical urban theory. *City* 15(2): 225–240.

- Dalton CM (2020) Rhizomatic data assemblages: Mapping new possibilities for urban housing data. *Urban Geography* 41(8): 1090–1108.
- Datta A (2023) The digitalising state: Governing digitalisation-as-urbanisation in the global south. *Progress in Human Geography* 47(1): 141–159.
- DeLanda M (2006) *A New Philosophy of Society: Assemblage Theory and Social Complexity*. London: Continuum.
- DeLanda M (2016) *Assemblage Theory*. Edinburgh: Edinburgh University Press.
- Deleuze G and Guattari F (1987) *A Thousand Plateaus: Capitalism and Schizophrenia*. Trans Massumi, B. Minneapolis: University of Minnesota Press.
- Dunleavy P, Margetts H, Bastow S, et al. (2006) *Digital Era Governance: IT Corporations, the State, and E-Government*. Oxford: Oxford University Press.
- Dwyer-Bond E, Casey B and Coates D (2019) Developing data sources for a commercial property statistical system (CPSS) in Ireland. In: *International Conference on Real Estate Statistics*. Luxembourg, February 20–22, 2019. https://www.cso.ie/en/media/csoie/methods/methodologicalresearch/commercialproperty/ICREA_Ireland_2019.pdf
- Edwards PN (2010) *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. Cambridge, MA: The MIT Press.
- Falk S, Rommele A and Silverman M (eds). (2017) *Digital Government*. Berlin: Springer.
- Foucault M (1980) The confession of the flesh. In: Gordon C (ed) *Power/Knowledge*. New York: Pantheon Books, 194–228.
- Gelhaar J, Groß T and Otto B (2021) A taxonomy of data ecosystems. In: Bui TX (ed) *Proceedings of the 54th Hawaii International Conference on System Sciences*. Honolulu: HICSS, 6113–6122.
- Halfmann G (2020) Material origins of a data journey in Ocean Science: How sampling and scaffolding shape data practices. In: Leonelli S and Tempini N (eds) *Data Journeys in the Sciences*. Cham, Switzerland: Springer Open, 27–44.
- Hearne R (2020) *Housing Shock: The Irish Housing Crisis and How to Solve It*. Bristol: Policy Press.
- Inman S and Ribes D (2018) Data streams, data seams: Toward a seamful representation of data interoperability. In: *Paper presented at Design Research Society*. University of Limerick, September 25–28, 2018. <https://dl.designresearchsociety.org/drs-conference-papers/drs2018/researchpapers/16/>
- Kaltenbrunner W (2015) Infrastructural inversion as a generative resource in digital scholarship. *Science as Culture* 24(1): 1–23.
- Kelly O (2015, August 31) Builders of one-off homes can opt out of regulations from Tuesday. *Irish Times*. <https://www.irishtimes.com/news/environment/builders-of-one-off-homes-can-opt-out-of-regulations-from-tuesday-1.2334297>
- Kennedy H, Poell T and van Dijck J (2015) Data and agency. *Big Data & Society* 2(2): 1–7.
- Kinkaid E (2020) Assemblage as ethos: Conceptual genealogies and political problems. *Area* 52(3): 480–487.
- Kitchin R (2014) *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. London: Sage.
- Kitchin R (2025) *Critical Data Studies: An A to Z Guide to Concepts and Methods*. Cambridge: Polity Press.
- Kitchin R, Davret J, Kayanan CM, et al. (2025a, online first) Data mobilities: Rethinking the movement and circulation of data. *Mobilities*. DOI: 10.1080/17450101.2025.2481309.
- Kitchin R, Davret J, Kayanan CM, et al. (2025b, online first) Genealogy of a data ecosystem: The digitalisation of the Ireland's planning development and control, 2000–2024. *Digital Geography & Society*. DOI: 10.1016/j.diggeo.2025.100124.
- Kitchin R and Lauriault T (2014) Towards critical data studies: Charting and unpacking data assemblages and their work. *The Programmable City Working Paper 2*: 1–19. <https://mural.maynoothuniversity.ie/id/eprint/5683/>.
- Kitchin R and Moore-Cherry N (2021) Fragmented governance, the urban data ecosystem and smart cities: The case of Metropolitan Boston. *Regional Studies* 55(12): 1913–1923.
- Kitchin R, O'Callaghan C, Boyle M, et al. (2012) Placing neo-liberalism: The rise and fall of Ireland's Celtic Tiger. *Environment and Planning A* 44(6): 1302–1326.
- Mayer-Schönberger V and Cukier K (2013) *Big Data: A Revolution That Will Change How We Live, Work and Think*. London: John Murray.
- Muller M (2015) Assemblages and actor-networks: Rethinking socio-material power, politics and space. *Geography Compass* 9(1): 27–41.
- Nadim T (2016) Data labours: How the sequence databases GenBank and EMBL-bank make data. *Science as Culture* 25(4): 496–519.
- Nail T (2017) What is an assemblage? *SubStance* 46(1): 21–37.
- O'Dowd E (2016) Building control management system: Supporting building control compliance. In *Paper presented at IBCI Conference*, Sligo. https://www.ibci.ie/docs/conferences/2016/2_Eoin_ODowd_BCMS_Supporting_Building_Control_Compliance.pdf
- Oliveira MLS and Loscio BF (2018) What is a data ecosystem? In: Zuiderwijk A and Hinnant CC (eds) *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*. New York: ACM, Article 74.1–9.
- Ruppert E, Isin E and Bigo D (2017) Data politics. *Big Data & Society* 5(2): 1–7.
- Scassa T (2019) Ownership and control over publicly accessible platform data. *Online Information Review* 43(6): 986–1002.
- Tanweer A, Fiore-Gartland B and Aragon C (2016) Impediment to insight to innovation: Understanding data assemblages through the breakdown–repair process. *Information, Communication & Society* 19(6): 736–752.
- Tarantino M (2019) Navigating a datascape: Challenges in automating environmental data disclosure in China. *Journal of Environmental Planning and Management* 61(3): 67–86.
- van Schalkwyk F, Willmers M and McNaughton M (2016) Viscous open data: The roles of intermediaries in an open data ecosystem. *Information Technology for Development* 22(S1): 68–83.