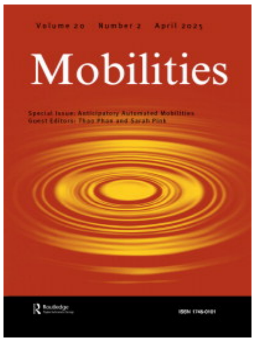

Mobilities

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rmob20

Data mobilities: rethinking the movement and circulation of digital data

Rob Kitchin, Juliette Davret, Carla Maria Kayanan & Samuel Mutter

To cite this article: Rob Kitchin, Juliette Davret, Carla Maria Kayanan & Samuel Mutter (09 Apr 2025): Data mobilities: rethinking the movement and circulation of digital data, Mobilities, DOI: 10.1080/17450101.2025.2481309

To link to this article: https://doi.org/10.1080/17450101.2025.2481309

Published online: 09 Apr 2025.

MOBILITIES
https://doi.org/10.1080/17450101.2025.2481309

Data mobilities: rethinking the movement and circulation of digital data

Rob Kitchin, Juliette Davret, Carla Maria Kayanan and Samuel Mutter

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

ABSTRACT

The mobility of data has been variously described as data: flows, streams, journeys, threads, transfers, exchanges, and circulation. In each case, data mobility is conceived as a movement from here to there; that data moves along a chain of receivers and senders. However, we contend that the metaphors of data flows (or journeys, threads, etc.) does not reflect well the processes by which digital data are shared. Rather, we propose moving from a metaphorical conceptualisation to a description of the actual mechanisms of mobility. Through a case study of the planning data ecosystem in Ireland, we detail how data replicate (replica copies produced), with the original source retaining the data and a new source gaining it, and data proliferate (multiply) across systems and sites when made available. As data replicate and proliferate, they are transformed through processes of data cleaning, data wrangling, and data fusion, producing new incarnations of the source data. Importantly, this rethinking of data mobility makes clear how and why various data incarnations are produced and, in so doing, create fundamental issues regarding the integrity of data sharing and data-driven work, the repeatability, replicability and reproducibility of science, and data sovereignty and the control of data use.

ARTICLE HISTORY
Received 1 July 2024
Revised 20 January 2025
Accepted 28 February 2025

KEYWORDS
Data mobilities; data journeys; replication; proliferation; data frictions; data seams

Introduction

Data are 'representative measures of phenomena captured through some form of measurement or observation, or derived or inferred values produced through calculations such as statistics or modelling', and form the 'building blocks from which information and knowledge are produced, and constitute the input for and output from computational processes' (Kitchin 2025: 45). Their production, circulation and analysis are crucial to the development of science, the operations of digital systems, and the functioning of bureaucracy and the economy. The digital transformation of society in recent decades has led to a proliferation of data systems and infrastructures, the mass datafication of all aspects of everyday life (van Dijck 2014), and the development of data capitalism with an enormous expansion of data markets and services (Sadowski 2019). In turn, significant conceptual and empirical work has been directed at the production and use of data, accompanied by the establishment of the new, interdisciplinary field of Critical Data Studies (Kitchin 2022).

CONTACT Rob Kitchin rob.kitchin@mu.ie Maynooth University Social Sciences Institute, Maynooth University, Maynooth, Ireland

Some of this work has focused on the movement and circulation of data since the mobility of data is vital to much of the work undertaken, and value produced, with data. Indeed, data mobility is not incidental to the functioning of IT systems and data ecosystems, but is a constituent feature. It is essential to: aspects of the data lifecycle enabling data enrichment and data fusion to take place; the operations of data assemblages, with constituent elements being bound together and interacting through datafied connections and data sharing; data re-use and repurposing; and the creation of new data products and services.

To help make sense of data mobilities, a range of metaphorically-grounded concepts have been utilised, including: data flows (McNally et al. 2012; Hoeyer et al., 2017), data streams (Dourish and Gómez Cruz 2018; Hilgartner and Brandt-Rauf 1994), data journeys (Bates, Lin, and Goodale 2016; Leonelli 2020), data threads (White 2017), data transfers (Glouftsios and Leese 2023), data exchanges (Weltevrede and Jansen 2019), data arrows (Flensburg and Lai 2023), data circulation (Beer 2016; Pelizza 2016), and data sharing (Borgman 2015). While these framings of data mobility have utility, our contention is that their rootedness in metaphors of movement is misleading and needs rethinking, particularly with respect to digital data. Indeed, we contend that digital data circulate in a quite different way to other materials given their non-rivalrous (more than one entity can possess the same data) and non-excludable (it is easily copied and it takes effort to block sharing) nature, and the cost of reproduction has a zero marginal cost (it is effectively free to copy) (Floridi 2010).

Metaphorical conceptualisations such as data flow and data journey barely hint at the non-rivalrous nature and versioning of data and its effects. Rather than using metaphors that intimate, but do not specify in detail the means or nature of data mobility, we believe that it is more productive to describe the actual processes by which data are transferred between sites (where a site is a digital locus – files, databases, software programmes – on the same or different digital devices located across geographic space). As we illustrate through our case example of data mobilities in the operation of the development and control functions of a planning system, digital data do not move in the conventional sense of a journey – leaving one place to travel to another. Rather digital data replicate (replica copies produced), with the original source retaining the data and a new source gaining it. Data proliferate and are transformed as they multiply, altered by data cleaning, data wrangling, and their mediation by technologies, protocols, and practices. Numerous versions of the original source data can thus be produced, which can themselves be shared and proliferate.

Adopting a more nuanced conceptualisation of data mobilities is important, we believe, as it illuminates some fundamental issues with respect to the circulation and use of data. For example, it is widely recognised that science is presently experiencing a replicability and reproducibility crisis (Andreoletti 2020). In part, this crisis is created because of the many incarnations of datasets produced as they are shared, circulated, and transformed. The effects of such versioning is evident in data debates within fields. For example, in relation to planning in Ireland, there are several on-going data debates related to veracity and integrity of planning and housing data in part caused by actors creating different derived data from the same sources. Recognising the non-rivalrous nature of data mobility also helps to explain why exerting data sovereignty (possessing control over how data are circulated and used; Kukutai and Taylor 2016) can be difficult once data replicate beyond owned devices, and why it is challenging to erase all traces of a dataset. As Thylstrup (2022) notes, original source data and its residues and derivatives often persist long after the original data has been deleted, which is why revenge porn, and other abusive and derogatory content, continue to haunt and circulate the internet. In the remainder of the paper we examine the conceptualisation of data mobilities and illustrate our own theorisation through its application to the data ecosystem of the Irish planning system.

Data mobilities

As noted, a number of ways of describing the mobility of data have been forwarded in both the scientific (e.g. Data Science, Computer Science) and social science literature (e.g. Geography, Science and Technology Studies, Media Studies, Sociology). While there are some differences in the conceptualisation of data and the epistemological approach to studying data mobility (see Kitchin 2022 for an overview), these studies all frame data movement in metaphorical terms, with the use of those metaphors aligning quite loosely with disciplines. For example, data flow and data pipeline would be commonly used in the scientific literature, whereas data journey would be used within Media Studies and the Philosophy of Science, with the difference in choice of metaphor mainly centring on contrasting understandings of data: as a technical object and as a socio-technical construct. Here, we provide a general overview of the different data mobility metaphors employed within the literature, before setting out our own conceptualisation of data mobility that does not utilise a metaphorical framing but rather describes directly the processes involved.

Data flow is mostly used in the literature as a metaphorical descriptor to denote a liquid-like movement of data from one place to another. Data is said to flow between nodes, moving along a path or through a network (McNally et al. 2012; Tarantino 2020). As McNally et al. (2012) detail, these flows have variable temporalities (e.g. duration, rhythms, synchronisation, prioritisation), with the flow rate altering with context; data might flow freely or be viscous (van Schalkwyk, Willmers, and McNaughton 2016). Flows also vary in the extent to which their topographies (e.g. routes, sequences) are stable over time and with context (McNally et al. 2012). Not all data necessarily flows, forming what Hoeyer et al. (2016) call 'nonflows' (e.g. confidential data with restricted access).

A related concept, which also uses a liquid metaphor, is *data pipeline*, which refers to the ordered sequence through which data are generated and processed, with data being channelled along a workflow in a controlled manner, mediated by various actors and systems, as they move through different stages of production. The pipeline metaphor indicates the inherent movement of data between sites and systems of production towards use and wider circulation (e.g. the field, the lab, a database, analytics software, publication software, an open data repository) (Plantin 2019). As the data moves, it is transformed and extended through processes of cleaning, wrangling (e.g. formatting, generalisation, standardisation), data fusion and enrichment, review, verification, adding metadata, and analysis.

For others, the notion of a data pipeline is too static and fixed; instead they contend that data and their production are a *data stream*; that is, the production and movement of data consisting of a stream of processes that are never-ending (Hilgartner and Brandt-Rauf 1994). For example, the work of producing official statistics never stops, with a regular schedule of monthly, quarterly and annual releases, and the processes used are constantly reviewed and refined (Ruppert and Scheel 2021). Similarly, real-time data compose an endless data stream. The stream metaphor then captures the constant movement of data, both in terms of the data flow never ceasing, and the stream shifting its route and nature (as with water streams that migrate across the landscape).

In contrast to conceptualisations that conceive of data as liquid-like, others contend that data do not have liquid qualities (Bates 2018; Borgman 2015) and do not pass along frictionless pipes or streams, disconnected from the politics and praxes of data assemblages and the infrastructures that connect them (Bates, Lin, and Goodale 2016; Pelizza 2016). As Bates, Lin, and Goodale (2016) and Pelizza (2016) contend, the movement of data is subject to choices, negotiations and decisions by the various actors along a route, and is mediated by data practices, protocols and regulations, infrastructural supports, and data politics that can create data frictions that slow and block its progress. Moreover, data are transformed at points along the route as they move. In other words, the movement of data is more akin to a *data journey* than a flow (Bates, Lin, and Goodale 2016).

There are two strands of thinking and empirical work concerning the notion of a data journey. The first strand is rooted in Philosophy and Social Studies of Science (Howlett and Morgan 2011; Leonelli and Tempini 2020) and focuses on the conceptual conditions of data movement, and the data practices, data infrastructures, and protocols and governance arrangements that facilitate and mediate data movement and data re-use, mainly with respect to knowledge production (Leonelli 2020). The second strand has its origins in Communications and Media Studies, and in particular in the work of Jo Bates et al. (Bates 2018; Bates, Lin, and Goodale 2016, Bates et al. 2019), and conceives of data in more socio-material terms, with its production and movement bound up within the nature and workings of data assemblages, and shaped by the materialities and contingencies of mediating technologies and their attendant data politics and data power (Bates, Lin, and Goodale 2016).

The notion of data journeys has been critiqued by White (2017), who instead proposed the concept of *data threads*. Data journeys, in his view, gives the impression that data movement is a largely linear, sequential process along a singular path from origin to destination. Instead, he posits, the path is more complex, with data taking circuitous routes, encountering dead-ends, and looping back on themselves, and the movement can occur with no clear sense of the destination. Moreover, data often does not travel alone, and the paths of different data can be threaded together to form knots or tapestries, or fray and split apart.

The discussion so far has largely conceptualised data movement as a singular passage that might detour and loop, but moves along a path, pipeline, stream or thread from origin to destination. The notion of *data circulation* denotes the mobility scaling effects caused by the non-rivalrous and non-excludable nature of data. Rather than consisting of a singular movement along a defined path, the concept of circulation recognises that a single source of data can rapidly multiply, bifurcate, and combine with other data, travelling along multiple paths to many destinations (Beer 2016). This is especially the case once data are published and shared (Borgman 2015). Data, and its derived products (e.g. analytics, information), circulate through archives, data repositories and data markets to potentially thousands or more users. Key datasets on open data sites might be downloaded tens of thousands of times, with key facts or derived data visualisations moving into news media that are viewed by potentially millions of people. Here, the challenge of making sense of data mobility grows from tracking a single flow, journey or a handful of threads, to tracing tens of thousands of movements across a complex network stretching from the local to the global, with this circulation always unfolding and in flux. The exact mechanisms by which data circulate is rarely articulated, however.

Rethinking data mobilities

All of the data movement and circulation concepts discussed so far are united by an assumption that digital data travels like a liquid, material object or passenger. That is, data leaves an origin point and moves along a path between a series of sites and systems to a destination (or many destinations), and it can transform along this route as it is processed. However, we contend that such a conceptualisation is flawed. Instead, data *replicate* and *proliferate*. As evident in the notion of data circulation, data most often travel through copying, with the source data remaining at the origin (e.g. on an open data site) and a replica transferred to a new site (e.g. on the computer of the person who has downloaded a dataset) (Figure 1(a)). At the new site the data might be transformed through data wrangling or data fusion or enrichment, or be converted into derived data, with these transformed data then transferred to the next site (with or without the original, copied source data). In some cases, the data transformation can be significant, with downstream versions being derived tertiary data or the products of analysis (e.g. visualisations). Through replication digital data proliferate across many sites, with local transformations creating multiple versions (Figure 1(b)). Data mobilities then predominately consist of processes of copying, transformation and erasure, as data are produced, processed, reworked, fused, used, re-used and re-purposed across sites (networked databases and devices located across geographic space) and actors (state agencies, companies, civil organisations, individual analysts).

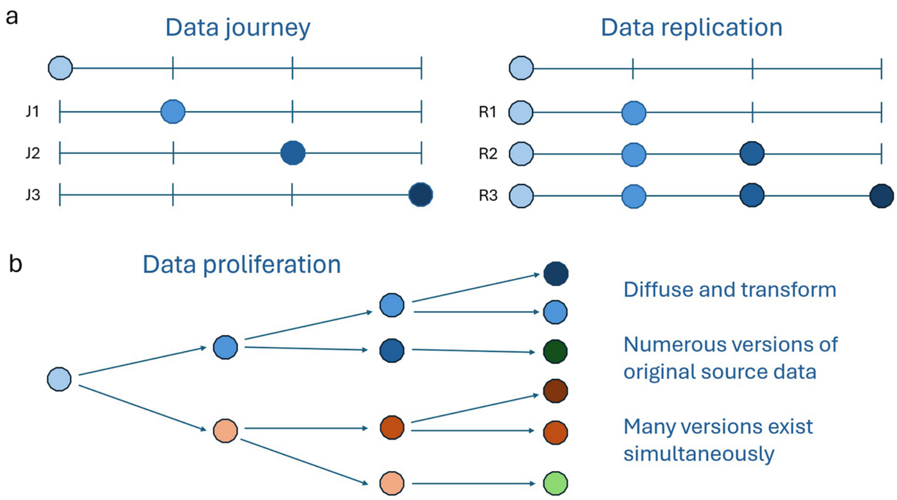

Figure 1. (a,b) Data journey, data replication, data proliferation and data transformation.

Data mobilities then, we contend, are quite different in nature to the movement of water (a flow) or transportation (a journey). Nonetheless, our conception of data mobilities shares some characteristics with data journeys, threads and circulation. In particular, data mobilities are contingent and contextual in nature (Bates, Lin, and Goodale 2016; Leonelli 2020). How data replicate and proliferate is never fixed or predetermined; rather, it is emergent and evolves in indeterminate ways depending on context, chance, uncertainty, and the agency of individual actors. With respect to the latter, individual planners undertake their data work in varying ways; for example, some planners comprehensively enter data into systems, whereas others will only enter the required fields, and all have latitude to compose open data fields in their own style and to select discretionary expert reviewers. As Metzler, Ferent, and Felt (2023, 3) notes, 'data mobility is always an effect of the relations in which data are entangled'. The pipeline of data production can be planned in detail through processes of articulation and scaffolding – identifying, assembling, scheduling, coordinating and monitoring all the tasks necessary to complete a job (Halfmann 2020; Nadim 2016). However, planning, and its implementation and operation is contingent, shaped by choices and decisions that are informed by personal experiences, negotiations, institutional politics, capacities, policies, and regulations, among other factors, along with the varying application of data practices, glitches, unanticipated interventions, and contextual matters (Loukissas 2019).

The contingency of data mobilities is evident in the effects of data frictions and seams, and the work of maintenance and repair. Data frictions are impediments or blockages that hinder the replication of data across sites, systems and actors (Bates 2018; Edwards 2010). Frictions can be caused by a number of human and technical issues, such as: incompatible data formats, standards, and systems that limit interoperability; mistakes, glitches, and disruptions; resistance and refusal of actors to cooperate and share data; cost, resource and skills capacities; and regulatory and legal limitations. A consequence of data frictions is inefficient or partial replications, error and noise (Lindsay 2017), or an inability to replicate creating stranded data (Kitchin 2022) or broken data (Pink et al. 2018). While data frictions are mostly framed as a hindrance, some exist for good reason; for example, to protect privacy or proprietary knowledge and to ensure data security. Data seams, in contrast, are the points of contact between component parts of a data system, or between data assemblages, that enables them to be conjoined or communicate (Inman and Ribes 2018; Vertesi 2014). The replication of data across seams is facilitated by data management and governance mechanisms, such as metadata, data dictionaries, data standards, and transfer protocols that enable interoperability (Gal and Rubinfeld

2019; Millerand and Bowker 2009). In addition to creating seams to smooth data frictions, processes of data maintenance and repair are employed, such as firmware patching and software updates, as well as repairs and replacement of physical system components (Mattern 2018). Without this maintenance and repair the systems underpinning data mobilities will atrophy and eventually stop working altogether.

In addition, our conception of data mobilities similarly recognises that digital data can be replicated in different forms (e.g. lists and tables using a variety of data formats, or assembled as hierarchical or relational databases), using different media, infrastructure and processes (e.g. CDs, DVDs, pen drives, external hard drives, email, file transfer, the cloud, and the use of APIs, ETL [extract, transform, load], scraping, and ingestion processes). Each of these work in varying ways, have different characteristics, and possess and create different data affordances. Likewise, our conceptualisation also acknowledges that data can be assembled into new forms and larger datasets, and data often replicate and diffuse alongside metadata, other datasets, and other entities (e.g. paradata, documentation, derived information, narratives) (Edwards 2010; Morgan 2011). As such, not only do data replicate, but so do their digital companions. As data are converted into information (e.g. visualisations, narratives) their forms and means of replication alter (e.g. via news and social media, and websites).

Case study and methodology

The planning system in Ireland, as in other countries, consists of three broad blocks of planning work: strategic planning (future development); development and control (planning applications, appeals and building control); and enforcement and compliance. These three blocks of planning work are organised and overseen by a multi-level, tiered system of governance, with planning practice and policy delivered by 31 local authorities (LAs) operating at the county scale, three regional authorities, and a handful of state agencies and government departments at the national scale. In all three blocks of planning work, the key stakeholders make extensive use of IT systems to undertake their planning function, and they generate, handle, process, analyse and share substantial volumes of data. Importantly, they do not work in isolation: their functioning are dependent on the sharing of data between planning IT systems, and linkages to other IT systems, such as financial systems (related to payment of fees), file management systems, and analysis systems (e.g. GIS). Here, we focus our analysis on the development and control function of the planning system, which has its own distinct, multi-level data ecosystem.

To make sense of the development and control data ecosystem in Ireland and its attendant data mobilities we undertook an empirical study using a number of related methods. The fieldwork was conducted between June and August 2023. Interviews were undertaken with 29 public sector officials involved in data work at local, regional and national scale within the planning system. A purposeful sample was employed, with interviewees drawn from all the main organisations involved operating planning IT systems within the data ecosystem, including 6 LAs, the Local Government Management Agency (LGMA), An Bord Pleanála (ABP; the national planning appeals body), the National Building Control and Market Surveillance Office (NBCMSO), the Office of the Planning Regulator (OPR), the Department of Housing, Local Government and Heritage (DHLGH), the Department of Environment, Climate and Communications, and the Central Statistics Office (CSO; Central Statistics Office 2023). In each case, the person being interviewed was overseeing the management of, or was actively undertaking, planning data work using IT systems. A number of the interviews were of a walk-through nature, with the participant demonstrating the workflow – data entry, data fusion, data management and data-informed decision-making – related to the use of an IT system. This was supplemented with a close reading of the user manuals for these systems. In addition, a full data audit was undertaken of the core planning application management systems used by LAs (iPlan, APAS and Odyssey), the Building Control Management System (BCMS), and planning.localgov.ie. We also examined a number of downstream open data sites and planning/housing data tracking tools (e.g. Dublin Housing Observatory, Housing Delivery Tracker, Housing for All dashboard). The walkthrough interviews and data audits were used to chart the data ecosystem as a whole (see Figure 2). Next, we constructed charts of data mobilities for each stage of the process. Each chart, organised along a timeline, detailed the replication processes employed and key data-informed decision points (see Figure 3 for an example). Collectively, this set of charts provided a detailed mapping of data mobilities in the planning and development system from application to completion of properties.

The data mobilities of development and control

The development and control function of the Irish planning system consists of a sequential and time-ordered set of phases and tasks for assessing planning applications, appealing decisions, and monitoring the construction process. Each phase has an associated workflow of tasks and at least one IT system that is used to undertake and manage these tasks. Just as an application passes through phases and systems, so too are data shared between them (as indicated by the blue and green arrows in Figure 2), in order to inform operations and decision-making.

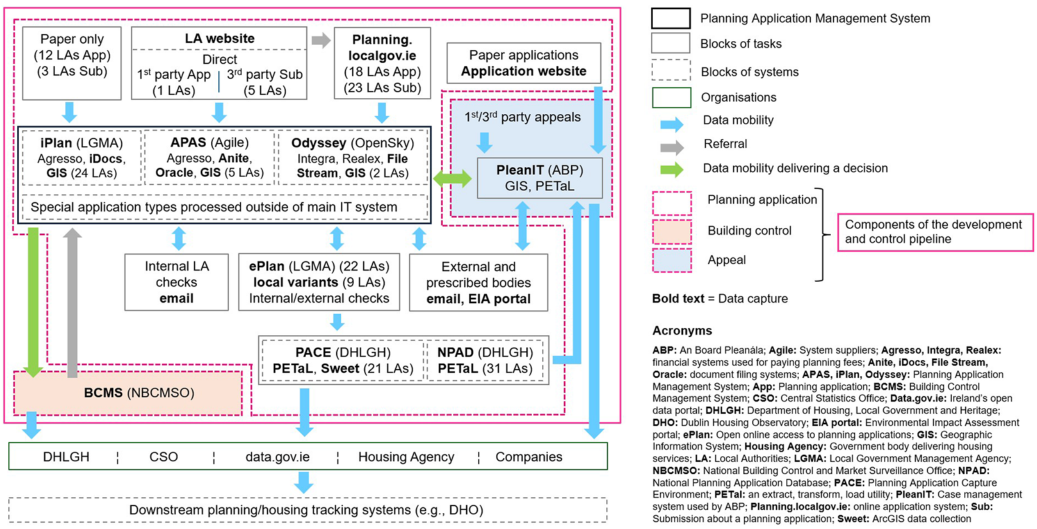

Figure 2. The development and control data ecosystem in August 2023.

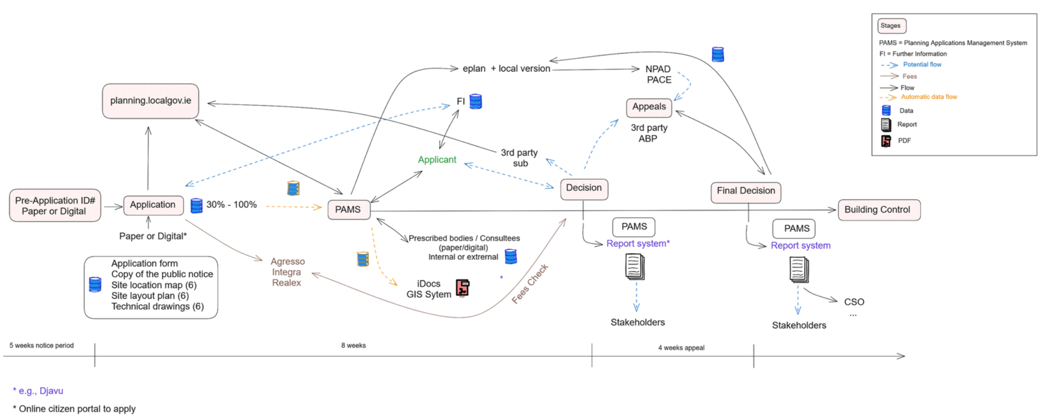

Figure 3. Data mobilities in the planning application phase.

As Figure 2 illustrates, this system of development and control is reasonably complex but, in brief, it is organised and operates in the following way.

Planning application

After an initial stage of pre-planning consultation, applicants are required to post a notification about the proposed development in a local newspaper and on a site notice at least two weeks prior to submission, thus sharing some outline data and information with the local community. After a five week notice period, they then submit a planning application that consists of a copy of the site notice, the site location map, site layout plan, technical drawings, and details of the site development (materials, infrastructure, services). If submitting by paper and post (required for 12 LAs), six sets of documents must be included. On receipt, a LA worker will scan and file the documents and manually enter key data into a Planning Application Management System (PAMS). 19 LAs allow an application to be made digitally, with 18 using a shared services web portal – planning.localgov.ie (administered by the LGMA) – which provides a data entry and upload facility. The uploaded documents are filed in a document management system and the data in the online application form imported into a PAMS. There are three main variants of PAMS used across the 31 LAs: iPlan (produced by the LGMA) used by 24 LAs; Odyssey (produced by Open Sky) used by 2 LAs; and APAS (produced by Agile) used by 5 LAs. In the case of the five LAs using APAS, each LA has locally configured the system to its own needs, with each instance quite different in terms of design and workflow. These systems are used to manage and track the progress of assessment, including sourcing relevant information, monitoring fee payment, tracking all communications with the applicant and third parties, noting observations and decisions, and ensuring that tasks are undertaken on the proscribed timeline, and this involves a substantial amount of processing and sharing data (Figure 3).

After initial processing and assessment, a case manager will seek feedback and an assessment of the application. Requests for expert views and additional information are sent by email to internal LA units (e.g. transportation, environment, and archaeology and heritage departments) and selected external bodies (e.g. OPR, Land Development Agency) and prescribed bodies (e.g. DHLGH, Transport Infrastructure Ireland, Health Service Executive, Environmental Protection Agency). Prescribed bodies have to be consulted and are defined by legislation. Checks are also made against the registry of protected structures using an Environmental Impact Assessment (EIA) portal, and with respect to strategic plans (e.g. local area plans, land zoning). The responses to these queries are recorded as data fields in the PAMS.

To enable the general public to assess and make submissions on proposed developments, selected details of the application are replicated on a publicly accessible website and are also made available to walk-in visitors to the LA planning department. 22 LAs use the ePlan system (eplanning.ie, produced by LGMA), with the four Dublin LAs using an Agile platform (planning.agileapplications.ie), the two Galway LAs using a shared site, and the remainder using their own dedicated sites. These systems are purely used for communication, with third party submissions on applications made in three ways depending on the LA: 23 LAs using planning.localgov.ie, 5 LAs using their own dedicated website, and 3 LAs via paper submission only. Key information from third party submissions are added to the PAMs and all materials stored in the document file system. In addition to each LA replicating the data on an eplanning website, a selection of data for all LAs are collated within a single site, the NPAD (National Planning Application Database) produced by the DHLGH. NPAD is a national, publicly accessible, online GIS mapping tool that displays the location of all planning applications since 2012 for all 31 LAs, along with summary planning application information (25 fields), and a link to the planning files in ePlan or its variants. The replication of data into NPAD is automated, using a custom ETL process called PETaL (Planning, Extract, Transform and Load). In addition to NPAD, Tailte Éireann and DHLGH provide LAs (used by 17) with a standardised tool for digitally capturing the site boundaries of planning applications, pulling all these data into a centralised system, PACE (Planning Application Capture Environment).

LAs are mandated to process and make a decision on applications within eight weeks of receiving them. If at any point further information is sought from the applicant by the case manager, this can reset the clock and extend the time for making a decision. Once a decision has been made, there are two routes forward to the next stage. Either to appeal the decision (the first party objecting to the denial or by a third party objecting to approval or denial), or to proceed to development and building commencement.

Appeals

An Bord Pleanála (ABP) is the national planning appeals body and has two principal remits. First, to investigate and adjudicate on first or third party appeals to planning permission decisions. Second, to process specific types of planning applications that are sent directly to ABP (e.g. Strategic Infrastructure Developments (SIDs) such as motorways and hospitals, and specific cases regarding the Development (Emergency Electricity Generation) Act 2022). ABP remains a paper-based organisation for legal and statutory reasons (an amendment to existing legislation is required to enable the adoption of a digital approach). While it uses IT systems, notably PleanIT (its own planning applications management system), a document management system, and a GIS, appeals and applications are made through a paper submission and it has paper-based versions of all files and it prints out, thus replicating, all emails, adding them to its paper-based document filing system. For large applications and appeals the applicant is expected to create a website that contains all pertinent information that ABP can then access and replicate.

On receipt of an appeal details are entered into PleanIT, a unique ID assigned to track progress, and a case inspector assigned to assess the appeal. ABP will at this point also correspond with the responsible LA to notify them of an appeal and request relevant planning documents, replicate information from its ePlan or variant system and from NPAD via PETaL. Basic data on the planning application, along with technical drawings, maps, and documents, are replicated automatically in PleanIT, while other details are entered manually. A similar process to LA assessments then takes place in which prescribed bodies and the public are invited to provide observations about an application/appeal, with the provided information being recorded in PleanIT. ABP has 18 weeks, organised into a number of timed phases, in which to investigate and make a decision on an appeal (Figure 4). When ABP makes a decision, it notifies the applicant and appellant, as well as the relevant LA. The appeal decision will be recorded in the LA PAMS, along with the ABP ID number, the date of appeal, appeal decision, and other relevant information.

Building control

Once planning permission has been granted the applicant is able to move to the development phase. All development in Ireland must be compliant with the regulations and standards set out in Building Control Act 1990, with this being monitored through the BCMS, first introduced in March 2014 as a shared service for all LAs and administered by the NBCMSO. The data and documentation to be submitted at commencement stage differs depending on the nature of the works. One-off houses and extensions are exempt from tracking using BCMS, though the intent to proceed with an exemption needs to be registered via the submission of an opt-out declaration. For these works, no other documentation is required at this or any subsequent stage of development. All other development must register and upload relevant data at set points along the construction process. These data enter BCMS in the form of a series of notices, certificates, and statutory documents filled out by building owners and assigned certifiers, and validated by the relevant Building Control Authority (i.e. LA). In the simplest case, this occurs

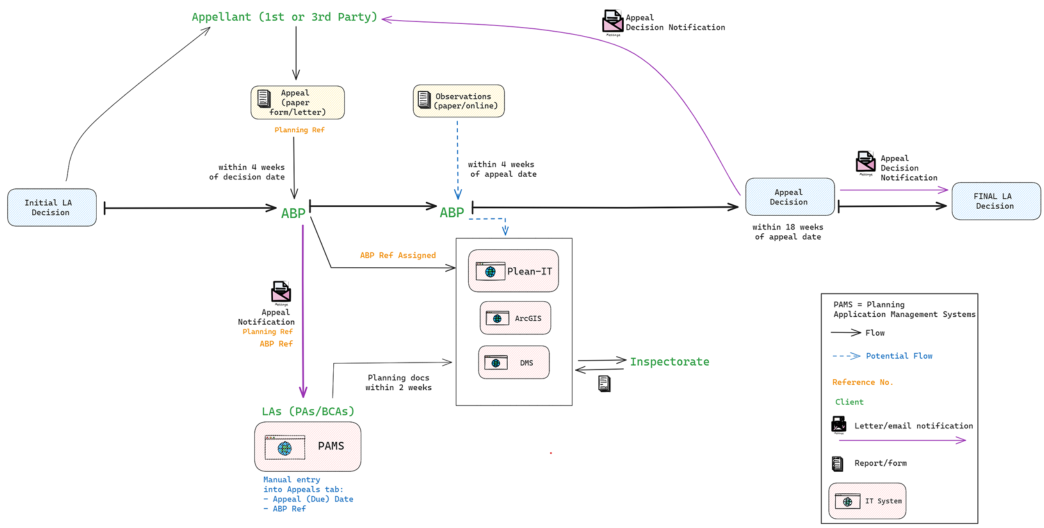

Figure 4. Data mobilities in the appeal process.

at two main points in the development timeline. First, between 28 and 14 days prior to works being commenced (Commencement Notice); second, upon completion of the works (Certificate of Compliance on Completion). While online submissions are encouraged for all types of notice, paper submissions are accepted, with relevant information then entered into the system. For works which meet the criteria as set out in legislation, additional compliance documentation is required, including the nomination of a designer and assigned certifier, a granted Fire Safety Certificate (FSC), a certificate of design, and technical drawings/maps which demonstrate the proposed development's compliance with the regulations (Figure 5).

Beyond the planning system

Some of the data generated within the planning application and assessment process, as well as the building control process, are made available as open data and are published as official statistics in an aggregated form. The CSO are responsible for producing official planning permissions statistics and receive details on 14 variables from LAs/ABP monthly, which are compiled into aggregated data at LA, regional and national scale. These data are also published on the national open data portal, data.gov.ie. The DHLGH receive regular cuts of development and control data from LAs, and can also request LAs to compile planning data as needed for the monitoring of planning and policy. A limited set of compiled data are shared via the DHLGH open data portal and through NPAD, and also by the Housing Agency. ABP publish a set of data relating to planning appeals and applications in their annual report. The data made available through open data sites, or through specially prepared cuts of the data, are also accessible to sector stakeholders and the public through a number of data visualisation sites that are designed to enable the monitoring and tracking of planning and development. These include:

- DHLGH Housing Delivery Tracker (https://storymaps.arcgis.com/stories/ab12ed6d50a540e2891170c24955ff49)
- Housing for All dashboard (https://public.tableau.com/app/profile/statistics.unit.housing/viz/HousingforAll/0_Overview)
- OPR Digital Planning Hub (https://opr-hub-oprgis.hub.arcgis.com/)
- Dublin Housing Observatory (https://airomaps.geohive.ie/dho/)

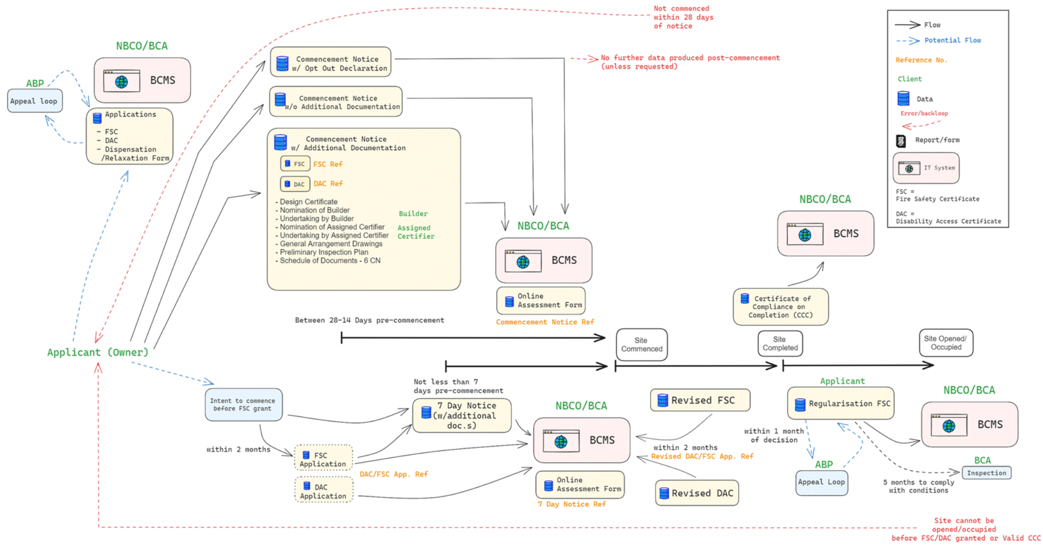

Figure 5. Data mobilities in building control.

- Dublin Housing Task Force mapper (https://housinggovie.maps.arcgis.com/apps/View/index.html?appid=3fa56a71ee774f9487d14a9e5336b00c)

In addition to the official, mandated sites for accessing data, unofficial mobilities exist through data scraping by private companies. Construction Information Services Ireland (CIS Ireland) and Building Information Ireland (BII) are two companies that specialise in harvesting, validating, wrangling, and enriching such data to create data products and tools. Each day (CIS), or each week (BII), scrape data and documentation relating to new applications from the ePlan and variant systems, and new registrations in BCMS. They extract relevant data from the documents, link data with other relevant datasets, such as procurement data sourced through e-tender websites, and compile it with respect to each proposed development. The data are also used to compile a number of high-level planning and construction statistics at LA, regional and national level, and across different sectors of activity (e.g. residential, commercial, industrial, etc.).

The nature of data mobilities

As made clear in the previous section, the development and control process in Ireland involves extensive data mobilities to sustain its organisation and operation. Indeed, it is reliant on such mobilities to function as a constituent whole given that data needs to be sourced at various points in the application assessment and data generated or assembled at one phase is required for use in other phases. Here, we consider the nature of these data mobilities, how they work in practice through replication and proliferation, the various data frictions at play, and how data mobilities transition in organisation and form.

Forms and entanglements of replication

Throughout the development and control process each actor and IT system, at each stage, retains the data shared with other actors and IT systems. For example, the applicant retains a copy of their application at submission (regardless of whether it is a paper or digital submission), and a PAMS retains a copy of the material shared with the ePlan systems. In other words, data are replicated, not moved. As documented, how replication takes place is multiple in form, including post, email, manual data entry through typing and cut-and-paste, manually-directed file transfer, data scraping, and automated APIs and ETL processes. These replication processes are not fixed and invariant in nature, but contingent and contextual. For example, manual data entry practices vary across operatives, with some staff entering only the required fields and others entering as many fields as possible, with most somewhere between; automated ETL processes can be glitchy, failing to work if there are changes to firewall settings, server configurations, or database design, and need constant maintenance and repair. Replication processes are stream-like in that they are continually taking place – every day thousands of replications are occurring between the various systems. As such, Figures 2–5 would ideally be animations rather than fixed images to illustrate the dynamism of their data mobilities. In addition, replication is rarely a singular occurrence, but involves companions, such as correspondence and instructions (e.g. email, forms), other data (e.g. metadata, derived data, additional datasets), and information (e.g. documentation, narrative, visualisations), and data are often assembled into, and circulate as, larger datasets (e.g. databases) (Edwards 2010).

The replication of data within the planning system has a relative degree of path dependency (a defined and self-reinforcing sequence of tasks), due to the rule-set for evaluation, scripted practices, established protocols, embedded institutional workflows, and technical configuration of systems (Payne 2014; Poirier 2022). Nonetheless, as Figures 2–5 make clear, the mobility of data within the planning data ecosystem does not consist of a neat, linear sequence of replication, transformation and proliferation. Rather, how development and control works is through entangled sets of replications used to assemble data, populate IT systems, and make decisions. Some of these replications are unidirectional, data replicated from one system into another (e.g. data from PAMS into ePlan). In other cases, replications are bidirectional, with data from one system replicated into another, and other data replicated in response (e.g. application data from PAMS is shared with internal and external units, along with a request for evaluation, with an assessment and associated relevant data returned, which is added to the PAMS). In some cases, the data pass through a number of replications and transformations before looping back to an origin point (e.g. from PAMS to ePlan to NPAD to PleanIT to PAMS). At certain points, there can be loop backs, with a process re-set and re-performed (e.g. at any stage of the planning application phase (Figure 3), a case officer can ask the applicant to revise their application to address identified issues or to supply additional data and documentation, then re-run the tasks already performed). The data mobilities for each planning application are contingent and variable: despite being evaluated against the same criteria, using the same workflows, no one application consists of the same set of replications and processes.

Transformation and proliferation

Once replicated, new data in a system are often transformed in a number of ways to make them amenable to how they are to be managed, analysed and used. For example, data might be subjected to data cleaning to remove personal and sensitive data to comply with data protection regulations and to improve data quality, data wrangling to restructure data and produce derived data, data fusion to merge and assemble more extensive datasets, and data analysis to create information (e.g. visualisations and data narratives). As a result, multiple new versions of the data are produced that differ from the data retained at the source. In turn, these transformed datasets are themselves replicated. These process of mutation and data transformation are illustrated in Figure 6, with v1 representing the original source data. This data is replicated in the iPlan system with personal data removed (v2). A selection of these

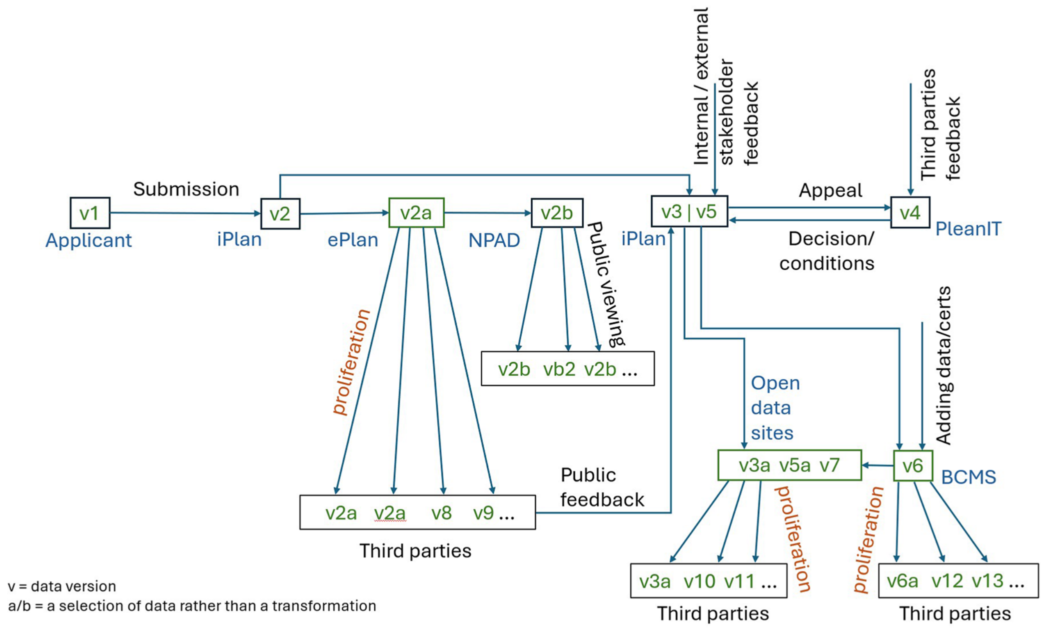

Figure 6. Data replication, transformation, proliferation.

data are then replicated in ePlan (v2a), and a further refined selection are replicated in NPAD (v2b). The addition of new data replicated via internal/external stakeholder feedback and via third party submissions in response to the ePlan data and documents produce an extended set of data (v3). If a decision is appealed, a selection of the v3 data is shared with PleanIT, which is subsequently extended by third party feedback (v4), with some data fed back to iPlan (v5). A selection of the v3/v5 data is also shared with the open data sites (v3a/v5a) and with the BCMS system (v6). Data mobilities then lead to multiple incarnations of an original source dataset (v1–v6), each of which has different data affordances (what processes and outcomes they enable) (Fjørtoft and Lai 2021), though not all of these incarnations exist simultaneously (e.g. the v1 dataset has been progressively extended to become v3).

For much of the development and control phases, replication occurs in defined and stable forms that involve specific actors, limiting proliferation. Proliferation occurs, however, at three points that enable the data to be downloaded and used by anybody globally with an internet connection for re-use and repurposing. When replicated onto ePlan, key data relating to the planning application are open to viewing and download by the general public and to scraping by private companies, who can subsequently transform the data (v8, v9). The same is the case for the data recorded in the BCMS (v12, v13), which is openly accessible. Likewise, proliferation can occur when selected, aggregated data are published on open data sites (v10, v11), from which the data might be replicated hundreds or thousands of times. These data, and their transformed incarnations, might then be hosted on other sites. In contrast, the data on NPAD (v2b) is viewable by the general public but not available to download, meaning they cannot easily be replicated and hence are less likely to proliferate or transform. A consequence of transformation and proliferation is that: two sets of analysis supposedly of the same source data might be undertaken on data sets that vary, producing discrepancies in outcomes and interpretation (Kitchin, 2022); there is no end to the data lifecycle as data incarnations persist across many systems and continue to be mobile outside the control of the data processor/controller (Thylstrup 2022). It should be noted that datasets rarely circulate virally, with mass replicating downloads or active sharing, given the specialist skills needed to process and analyse data, though data visualisations and other derived information might be shared and circulate as such via social and traditional media.

The versioning of data raises questions about the status of each version and if one version constitutes the 'official' or 'final version'. In the case of the Irish planning system there are four official versions of datasets. The data in the PAMS and PleanIT are the official, legal data relating to planning permissions for LAs and ABP, respectively. The data in the BCMS is the official building control dataset. The summary data published by the CSO are the official planning statistics. However, in practice, these official datasets are little used by those outside of these organisations. PAMS and PleanIT data are not open, though a selection of PAMS data can be scraped from ePlan websites. The BCMS is available as open data, but it is widely acknowledged that it constitutes a 'dirty database' and needs extensive cleaning and checking before use. The CSO planning statistics are tertiary, derived data, aggregated to the local authority scale and published monthly and lack granularity and detail. Consequently, official data are often not the 'final' version of a dataset used by planning and development professionals, academics and other analysts. Instead, they might use the datasets that are transformed versions of scraped PAMS data or BCMS data produced by planning data companies such as CIS or BII, whose business model is to sell a 'final', higher quality (the data are cleaned and validated) version of the data (which is sometimes also used by local authorities and state agencies rather than using their own versions). Or they might perform their own bespoke data cleaning and wrangling to create their own 'final' version. As a result of official datasets being closed or lacking in veracity, and analysts using multiple versions of 'final' datasets, there can be discrepancies between datasets, particularly if the data are transformed through classification or categorisation where different typologies are used. Such variation can create different impressions about planning outcomes, and raises questions about whose dataset to trust.

Data seams and frictions

A number of mechanisms, such as internal and shared data management and governance, standardised forms and templates, standardised workflow within organisations, and the use of APIs and ETL processes, have been put in place throughout the data ecosystem to try to ensure that replication and proliferation occurs as intended and to minimise data frictions. How these processes work in practice to facilitate replication across data seams varies depending on personnel and institutional cultures and priorities, and technical specifications. In the case example, the seams linking data systems can be unstable and break; for example, changes to system or firewall configuration can block the functioning of ETL processes. There are several active data frictions that hamper the smooth operation of replication and proliferation. For example, the continued use of paper and manual data re-entry is inefficient and weakens data quality through mistyping and miscodings. The lack of a consistent ID reference number across systems, with a planning application receiving unique IDs at pre-planning, planning, appeals, and construction phases, impedes the ability to track the development pipeline.

The lack of standardisation of data form and availability across PAMS hinders the production of harmonious national-scale datasets and official statistics by making it difficult to conjoin data (e.g. iPlan and Odyssey make extensive use of free text fields, and Odyssey and APAS make strong use of check boxes and drop-down selections; iPlan has 65 required fields, whereas Odyssey has 40 and APAS 21). It presently takes a planning officer approximately a week every month to extract the required data from PAMS for submission to the CSO for the compiling of official planning statistics. In turn, the CSO then spends a considerable amount of time cleaning and wrangling these data into the required standardised measurement units, types and classes. There are a number of specialised planning applications handled by local authorities that are not processed through the standard development and control process, and in some LAs these are not administered by IT systems identified in Figure 2. These include applications under

Sections 5, 35, 42, 44, 44, 57, 247 and Parts V, VII and XI of the Planning Act. A consequence of their separate administration is that special procedures are required in order for the data recorded to be passed into downstream systems. ETL processes require on-going maintenance and repair. Data frictions expose the contingent nature of data mobilities revealing how they unfold varies; that they are always in the process of taking place.

Data mobility transitions

How the data ecosystem is assembled and how its data mobilities operate is not static, but is subject to alteration as new versions of software, new work processes and new IT systems are put in place, and new regulations, policies and legislation alter the rules and processes of assessing planning applications. In general, changes occur through slight, incremental shifts in workflows, system configurations, and data practices over time, though occasionally a more radical alteration can occur through a critical juncture (e.g. regime change, a crisis, or a 'game-changing' new innovation) (Rast 2012). At the time we were undertaking our fieldwork parts of the development and control process were undergoing digitalisation. Twelve LAs still only accepted planning applications in a paper form and one LA used its own portal. In the 12 months afterwards, 9 LAs transitioned to using the planning.localgov.ie portal and the remaining 4 were due to phase out paper applications by the end of 2024. At the start of 2024 the two Cork LAs transitioned from using Odyssey to APAS as its PAMS. During the first half of 2024 four LAs transitioned to using ePlan for sharing planning application documentation with the public, raising the number of participant LAs to 26. The remaining five LAs were planning to transition. In addition, four more LAs (all iPlan users) adopted the use of PACE.

Since we undertook our fieldwork, the LGMA have undertaken a review of the IT support and processes for development and control and throughout 2024 were working on a business case for a new national planning system designed to significantly reduce data frictions and improve data harmonisation and quality (Kitchin et al. 2024). The new system, if adopted, would replace iPlan and the various iterations of APAS with a single planning application management system that has a single workflow and data schema. This would mean all 31 LAs using in a consistent manner planning.localgov.ie, the new PAMS, ePlan, PACE and NPAD. Further, the Department of Taoiseach (Prime Minister's office) and the Housing for All inter-departmental group have been undertaking a review of BCMS with a view to significantly improving its data capture processes to ensure more accurate and consistent data and to improve the ability to trace a proposed development from planning application to turn-key. These reconfigurations would mean the associated data mobilities would be somewhat reorganised, streamlined and harmonised, with consistent, higher quality data becoming available as official statistics and open data.

Conclusion

The mobility of data is vital to the work of data assemblages and data infrastructures, linking them together into functioning data ecosystems. Data mobilities enable data enrichment and data fusion, new forms of data analysis, enhanced data-informed decision-making, and the creation of data products and services. In the case of planning, data mobilities are fundamental to the operation and delivery of the state's planning function, binding a set of mutually constitutive data assemblages together into a data ecosystem. Data mobilities are the means of sourcing, assembling and sharing evidence in the form of plans, facts, figures, views and opinions, assessing these evidence, and enacting evidence-based decision-making, to ensure that built structures in an environment comply with regulations and policy at all stages of the development pipeline from pre-planning to turn-key.

Up until now, how data are shared and circulate has been conceptualised metaphorically as a movement from here to there, where the data flows or journeys along a path, pipeline, stream or thread from one system or place to another. In contrast, we have argued that data do not flow (like water) or journey (like other goods and passengers), rather data replicate and proliferate. Data are copied, with the source retaining the data and the destination receiving a replica. As illustrated through our case study, at all phases of development and control in planning, data are replicated across sites (various networked databases and devices located across offices) and actors (local authorities, state agencies, individual analysts) using a variety of processes (post, email, manual data entry through typing and cut-and-paste, manually-directed, file transfer, data scraping, and automated APIs and ETL processes). At selected points the data are made openly available, enabling them to proliferate by being replicated (downloaded or scraped), stored and shared by the other actors. Once replicated, the data are often transformed through processes of cleaning, wrangling, fusion and analysis. This produces various incarnations of an original source dataset, with multiple versions existing simultaneously.

Moving beyond metaphorical conceptualisations to identify the mechanisms by which data become mobile is important as it starts to shed light on a number of data issues. The replication, transformation and proliferation of data means that multiple incarnations of datasets are simultaneously being used as sources of evidence for formulating knowledge and policy, and informing decision-making. While there might be an 'official' version of a dataset, this is by no means the 'final' version even within the organisation holding the official dataset as they will often process and transform the data to perform analysis. Moreover, there might be multiple 'final' versions, each unique to those that produced them. The 'final' versions might be designed to investigate the same issue, or to answer a different question. When designed to produce insight on the same issue, divergences in any transformations applied will lead to variations across the datasets. These variations raise questions of veracity, validity and trust in the dataset and its use, and pose challenges for replicability and reproducibility in science. Replication and proliferation also reveals why maintaining data sovereignty can be difficult, and why it is difficult to erase all traces of replicated data.

Our analysis suggests three lines of future work are required in relation to data mobilities. First, we need more attention paid to data mobilities as a constitutive feature of data work, recognising that the creation and use of data are never static but inherently consist of mobile practices. In particular, there is a need to conduct further empirical analysis of data mobilities in action and their characteristics and nature through case studies relating to different contexts. Second, we need to more thoroughly ground an analysis of data mobilities within a mobilities perspective. To date, consideration of how data replicates and proliferates has largely developed independently of the mobilities literature, though it is occasionally cited, and mobilities scholars have rarely made data the central focus of their analysis despite significant attention being paid to the mobilities enabled and enacted through digital technologies and platforms (Sheller and Urry 2006; Stehlin, Hodson, and McMeekin 2020). Third, we need to continue to develop mobilities theory with respect to data. We have sought to advance the conceptualisation of data mobilities, arguing that it is more productive to consider the mobility of data as replication and proliferation, rather than as a flow, journey, chain, thread, or circulation. This conceptualisation needs to be applied, tested and refined with respect to other cases, and no doubt its veracity and utility will be challenged by new ideas and concepts that will further advance our understanding of data mobilities. Given the critical role of data mobilities to the management and governance of society and the functioning of the global economy, such theoretic and empirical work, we believe, is vital.

Acknowledgements

During the research regular meetings were held with the LGMA Research Advisory Group and we are thankful to Dr. Bernie O'Donoghue Hynes, Tom Elders, Tara Kerrins, and Holly Morrin for facilitating access and providing direction and constructive advice and feedback. We are grateful to 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 Oliver Dawkins for feedback on an initial draft.

Ethical approval

The research received ethical approval from Maynooth University Research Ethics Committee and the Ethics Review of the European Research Council. Interviews were undertaken with consent but a condition of one of the project funders was that the interview transcripts could not be archived or quoted..

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The research reported in this paper was funded by the European Research Council [no. 101052998] and the Local Government Management Agency.

Data availability statement

The interviews for the study were recorded, but participants did not give written consent for their data to be shared publicly.

ORCID

Juliette Davret http://orcid.org/0009-0005-7778-6595
Carla Maria Kayanan http://orcid.org/0000-0002-4359-3534
Samuel Mutter http://orcid.org/0000-0002-9246-7147

References

Andreoletti, M. 2020. "Replicability Crisis and Scientific Reforms: Overlooked Issues and Unmet Challenges." *International Studies in the Philosophy of Science* 33 (3): 135–151. https://doi.org/10.1080/02698595.2021.1943292.

Bates, J. 2018. "The Politics of Data Friction." *Journal of Documentation* 74 (2): 412–429. https://doi.org/10.1108/JD-05-2017-0080.

Bates, J., P. Goodale, Y. Lin, and P. Andrews. 2019. "Assembling an Infrastructure for Historic Climate Data Recovery: Data Friction in Practice." *Journal of Documentation* 75 (4): 791–806. https://doi.org/10.1108/JD-08-2018-0130.

Bates, J., Y.-W. Lin, and P. Goodale. 2016. "Data Journeys: Capturing the Socio-Material Constitution of Data Objects and Flows." *Big Data & Society* 3 (2): 1–12. https://doi.org/10.1177/2053951716654502.

Beer, D. 2016. *Metric Power*. London: Palgrave.

Borgman, C. L. 2015. *Big Data, Little Data, No Data: Scholarship in the Networked World*. Cambridge, MA: MIT Press.

Central Statistics Office. 2023. "CSO Statement on New Dwelling Completions." https://www.cso.ie/en/csolatestnews/csostatementonnewdwellingcompletions/.

Dourish, P., and E. Gómez Cruz. 2018. "Datafication and Data Fiction: Narrating Data and Narrating with Data." *Big Data & Society* 5 (2): 1–10. https://doi.org/10.1177/2053951718784083.

Edwards, P. N. 2010. *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. Cambridge, MA: The MIT Press.

Fjørtoft, H., and M. K. Lai. 2021. "Affordances of Narrative and Numerical Data: A Social-Semiotic Approach to Data Use." *Studies in Educational Evaluation* 69: 100846. https://doi.org/10.1016/j.stueduc.2020.100846.

Flensburg, S., and S. S. Lai. 2023. "Follow the Data! A Strategy for Tracing Infrastructural Power." *Media and Communication* 11 (2): 319–329. https://doi.org/10.17645/mac.v11i2.6464.

Floridi, L. 2010. *Information: A Very Short Guide*. Oxford: Oxford University Press.

Gal, M., and D. L. Rubinfeld. 2019. "Data Standardization." *NYU Law Review* 94 (4): 737–770.
Glouftsios, G., and M. Leese. 2023. "Epistemic Fusion: Passenger Information Units and the Making of International Security." *Review of International Studies* 49 (1): 125–142. https://doi.org/10.1017/S0260210522000365.
Halfmann, G. 2020. "Material Origins of a Data Journey in Ocean Sciences: How Sampling and Scaffolding Shape Data Practices." In *Data Journeys in the Sciences*, edited by S. Leonelli and N. Tempini, 29–44. Cham, Switzerland: Springer Open.
Hilgartner, S., and S. I. Brandt-Rauf. 1994. "Data Access, Ownership, and Control: Toward Empirical Studies of Access Practices." *Knowledge* 15 (4): 355–372. https://doi.org/10.1177/107554709401500401.
Hoeyer, K., A. Tupasela, and M. B. Rasmussen. 2016. "Ethics Policies and Ethics Work in Cross national Genetic Research and Data Sharing: Flows, Nonflows, and Overflows". *Science, Technology, & Human Values* 42 (3): 381–404. https://doi.org/10.1177/0162243916674321
Howlett, P., and M. S. Morgan. 2011. *How Well Do Facts Travel?* Cambridge: Cambridge University Press.
Inman, S., and D. Ribes. 2018. "Data Streams, Data Seams: Toward a Seamful Representation of Data Interoperability." Paper presented at Design Research Society, University of Limerick, Limerick, September 25–28. https://dl.designresearchsociety.org/drs-conference-papers/drs2018/researchpapers/16/.
Kitchin, R. 2022. *The Data Revolution: A Critical Approach to Big Data, Open Data, and Data Infrastructures*. Second edition. London: Sage.
Kitchin, R. 2022. *The Data Revolution: A Critical Approach to Big Data, Open Data, and Data Infrastructures*. 2nd edition. London: Sage.
Kitchin, R. 2025. *Critical Data Studies: An A to Z Guide to Concepts and Methods*. Cambridge: Polity Books.
Kitchin, R., J. Davret, C. Kayanan, and S. Mutter. 2024. *A Genealogy of a Data Ecosystem: The Digitalisation of the Ireland's Planning Development and Control, 2000–2024* (Data Stories Working Paper 5). Ireland: Maynooth University. https://mural.maynoothuniversity.ie/id/eprint/19264/
Kukutai, T., and J. Taylor. 2016. "Data Sovereignty for Indigenous Peoples: Current Practice and Future Needs." In *Indigenous Data Sovereignty: Towards An Agenda*, edited by T. Kukutai and J. Taylor, 1–22. Canberra: Australian National University Press.
Leonelli, S. 2020. "Learning from Data Journeys." In *Data Journeys in the Sciences*, edited by S. Leonelli and N. Tempini, 1–26. Cham, Switzerland: Springer Open.
Leonelli, S. and N. Tempini (eds). 2020. *Data Journeys in the Sciences*. Cham, Switzerland: Springer Open.
Lindsay, J. R. 2017. "Target Practice: Counterterrorism and the Amplification of Data Friction." *Science, Technology & Human Values* 42 (6): 1061–1099. https://doi.org/10.1177/0162243917727353.
Loukissas, Y. 2019. *All Data Are Local: Thinking Critically in a Data-Driven Society*. Cambridge, MA: MIT Press.
Mattern, S. 2018. "Maintenance and Care." *Places Journal* (2018). https://doi.org/10.22269/181120.
McNally, R., A. Mackenzie, A. Hui, and J. Tomomitsu. 2012. "Understanding the 'Intensive' in 'Data Intensive Research': Data Flows in Next Generation Sequencing and Environmental Networked Sensors." *International Journal of Digital Curation* 7 (1): 81–94. https://doi.org/10.2218/ijdc.v7i1.216.
Metzler, I., L.-M. Ferent, and U. Felt. 2023. "On Samples, Data, and Their Mobility in Biobanking: How Imagined Travels Help to Relate Samples and Data." *Big Data & Society* 10 (1): 1–13. https://doi.org/10.1177/20539517231158635.
Millerand, F., and G. C. Bowker. 2009. "Metadata Standards: Trajectories and Enactment in the Life of an Ontology." In *Standards and Their Stories: How Quantifying, Classifying and Formalizing Practices Shape Everyday Life*, edited by S. L. Star and M. Lampland, 149–167. Ithaca, NY: Cornell University Press.
Morgan, M. S. 2011. "Travelling Facts." In *How Well Do Facts Travel? The Dissemination of Reliable Knowledge*, edited by P. Howlett and M. S. Morgan, 3–39. Cambridge: Cambridge University Press.
Nadim, T. 2016. "Data Labours: How the Sequence Databases GenBank and EMBL-Bank Make Data." *Science as Culture* 25 (4): 496–519. https://doi.org/10.1080/09505431.2016.1189894.
Payne, R. 2014. "Frictionless Sharing and Digital Promiscuity." *Communication and Critical/Cultural Studies* 11 (2): 85–102. https://doi.org/10.1080/14791420.2013.873942.
Pelizza, A. 2016. "Disciplining Change, Displacing Frictions: Two Structural Dimensions of Digital Circulation Across Land Registry Database Integration." *Tecnoscienza: Italian Journal of Science and Technology Studies* 7 (2): 35–60.
Pink, S., M. Ruckenstein, R. Willim, and M. Duque. 2018. "Broken Data: Conceptualising Data in an Emerging World." *Big Data & Society* 5 (1): 1–13. https://doi.org/10.1177/2053951717753228.
Plantin, J.-C. 2019. "Data Cleaners for Pristine Datasets: Visibility and Invisibility of Data Processors in Social Science." *Science, Technology & Human Values* 44 (1): 52–73. https://doi.org/10.1177/0162243918781268.
Poirier, L. 2022. "'Accountable Data: The Politics and Pragmatics of Disclosure Datasets', *FAccT '22*." In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 1446–1456. New York: ACM.
Rast, J. 2012. "Why History (Still) Matters: Time and Temporality in Urban Political Analysis." *Urban Affairs Review* 48 (1): 3–36. https://doi.org/10.1177/1078087411418178.
Ruppert, E., and S. Scheel. 2021. "Introduction: The Politics of Making up a European People." In *Data Practices: Making Up a European People*, edited by E. Ruppert and S. Scheel, 1–28. London: Goldsmiths Press.
Sadowski, J. 2019. "When Data is Capital: Datafication, Accumulation, and Extraction." *Big Data & Society* 5 (1): 1–12. https://doi.org/10.1177/2053951718820549.

Sheller, M., and J. Urry (eds.). 2006. *Mobile Technologies of the City*. London: Routledge.
Stehlin, J., M. Hodson, and A. McMeekin. 2020. "Platform Mobilities and the Production of Urban Space: Toward a Typology of Platformization Trajectories." *Environment and Planning A: Economy and Space* 52 (7): 1250–1268. https://doi.org/10.1177/0308518X19896801.
Tarantino, M. 2020. "Navigating a Datascape: Challenges in Automating Environmental Data Disclosure in China." *Journal of Environmental Planning and Management* 63 (1): 67–86. https://doi.org/10.1080/09640568.2019.1659132.
Thylstrup, N. B. 2022. "The Ethics and Politics of Data Sets in the Age of Machine Learning: Deleting Traces and Encountering Remains." *Media, Culture & Society* 44 (4): 655–671. https://doi.org/10.1177/01634437211060226.
van Dijck, J. 2014. "Datafication, Dataism and Dataveillance: Big Data Between Scientific Paradigm and Ideology." *Surveillance & Society* 12 (2): 197–208. https://doi.org/10.24908/ss.v12i2.4776.
van Schalkwyk, F., M. Willmers, and M. McNaughton. 2016. "Viscous Open Data: The Roles of Intermediaries in an Open Data Ecosystem." *Information Technology for Development* 22 (sup1): 68–83. https://doi.org/10.1080/02681102.2015.1081868.
Vertesi, J. 2014. "Seamful Spaces: Heterogeneous Infrastructures in Interaction." *Science, Technology & Human Values* 39 (2): 264–284. https://doi.org/10.1177/0162243913516012.
Weltevrede, E., and F. Jansen. 2019. "Infrastructures of Intimate Data: Mapping the Inbound and Outbound Data Flows of Dating Apps." *Computational Culture* 7. http://computationalculture.net/infrastructures-of-intimate-data-mapping-the-inbound-andoutbound-data-flows-of-dating-apps/.
White, J. M. 2017. "Following Data Threads." In *Data and the City*, edited by R. Kitchin, T. P. Lauriault, and G. McArdle, 85–97. London: Routledge.