



FICTIONS, FRICTIONS & EVICTIONS

HOUSING DATA STORIES

ROB KITCHIN

FICTIONS, FRICTIONS & EVICTIONS

HOUSING DATA STORIES

ROB KITCHIN

Copyright © 2026 by Rob Kitchen

This is a work of fiction. Names, characters, places, and incidents are either the product of the author's imagination or are used fictitiously.

ISBN: 978-1-0683797-1-0

Data Stories project

Maynooth University Social Sciences Institute

<https://datastories.maynoothuniversity.ie/>



**Maynooth
University**
National University
of Ireland Maynooth

**Social
Sciences
Institute**



European Research Council
Established by the European Commission

Introduction

Housing, property and planning data are a key to how places are understood and managed. They inform government policy, shape public perception, and guide billions of euros of investment relating to land use and development, public and private housing, homelessness, commercial real estate, and infrastructure. Yet little attention has been paid to such data beyond some technical considerations related to data access, coverage, and quality, with an absence of more critical analyses of the social and political nature of data production, circulation and use.

Since August 2022, the Data Stories project has been addressing this lacuna by examining the nature, operation and effects of housing, property planning data regimes across state, business, and civil society actors. To undertake the research, the project team have employed traditional social science methods, as well as using arts-based methods in collaboration with seven professional artists and creative writers, to deliver a core objective: to tell informative and insightful stories *about* and *with* data.

Some of these stories are scientific, seeking to contribute to academic and policy debates. Others are more creative and imaginative, using fictional narratives or artistic media to illuminate and reflect on the same issues and phenomena.

This collection of drabbles – stories of exactly 100 words – forms part of the latter. They are inspired by conversations and observations while undertaking empirical research over the past four years, with one drabble penned each week (hence 208 drabbles in the collection). Each drabble is designed to cast a spotlight on a specific data issue.

As the drabbles reveal, the housing, property and planning evidence base in Ireland is less than ideal, suffering from several shortcomings. Such weaknesses are particularly troubling given the extent and effects of the current housing crisis and the need to understand what is happening and to develop evidence-informed responses. Hopefully the stories demonstrate the scope and scale of these data issues and point to what needs to be done to address them.

Acknowledgements

The Data Stories project at Maynooth University has been a collaborative affair and I would like to extend heartfelt thanks to all the team members for their support, camaraderie and observations – the researchers: Dr Leah Aaron, Dr Carla Kayanan, Dr Juliette Davret, Oliver Dawkins, Dr Jennie Day, Dr Danielle Hynes and Dr Sam Mutter; the artists/creative writers: Augustine O'Donoghue, Joan Somers Donnelly, Mel Galley, Dr Ella Harris, Hannah Mumby, Helen Shaw and Sean Borodale; the research support officer: Rhona Bradshaw; the curator: Brendan Fox; and the student interns: Lorena Dias, Kerry-Ann Darcy, Elina Musa, Sabreen Lyman, Leah Fennell Neville and Ellie Hogan. Thanks also to the other administrative staff in Maynooth University Social Sciences Institute for their support: Anne Hamilton Black and Orla Dunne.

The stories in this collection are inspired by interactions with over 300 actors from 105 state, business and civil society organisations. I am very grateful to all those who spoke to us candidly on and off the record.

The cover photo was taken by Samuel Mutter and is used with permission. The cover design was kindly formatted by Leah Aaron.

The Data Stories project is funded by European Research Council (grant no. 101052998). Views and

opinions expressed are however those of the author only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Acronyms

3D: three-dimensional
AHB: Approved Housing Body
AI: Artificial Intelligence
APAS: Agile Planning Application System
API: Application Programming Interface
BCMS: Building Control Management System
BER: Building Energy Rating
CIS: Construction Information Services
CSO: Central Statistics Office
DIY: Do-it-yourself
DoE: Department of Education
ED: Electoral Division
EU: European Union
FOI: Freedom of Information
GDPR: General Data Protection Regulation
GIS: Geographic Information System
HAP: Housing Assistance Payment
HNDA: Housing Need Demand Assessment
ID: identification
IMF: International Monetary Fund
iPlan: a planning application system
IT: Information Technology
LA: Local Authority
LDA: Land Development Agency
LEA: Local Electoral Area
LGMA: Local Government Management Agency
LLM: Large Language Model
NAMA: National Assets Management Agency
NCT: National Car Test
NGO: Non-Governmental Organisation
NOAC: National Oversight and Audit Council

NPF: National Planning Framework
NTMA: National Treasury Management Agency
OECD: Organisation for Economic Co-operation and Development
PASS: Pathway Accommodation & Support System
PDF: Portable Document Format
POWSCCAR: Place of Work, School, College or Childcare - Census of Anonymised Records
PQ: Parliamentary Question
QED: quod erat demonstrandum (the proof is demonstrated)
PRAI: Property Registration Authority of Ireland
RAS: Rental Accommodation Scheme.
RTB: Rental Tenancies Board
SCSI: Society of Chartered Surveyors Ireland
SI243: European Union (Energy Performance of Buildings) Regulations
SILC: Survey on Income and Living Conditions
TD: Teachta Dála (a member of the Dáil, the Irish parliament)
TLC: Tender Loving Care
VAT: Value Added Tax
X: social media platform (formerly Twitter)

Are stories evidence?

'You're obsessed with numbers. As if they're the only evidence worth considering.'

'They're clear and they're factual and they reveal the size and scope of an issue. Difficult to argue with a number.'

'Except we do it all the time – completions, homelessness, dereliction. When is a fact not a fact? When it's vacancy data!'

'Hilarious.'

'The number of people homeless is just a statistic. It tells us nothing about the experience of being homeless. The rental statistics say nothing about being part of Generation Rent. Stories change minds, not numbers.'

'Stories aren't evidence.'

'So, testimony in court is not evidence?'

Surely somebody knows?

'Maybe the Housing Agency might know where we can source the data?'

'We can try, but I'm not holding my breath. My sense is that nobody in government has a good understanding of the whole data terrain. It's all siloed across multiple departments and agencies.'

'Surely someone has it all mapped out?'

'Most agencies don't even have their own data dictionaries mapped out. Or know how their data is transformed into derived data by others.'

'What about a data intermediary?'

'Yeah, they might be worth asking, but I bet it's mostly in their heads and not systematically documented for others.'

Working blind

'There's no state data on commercial property?'

'Uh-huh.'

'Seriously? Nothing on how much, what's vacant, commencements, prices, value, rent?'

'Yep.'

'But surely that's key data for planning and development?'

John laughed. 'Handier for the developers and politicians not to know. Or the public. Evidence can be a real hindrance.'

'So, the planning system is working blind?' Kevin asked, exasperated.

'Let's say partially sighted. Some real-estate companies have some data, if you really need it. For a fee.'

'You had a huge property crash and no-one fixed this? It's like you want it to happen again.'

'Just don't tell the IMF.'

Fairy Tales

'So, what's the initial assessment?'

'The pundits were right; the housing completions data are fairy tales.'

'Fairy tales?'

'Their assessment is we've been over-counting by 58 percent.'

'58 percent!'

'That's what happens when you use a proxy. Just because a unit's been connected to the electricity grid doesn't mean it's finished.'

'But 58 percent?'

'They think the over-estimation is mostly due to unfinished estates, one-off houses, do-it-uppers; units that have been connected, disconnected and reconnected again later on.'

'The Minister's going to have a meltdown. He's been telling everybody that we've been meeting housing targets and thousands have just vanished!'

A wicked problem

'I know you want me to tell you exactly what's going on, but the reality is nobody knows. No commentator. None of your listeners. Some of us are pretty well informed. We've spent years analysing the data and can give you a reasonable sense of things. But that's it.

'Housing's a wicked problem: you try to fix something here, something else breaks over there. It's just so complex. Dozens of actors, loads of legislation and policies, tens of issues such as homelessness, affordability, planning, construction, social housing, finance. Lots of externalities – the economy, global politics.

'And it's constantly in motion.'

Needs a lot of TLC

'You're recording it as vacant?'

'It's clearly empty.'

'It's not even habitable.'

'I'd live in it. It just needs a tidy-up and bit of DIY to make it liveable. A month's work at most.'

'I think we should put it down as derelict.'

'It's not falling down or in such disrepair that it needs partial or full reconstruction. If you were to put this on the market, an estate agent would say it's "in need of modernisation".'

'In need of knocking and starting again. There should be a category titled 'Needs a lot of TLC'.

'Let's leave it as vacant.'

Simpler days

'I miss the simpler days when housing stats were easy. You'd look out the window, count the cranes, note they'd increased, and declare that loads of stuff was being built and all was well with the world.'

'All those cranes led to the property crash!'

'As I said simpler days.'

'And now?'

'You need evidence for any little decision. I used to be able to walk into a bank and an hour later leave with 30 million euro. Now, it takes several months and a business plan stacked with data and stressed tested to within an inch of its life.'

Secondary job

'There's no point comparing us to the CSO or Tailte Éireann. They're professional data organisations. Their primary job is to produce, manage and publish high quality data. That's why they're so good at it. Our job is to provide services to citizens.'

'But you generate massive amounts of data doing that.'

'Yes, but it's a byproduct of what we do. Our staff are planners or project managers or community liaison officers or whatever. None of them are trained data or IT specialists. You can't expect the same processes and outcomes. Well, maybe you can, but it's a completely unrealistic expectation.'

Juking the stats

‘Something’s not right with these data. 1,000 unfinished estates were completed and occupied in a year?’

‘Someone’s been juking the stats. Has to be.’

‘You check through the spreadsheet; I’ll take a look at the report.’

Ten minutes later, Sarah pointed at her screen. ‘It’s in the small print, buried in a note in an appendix. They’ve changed the definition of an unfinished estate.’

‘They’ve done what?’

‘They’ve altered the parameters. They’ve removed the vacancy measure and changed the percentage of units under-construction.’

‘And magically a big chunk of the problem disappears.’

‘Unless you’re a poor sod living on one.’

Intuition and belief

'There's no way that's an accurate picture of what's going on.'

'Why not?'

'Because the numbers feel ... wrong.'

'They're the official statistics.'

'And I don't believe them. Thirty years of experience is telling me they're off. And not by a small amount. If you want to invest based on those numbers, go ahead, but count me out.'

'You trust your gut more than the data?'

'I always trust my gut, unless there's hard data that proves it's misreading the situation. But that's not hard data.'

'It's only your gut telling you that.'

'Intuition and belief, never under-estimate them, Liam.'

A stopped clock

'Surely a fact is a fact? How can four different official sources have such divergent values? The census says there are 166,752 vacant units, Geodirectory that there are 86,708, Revenue say it's 57,206, and VacantHomes.ie 3,460.'

'You're never going to pin this down, they've been arguing about it for years.'

'So, none of them are right?'

'Well, a stopped clock is right twice a day.'

'Meaning?'

'Based on how they're defining and measuring it, the data might have been right at the time it was collected.'

'But who is producing the best measure?'

'Your guess is as good as mine.'

Additive but never subtractive

‘My sense is the number of planning permissions is hugely over-inflated. They’re double and triple counting them.’

‘Double counting?’

‘They’re just adding up all the permissions in the system not realising that a single site might have multiple planning permissions. For example, a landowner might apply for three permissions to build: 100 houses; 100 houses plus garages; or 40 houses and 60 apartments. Whoever buys the site can then pick the option they want. That’s 300 permissions, but only 100 will get built.’

‘The calculation is additive but never subtractive?’

‘That’s my hunch. The true number’s probably half the estimate.’

Just grin and bear it

‘Next time, anything related to BCMS, say “no”.’

‘It’s just one bit of the project.’

‘The bit that’s going to take the longest and be the most frustrating. Just factor in buying the cleaned data from CIS.’

‘That would have inflated the budget.’

‘But the dozens of hours I’m going to spend cleaning, wrangling and making it useable is free labour?’

‘It’s not that bad.’

‘It’s the dirtiest database we use, Paul. Self-reporting with open text fields. And it not being linked with the planning permission database is nuts.’

‘Just grin and bear it.’

‘As if I have a choice.’

Clueless

Sarah and Hannah hurried out of the EU building.

'Jesus, that was embarrassing.'

'I know, I kept sliding down my seat hoping that the Chair would skip over us.'

'Unless someone can provide us with some data, I'm not coming next time. It's embarrassing having to say that we don't know. We look like we're clueless.'

'We are clueless.'

'No, we're data-less. We know what's going on, but we've no way of proving it. Unlike the Dutch and Danes who've got data coming out their ears.'

'Even a table or graph would help.'

'What would help right now is wine!'

A confetti of facts and figures

'Sean, have you got that analysis ready yet?'

'I'm working on it.'

'You've been working on it for weeks.'

'Every time I think I'm nearly finished another report is published filled with a confetti of facts and figures about house prices, or rents, or vacancy, or build cost, or whatever, and conclusions that contradict a different organisation's findings.'

'You need to draw a line under it and submit it.'

'Okay, but I still need to work out what it all means.'

'We know what it means, Sean. What you need to do is find the evidence that backs that up.'

What do we need to know?

‘What I can’t figure out is the logic as to what’s collected. There’s some stats about housing, but practically nothing about land or commercial property. Has anyone ever sat down and worked out what we really need to know?’

Conor laughed. ‘Not in my time.’

‘Then why are some things compiled, but not others?’

‘Most probably because someone requested it – the Dáil, the EU, FOI – probably several times and we’ve decided to collect it regularly.’

‘So, a review of what we should collect is probably a sensible idea?’

‘You’re writing the new data strategy, but ...’

‘It’s unlikely to happen?’

Public but not open data

'This is driving me nuts, why would anybody publish data in a locked-down PDF?'

'You know why.'

'And why only publish this month's figures? They could publish a spreadsheet with the full time-series. Instead, we have to re-type this month's data into our own spreadsheet.'

'Hannah, you say ...'

'It's like it's designed to piss people off, as if the data doesn't itself. There's this many people homeless and by-the-way, we're releasing the data in the shittiest way possible.'

'Better than them not being released at all.'

'Which is what they'd like. As if that would make the problem disappear.'

Transparent government

'You need to talk to the housing staff in local authorities. They'll be able to tell you about the systems they use and what's recorded.'

'But how does the Department influence what data they collect?'

'The systems are procured and configured by them. They can explain that best. You could also try the LGMA; they oversee their data work.'

'You're in charge of housing data policy and you publish housing statistics. You direct the local authorities and compel them to give you certain data. You're really ...'

'I'm sorry, but we can't discuss this without the permission of the Minister.'

Constant tinkering

'It's got to be deliberate, right?'

Jackie nodded. 'Run through it again for me.'

'So, NOAC uses 18 indicators for measuring and tracking housing delivery in each local authority. Over the past six years, only nine of them have been measured in all years, and of those only four have not had their methodology altered.'

'There's zero stability in the measures?'

'Uh-huh. They're constantly being tinkered with. Most are almost the same, but they're not perfectly like-for-like.'

'Do they give a reason?'

'They don't even tell you. You only know by comparing each annual report and reading the small print.'

They exist because we measure them?

'You're saying that unfinished estates don't exist?'

'No,' Jane said patiently, 'but how we understand their existence depends on how you categorise them and the data generated about them. Unfinished estates exist through their definition and the decisions made by their measurer.'

'Seriously?' Tom snapped. 'They exist because we measure them?'

'How we measure them defines them; what is and isn't an unfinished estate. Is a housing estate with fifty percent of units under-construction unfinished?'

'Of course.'

'How about fifteen percent?'

'No.'

'Why not?'

'Because it's near completion.'

'So, where's the threshold?'

'Thirty percent.'

'Why not twenty percent? Or forty?'

A kind of digital twin

'It's a kind of digital twin; a photorealistic, 3D landscape in which you can display administrative and real-time data and traverse through it.'

'Cool. What it's showing here, building use?'

'Yeah, and if you click on a building it pulls up an information panel that tells you more about it. And we can drop in other related data as well, such as property price, rateable value, BER, whether it's an Airbnb property, or it's had an illegal eviction.'

'So, anything property-related?'

'We wish. Very little data is available at property or street-level. It's mostly aggregated at LEA or county level.'

At least it's every five years

'So, what you're telling me is that we only measure this every five years?'

'Through the census.'

'And we've no way of knowing the extent of hidden homelessness, or changes to household size or tenure type?'

'Homeless figures are just PASS registrations and until last year the RTB only recorded new private rental tenures, not on-going ones.'

'And a lot can change in five years, such as 100,000 Ukrainians turning up.'

'At least our census is every five years and not ten like the UK, or never like Germany.'

'The "we're not as bad as" argument doesn't provide answers, Sarah.'

Is a fact, a fact?

‘It’s a fact that the number of commencements is growing.’

‘First, the commencements data aren’t accepted facts; a number of experts have contested how they are counted. Second, even if a fact is a fact – it truly represents the present state of play – that doesn’t mean that everyone will interpret it in the same way. One political party will use it to say we’re on track to meet construction targets, another that we’re failing to meet demand.’

‘So, we should ignore facts?’

‘You’re not listening. What I’m saying is you need to be mindful of how facts can get distorted.’

Redundant question

'I think you can remove the question – when was your house, flat or apartment first built? – to make space on the form.'

'It's to track the age of building stock.'

'Do you know when your house was built? Does a tenant? I bet it's a noisy variable. If we compared that measure for estates, it would vary even though the houses were built at the same time.'

'It's not captured within other administrative data though.'

'Except Geodirectory has been recording new units quarterly since 1998 and every census since has asked the same question. Do houses change their build date?'

Two birds, one dataset

'I think we can kill two birds with one stone: improve housing quality and get some decent data on housing conditions.'

'Go on, I'll bite, what's your big idea?'

'Introduce a NCT for social and private rented housing. Every two years the house is inspected and if it doesn't meet minimum standards the landlord faces penalties. The inspection reports can be used to compile a national housing conditions survey.'

'But only for rented properties.'

'Better than no survey at all. And it'll improve tenants' lives.'

'Do you really think politicians are going to introduce this? Most of them are landlords.'

Crisis repurposing

'This is going to be a bigger crisis than it already is if we don't find a solution quickly. There's going to be thousands of Ukrainians arriving here with nowhere to live.'

'It might be possible to repurpose the homeless PASS system. It's designed to record household details and find those registering accommodation within 24 hours. We enrol them as they arrive.'

'Will that work?'

'Apparently. And we can link it up with the DoE's school's database. We're going to need to locate families in communities where there are available places. Or at least spread the kids out across schools.'

Shield

'GDPR has become the bulwark against open data. If they want to hide the data because of what it reveals, then they cite GDPR as the reason.'

'Even if you FOI it?'

'Especially if we FOI it! They won't even consider supplying it in a de-identified or aggregated form.'

'It's a suppression tactic?'

Julie laughed. 'GDPR is the get out of jail card! It's meant to protect the public from harms arising from nefarious use by third parties, but it does nothing to protect them against harms arising from government use or hiding what's really going on. It's a shield.'

Under controlled circumstances

‘So, this is the CSO’s sandbox. It’s a controlled space in which you can play around with data and run various scenarios and analysis. I want you to use POWSCCAR data to model what future housing demand in places outside the cities might mean for commuting.’

‘And it can only be performed in the sandbox?’

‘Given the sensitivity of the data, individual household and workplace records, they want to keep track of every calculation we do. All we can extract are aggregated models, and only if they meet their confidentiality thresholds.’

‘That’s ...’

‘It’s that or no access at all.’

Deserving and undeserving homeless

'Clara, can I have a word? It's about data entry on the PASS system.'

'Sure, what's up?'

'There appears to be an issue with how some staff are recording details as people enter the system; personal views are creeping in.'

'They're not following the protocol?'

'It seems as if some are making a judgement as to whether someone is a deserving or undeserving applicant, which then affects how they're treated and what accommodation is offered.'

'They're prioritising those losing homes due to rent increases?'

'That's my sense.'

'We need to refresh our training?'

'If we're going to treat everybody fairly.'

Special cuts

'I've asked the CSO for a special cut.'

'Do you think they'll do it for us?'

'Depends on whether they think the analysis is worth the effort of reclassifying the data. Last year they rejigged mortgage data into different age classifications, and before that we had an anonymous cut of SILC microdata.'

'Do you think they might give us a cut of the under-occupancy data at small area level?'

'I'd say more likely ED. If you can give them a good reason.'

'You'd think it's obvious.'

'That might be so, but they've got a lot of demand and fixed resources.'

What's coming down the tracks locally

'That's the largest spatial scale for planning permissions? Local authority level?'

'It's the same for most datasets. In some cases, it's just Dublin, not-Dublin, as with the residential property price index. It's only really census data below the sub-local authority scale. Some they do at LEA scale, which are the constituencies for council elections.'

'But for our submission we need to know what's going on at town or neighbourhood scale.'

'You just have to use the census and what LEA data there is.'

'But that doesn't include planning data?'

'No.'

'Meaning we've no idea what's coming down the tracks locally.'

Buying back data

‘Wait a minute. They’ve scraped the data from local authority websites and now we’re buying our own data back?’

‘It’s not quite that simple, councillor. The company is doing a lot of cleaning and validation, joining datasets together so you can track the development pipeline, including who is involved, and providing a suite of visualisation tools. It takes time and resources.’

‘But why the hell can’t the local authority staff do it?’

‘Because it’s not easy to do and they don’t have the time or resources.’

‘But buying our own data.’

‘Think of it as buying a service. Like consultancy.’

Wipeout

'You got a minute?'

'What's up?'

'I got an email yesterday from someone in the Department. It said they couldn't find a housing finance spreadsheet from 2009 and they wanted to know if we had a copy.'

'They don't have copy of their own data?'

'Seems it was lost when they transferred to the gov.ie domain. It appears they were using their old website as a de facto archive.'

'And when they wiped the site ...'

John clicked his tongue.

'And they didn't make a backup?'

'Seems not. I've had a search, but no luck.'

'Do you have a filename?'

Data dump

‘It’s a data dump. A jumble of different data uploaded at various points that lack time series and metadata.’

‘But the EU has us ranked first in Europe for open data.’

‘And you believe that? When you search data.gov.ie for housing the first two pages are data about fish and animals! And the sites clogged up with zombie data that’s never been downloaded and never will be.’

‘We must be doing something right.’

‘Well, it’s a lot better that it used to be. Twenty years ago, practically nothing was open. Problem is a lot of potentially useful stuff still isn’t.’

Published then gone

'The strange thing is the government doesn't archive its own reports and policy documents. The Oireachtas archives legislation and directives, Merrion Street press releases, everything else just vanishes into the ether, especially anything digital.'

'I'm not sure ...'

'Search for a PDF of a policy document from ten years ago. It's most likely gone. The only way you'll find it is on the Wayback Machine. It's a disaster for future historians. A hundred years from now nothing will remain – documents, emails, databases, websites. Every official I've spoken to is not only unaware of the problem; they don't care about it.'

It's politics all the way down

'Look, the data is a support; it backs up what you want to say. It doesn't actually drive the policy.'

'But I thought that was the whole point of evidence-based policy? The data drives response.'

'Evidence-informed. It informs, not dictates. Data helps us understand the issue, but it's just context.'

'So, you cherry-pick the bits that suit your argument?'

'No, no, hang-on. Don't get the wrong idea. We have to balance many factors when creating policy – public sentiment, market conditions, private interests, existing legislation and regulations.'

'Politics trumps evidence and analysis?'

'This is government, it's politics all the way down.'

Meeting targets

'Yet again, the government have failed to meet the housing construction target. It's become an annual mark of shame that their policy has further deepened the housing crisis. We call on ...'

'The figures released this morning clearly show that we not only met, but exceeded, our target and that our policies are working. As usual, we're getting a distorted view of reality by ...'

'Except, the figures include houses bought rather than constructed, Minister, and the target is well short of what's actually needed. And the homeless numbers continue to grow and first-time buyers are locked out the market.'

Gatekeeper

'We can't give you that data without a data sharing agreement.'

'I'm not looking for personal information; I want aggregated data, preferably at ED or LEA scale.'

'We still need an agreement as we want to be able to control how our data are used.'

'You want to be able to control what can be said with state data?'

'We want to make sure that it's not used for an unintended purpose.'

'That goes against the spirit of open data.'

'It's not open data.'

'So, what's the process of setting up an agreement?'

'We only sign them with approved bodies.'

The data working for us

‘What do I want? To reorganise around the data. It should be working for us; instead, we’re working for it and we’re getting little in return.’

‘Look, Steve, that’s unlikely to happen. We’ve barely got our heads above water as it is.’

‘Yet these systems could be our life raft. They’re full of intelligence that we can use to work smarter. Instead, we treat data work as a necessary burden. We’ve never used the data to consider and improve our work processes.’

‘I hear you, Steve, but ...’

‘Do you think Amazon ignores its own data? Any successful business does?’

Firefighting

'I know it's easy to blame the planners.'

'I'm not blaming them, I'm just saying it would be good to have better planning data.'

'But they're the people who produce the data by processing planning applications.'

'Yes, well ...'

'What you need to remember is local authority planners are under massive pressure. They're under-staffed. They've got tight, defined deadlines. The legislation and procedures are constantly altering. The IT used are decision support, not data, systems. The Department is constantly sending them one-off data requests, plus the stream of FOIs. They're operating constantly in firefighting mode.'

'I get that, but ...'

Tacit know-how

'You can get so far with knowledge and skills gained through education and training, but you can't beat tacit know-how. Take Sean. He's straight out of college. He can perform basic analytics and produce an okay dashboard. Susan, however, she has a feel for the data. She intuitively knows how best to handle and make sense of it. You can't teach that. It's talent and years of practice. Three Sean's could work together and Susan would still produce better insights. The difficulty is retaining Sean long enough that he becomes Susan-like when the private sector pays double what we do.'

Resisting change

'We're moving to a new system?'

'That's what management have decided.'

'Then they've got to be joking. That means double-jobbing as we continue to use the existing system while also setting up the new one. Plus, we'll all need to be retrained. Where's the time for that coming from?'

'Tim said there might be some pain in the swap over, but the new system will make the work quicker and easier.'

'And you believe him?'

'There's got to be a better way of doing things than the way we're currently working.'

'Well, I'm happy to keep things as they are.'

That's just your opinion

'But this graph highlights that rent levels are stabilising.'

'It shows that they're still going up by three to ten percent per annum, even in rent pressure zones.'

'But, we ...'

'Because you've displayed the graph on a full 0-100 scale it looks like a flat line; you need to zoom in.'

'Oh. Okay, yeah, yeah. That's deceptive.'

'But that's how you published it in your recent report.'

'Right, right.'

'And you also claimed rents were stabilising.'

'Well, that's what our analysis shows.'

'I've just illustrated that's not the case.'

'But that's just your opinion.'

'It's what the data shows.'

Normalising a crisis

'You're just trying to normalise what the data shows to be a national crisis.'

'I'm not saying that there isn't a problem. But any functioning housing system has homeless people. That's just a fact.'

'But it shouldn't have and that should be our aspiration.'

'That's just wishful thinking. Every country has a homelessness issue. We're no different.'

'Ten years ago, our rate was a couple of thousand, now its 14,000. That's a dysfunctional housing system.'

'14,000 out of 1.8 million households is a tiny proportion and most are in-and-out of homelessness within months.'

'That doesn't make it a normal situation!'

Neither mandatory or compulsory

'When were Eircodes introduced again?'

'2015. Is this going to be another rant?'

'Nearly a decade later and government agencies still don't have it as a mandatory field on all forms.'

'Uh-huh.'

'It should be a base register. Enter the Eircode and the rest of the address fields populate automatically. It would solve the partial and indecipherable addresses people enter and make geolocating the data easy.'

'You do know that we make a lot of money georeferencing shitty government data, right?'

'They should be asking us to map what's going on where. That's where the real value could be made.'

The answer's simple

'The answer's simple – there's 167,000 empty properties around the country and another 66,000 holiday homes – make them available for the homeless.'

The audience clapped.

'Except everybody knows there aren't 167,000 vacant units! Tom does a good line in weaponizing the data, but it's a fantasy. Most empty properties in the census were for sale or rent, or were being renovated, or the owner was temporarily working overseas or in hospital or a nursing home. And those long term vacant are in the wrong places, with no public transport.'

'There's still more units vacant in the right places than household's homeless.'

A cheap fix

'I know we're paying you big bucks to review our systems and make recommendations but just keep in mind that we didn't have the luxury of being able to design and build it all in one go. It's been jerry-rigged together over a number of decades by different actors.'

'At least it's a chance for a fresh start.'

'Not exactly. It might not be pretty, but it's functional. It does what it's meant to do, though perhaps not as efficiently as it might. We're looking to reconfigure and streamline.'

'You want a cheap fix?'

'You think your bill is cheap?'

Moving at the speed of bureaucracy

'I know you're new, Alan, and you want to make a good first impression, but you need slow down and read the room.'

'If we can get the review completed and revisions made to the data systems then ...'

'That's what I'm talking about. There are several teams involved, including an inter-departmental group, and you need to bring everyone along with you.'

'But if we ...'

'What I'm saying is, we can't fix everything all at once. And there are competing priorities. It ...'

'But I was hired to ...'

'And it will be fixed, but at the right pace.'

Doing the work the state fails to do

'I just got accused again of stealing the state's data. Can you believe that? It's public data! It's available to scrape from open websites.'

'I told you to just ignore them.'

'If they were doing their job properly, we wouldn't need to trade. We exist because the state provides nothing but individual case files. We take their crap and turn it into valuable insights.'

'I co-founded the company, remember?'

'But, no, we're stealing their data! If we're stealing anything, it's the stuff before it becomes data. We make it data.'

'Stop referring to it as theft. It's accessible, public data.'

Competitive collaboration

'Remember, it's our turn to compile the data, so make sure they give it to you in useable formats. And don't share anything other than office vacancy data.'

'I know.'

'But sniff out how they're pricing stuff in the north docks. Talk to Anders, he used to work here.'

'I'll try my old boss as well.'

'I forgot; you used to do this for them, right?'

'Only on updating the joint database between the six agencies, not the writing the report.'

'I'll help you with report. I want to use it to pitch an op ed to the Irish Times.'

Pundit

'What the hell was that about?'

'What?' Jamie handed the clip-on microphone to a technician.

'You know damn well what. Those numbers were pure bullshit.'

'You do know that this is just theatre, right? It's television, people expect a performance.'

'That's why they shouldn't have a real-estate agent masquerading as a commentator.'

'I think the correct title is pundit. And I make no apologies for being one. They ask me back because I provide some drama.'

'You're a spoofer.'

'Look, I'm deadly serious about the data in my business decisions, but this is just entertainment. Do you want a pint?'

What's going on right now?

'The data's too slow for market actors. I need to know what's going on right now. What's happening with market activity, the development pipeline, mortgage approvals and drawdown.'

'There's no way we can ...'

'Am I going to go ahead as planned, or stall development, or change the mix of apartments – reduce the number of two-bed and increase studios? By the time we get some of the data, market conditions have already shifted.'

'The data has to be compiled, validated and ...'

'There's a lot of money at stake. In the big data age, it should be real-time or weekly.'

Priorities

'Okay, we're all agreed that things have to change. However, we can't do everything at once, we need to set some priorities. So, who wants to start? Claire.'

'My vote's for data standards. The data's too messy and it's taking too long to wrangle. We're creating unnecessary work for ourselves.'

'Tom.'

'Data governance. Put in place a revised management process. That would also cover standards. It's all too ad hoc at present and balls are being dropped.'

'Jenny.'

'Proper resourcing. And external help. There's only so much more for less you can do.'

'Priorities without additional resources only, I'm afraid.'

Data mature

‘Kenny’s facilitating this process. He’s helped implement data maturity frameworks in many organisations.’

‘Thanks, Greg. As we start, I want to emphasise, this works best if you’re brutally honest about the current state of play and your ambitions. The reality is, you’re a data organisation. It’s not a secondary task, it’s core business. You deliver countless services, all of which depend on data. You’ve got how many databases for managing clients, operations, staff and suppliers. You publish open data, implement evidence-based policy, produce official statistics. You’re all data workers. This process is about recognising that and becoming data mature.’

Known unknowns

'It was a known unknown. We knew Airbnb was an issue, but not the extent. It was hiding in plain sight. And that's a problem when short-term rentals are taking stock away from long-term use and distorting rents.'

'So now it's a known known?'

'Well, it's kind of unofficially a known known. The government doesn't collect any data. Or if it does, it's not open. But this crowd called Inside Airbnb scrape the data from their platform and provide some interesting analytics. Here ...'

'Jesus, look at all those dots.'

'6,175 of them. Over half are for a full unit.'

Procurement specs

‘Chloe, another FOI that needs answering.’

‘Get John to do it.’

‘He’s already working on one.’

‘Can we not explain the amount of effort it takes to do these and even then we’re not sure how accurate they are?’

‘You think the Indo cares?’

‘They would if they had to compile one. This would be a five minute job if the system wasn’t a steaming pile of ...’

‘The system does what it is meant to do, guide planning decisions.’

‘And nobody thought that running data reports might a sensible function?’

‘Nobody was probably distracted from procurement specs by FOIs.’

Endless reporting

Kelly dropped into her chair. 'Thank god it's Friday, except I hate Fridays.'

'A couple of hours and you'll be out of here.'

'It's all just so pointless; filling in reports, recording status updates, tracking work practices, completing timesheets. It's endless.'

'The place would fall apart without them.'

'They're a total waste of time. ... Well, you know, not really a waste, but ... You know what I mean. It's so frustrating.'

'I know, but don't skip fields. And try to be accurate. It's me that has to tidy up your mess to make sure all the metrics make sense.'

Moving at the speed of paper

'What the hell's the delay, Brian? I can't see our application on eplanning.'

'We're waiting for the local authority to create the case file, add in the data, and scan in all the paper documents.'

'I thought they'd all moved to online application? That's what we did last time, right?'

'There's still a third of local authorities where you can only submit a planning application by post.'

'Seriously? This is the country where every major IT company has its European headquarters, right?'

'Right.'

'And there's a need to build houses since there's a housing crisis.'

'You're shooting the messenger, Larry.'

Glitchy system

‘Let me guess, firewall settings or they messed with a server?’

‘Some bright spark created a virtual server and broke the API.’

‘And he didn’t consider the services connected to it?’

‘He was probably trying to fix a local issue and forgot we were pulling data from them.’

‘And what about these two?’ Kenny pointed at the map and two counties shaded red. ‘Their data’s been missing for almost two months now.’

‘Someone from IT’s visiting them next week to tweak security settings.’

‘One day this will actually function as a National Planning Applications Database and not a just-about-everywhere database.’

Wishful thinking

'Why's there so much office and retail space being proposed in the new local area plan?'

'Well, councillor, it's to complement the proposed housing, to provide jobs, services, places to shop and eat.'

'But half the retail units we have already are vacant and there seems to be an equal amount of offices empty.'

'That's not quite right.'

'Well, what are the rates then?'

'It's, well ... there's no available data.'

'On commercial vacancy?'

'No state data, that is. Some of the real-estate agencies sell it for Dublin.'

'So, we're proposing millions in supporting investment, on what basis? Wishful thinking?'

An analysis of what though?

'What I'm asking is, do you believe the data?'

'They're the largest housing platform and we've been covering their reports for years.'

'Maybe for selling houses, but it's all changed for rental. The market's so tight that hardly anything is now advertised there. Only the leftover crap. People find places now via word of mouth, or Facebook or Whatsapp groups. Anything advertised is mostly done straight through the letting agents, who all having waiting lists.'

'They still have nationwide coverage and decent analysis.'

'Analysis of what though when it is a very partial sample? Does it really represent the market?'

Personal agendas

‘What is going on with them? They’re like a collapsed soufflé. They were getting to grips with the data problem, now they barely seem to care.’

‘The person driving those projects has moved on. Sideways and upwards.’

‘And it’s all just stalled?’

‘It was his thing. Standardised data nationally and decent data analytics to aid local and regional policy. It won’t all disappear, but unless someone takes it on it’ll wither.’

‘One person can make that much difference?’

‘Unless it’s a core department policy with deliverables, then it’s all about personal agendas and personalities. Someone has to push it forward.’

A real dividend

‘Larry, I know you’re frustrated, but you need to be much more strategic about this. You can’t just point out all the faults and expect change.’

‘When it’s this much of a mess?’

‘Embarrassing folks is a not the solution. They’ll become bitter and resentful and they’ll dig their heels in. And you’ve no power; you’re not their boss.’

‘Then I need to talk to their boss.’

‘Change needs a careful balance of persuasion and compulsion and a dividend.’

‘The dividend is everyone gets better data.’

‘A real dividend, in which better data’s a by-product of making everyone’s lives easier.’

Ground-truthing

'Is anyone validating your data?'

'What, checking it?'

'Yeah, to see if it's actually true? There seems to be some discrepancies relating to some of our towns, so we've had teams out from our housing office trying to verify the records.'

'Well, they should be correct, though it's self-recorded data. We do a sample check where we talk to the main contact.'

'But you don't go to the locations to see?'

'No, no, we wouldn't have the personnel for that. There's just two people in the data team, so ...'

'The dataset has some errors?'

'But not too many. Probably.'

The crap we make others do

‘The amount of time wasted is ridiculous.’

‘And your solution?’

‘Get rid of the crap we make others do. Since the system is clunky to use, the planners are sloppy recording data. A senior planner then spends days fixing gaps and errors and compiling spreadsheets for the CSO. But because the spreadsheets ignore the template and desired data formats, they spend a month wrangling and cleaning the data. The CSO resent the planners, who resent their colleagues, who resent the system designers.’

‘And dealing with the resentment?’

‘Flip it around. Fix the system, improve data entry, and pay it forward.’

Protecting commercial interests

'It was your sector that lobbied for a minimal viable property register to protect commercial interests.'

'Well, that was probably a mistake. They should collect data, not just on price, but also floor area, plot size, number of beds/baths; information that'll help you make sense of property values.'

'And let's you model the market?'

'Exactly. Except the detailed data would only be available to registered sector interests. It would be aggregated to ED level for everyone else.'

'So, individual property details would be hidden from citizens?'

'They'd have enough generalised information for an informed opinion.'

'But you'd have much more?'

Charting the unknown

'I don't think the emails to proscribed bodies sit outside the system; they're managing all communications from within eplan.'

'Are you sure? My recollection is some emails are from personal accounts.'

'Okay, well leave that as a separate box for now.'

'Anything else?'

'There should be a back arrow from the BCMS to the planning application management systems. A developer can go back to retrospectively alter what is being sought.'

'This must be the fiftieth revision to this data flow diagram.'

'And I don't think it'll be the last. I'm not sure anyone actually knows how it all fits together.'

How difficult would it have been?

'I can't believe they don't have basic data standards.'

'This is pretty clean compared with some of the junk I've worked with previously.'

Seriously? Just look at the floor area data. You've got one, two and no decimal places; m², sq.m and no units all. Single values and stuff like 9.2+156+52.47. I mean, what is that? And loads of missing values.'

'At least all the values in the column appear to be about the same thing.'

'How difficult would it have been to enter it in the same way for all properties?'

'When they can pay you to fix it?'

If only it were so simple

'The simplest way to get more planning statistics is probably to request all the compulsory fields from the planning application management system for each local authority.'

'In theory, a great idea; in practice, it won't work.'

'But the planning forms were standardised in 2013.'

'They were, but they've started to drift. The bigger issue is that while there're 65 compulsory fields in iPlan, there're only 21 in APAS and not all of those match iPlan fields.'

'We could reclassify data?'

'Possibly, but data capture in iPlan is mostly open text, whereas it's dropdown menus in Odyssey, and checkboxes in APAS.'

Black and white

'The problem, Ruari, is you take everything presented to you as facts and print them as if they are true.'

'What else am I meant to do? I'm a journalist, I'm reporting based on an interview and press release.'

'Talk to someone who doesn't have skin in the game.'

'You all have skin in the game. You have an agenda like everyone else.'

'Maybe. But if you're using the data regularly, you know how dirty and noisy most data are and to take everything with a grain of salt. That story has made shades of grey into black and white.'

Noise

'I can't believe they're still doing this. A five minute segment on house prices dropping 0.1 percent. 0.1 percent is noise!'

'Not this rant again!'

'They're not comparing equivalent baskets of goods. Last month more units might have been sold in rural than urban areas, or more five bed houses.'

'They should show the standard deviation.'

'Exactly. 0.1 percent, plus or minus 0.4 percent.'

'People understand that, it's how they report political polls.'

'Sinn Fein 21 percent, plus or minus two percent. 0.1% is not indicative of anything, yet you'd think the market was imploding the way he's going on.'

Outside the system

'You mean, not every type of planning application is captured and processed using iPlan?'

'We have to process some special types of application using spreadsheets or Word docs or within a GIS. Section 5s, 35s, 42s, 44s, 57s, or Part 8s or 9s. iPlan doesn't have templates to deal with them. Section 247s as well.'

'So, if you're running data reports from iPlan?'

'We have to then manually add in the special applications' data. It's just a case of working through each file.'

'But the danger is that some data gets overlooked?'

'We try to make sure that doesn't happen.'

Have AI, need a problem to fix

‘How’d the meeting go?’

‘Another tech bro who thinks he’s going to fix the planning system with AI.’

‘Let me guess: he’s never worked in planning.’

‘Never even applied for planning. But he has a LLM that needs to be monetized and planning applications are documents that are assessed against a rule-base. QED. We can get rid of planners and speed-up decisions.’

‘And the non-standardised data, the politics, vested interests, consultative requirements?’

‘Easy. Standardise the data and make decisions solely on the rule-set. Or use existing decisions to train the model to make new ones.’

‘Difficult to spot the flaws.’

Bespoke solutions

'I know they've done it, and their site looks fantastic, and it's easy to use, and it's got a loads of datasets, but it's not a quick job to replicate for you. Unless you ask them to do it.'

'But if they've done it, surely you can?'

'It's really not that simple. Their operation consists of bespoke workflows and processes that handles the different data formats and structures, fixes errors and georeferences records, and parses the data into visualisations. The data's such a mess that you can't just download it, drop it in and its done. It's taken them years.'

But that's a hunch

'You're asking me to speculate. And that's tricky. I know that something's changing, because my business is shifting: apartments sales have slowed and it's flipping to build to rent. But I've no hard data for the whole market. Nobody does. I just know our portfolio. So, a vacuum's created where any old nonsense can be claimed.'

'So, should we start building or not?'

'If I was going to make a bet, I'd say the government will step in and buy or rent the units for social housing. But that's a hunch. I've no evidence to back it up at all.'

Group think

'I'm just sick of the group think. It's like the entire industry are lemmings. We do it this way, because they do it this way, because that's how it's always been done. Data says this, do that. But has anything changed?'

'In the market?'

'To fix the housing crisis.'

'What are you expecting? It's in the interests of developers, estate agents and banks for house prices and rent to rise. The group think works. At least, for them.'

'Until it crashes.'

'But then the government will bail them out and it'll all reset and start again.'

'With zero lessons learned.'

I just want a spreadsheet

'What is it now?'

'Do you think they designed this site to stop the average person from finding out things?'

'What?'

'You know, made it deliberately non-intuitive? I might not be a data analyst, but I'm not exactly stupid either. Yet I'm so bloody confused.'

'What is it that you're trying to do?'

'Download some tables. But it's asking for filters and codes and pivots.'

'But they're what makes that site so useful! You can even build viz or do API queries or use widget functions.'

'As if I'm meant to know what they are. I just want a spreadsheet!'

Chinese walls

'I don't understand why you're pretending that you've not set up an internal team to model urban density. I know that you have because you were the ones who told me you had.'

'I really can't discuss this with you, Lorna.'

'We're also modelling urban density. Don't you think it might be worth at least comparing notes?'

'Look, if it were up to me ...'

'You do understand that I'm a senior analyst for another government department? Does this need to be raised to ministerial level?'

'It's our senior minister's office that's telling us not to discuss it with anybody.'

Survey fatigue

'Just so you know, I'm setting a deadline of next Tuesday, then I'm cutting my losses on this survey.'

'You're still having difficulty getting responses?'

'It's the same as last year. Half our members respond straightaway, then I've begged and tormented the others to try to drag the data out of them, but they're ghosting me.'

'Survey fatigue.'

'That's the polite way of putting it. When they want something from us they expect it yesterday.'

'Just remember to make the sample size clear in the report.'

'And when it's published a dozen of them will complain their data wasn't included.'

Semantic slight-of-hand

‘You’re not going to get anywhere with finance institutions. They’ll give you the runaround, then ignore you. Or they’ll ghost you from the start.’

‘You’re talking from bitter experience?’

‘I spent a couple of years trying to talk to NAMA, the Housing Finance Agency, the Central Bank, NTMA, the Revenue Commissioners, the banks. They’ll tell you that their data is financial or they deal with financial assets and they don’t have anything to do with property or development. Plus their data are commercially sensitive.’

‘But the assets are property.’

‘They’ll perform any kind of semantic slight-of-hand to remain a black-box.’

Handover

'Did the person before me leave a handover document?'

'Handover document?'

'You know, something that would explain the file structure, database design, workflows, timelines?'

'I think there might be something on the share drive.'

'I've looked there and I've searched through the folders on the laptop. It's difficult to work out what's what. And there's multiple versions of the same files. Is there any way to talk to them?'

'I think Jim might have Dave's number.'

'So, it might be possible to meet Dave for a handover meeting?'

'I don't know.'

'Without it, we might be starting again from scratch.'

If we have enough data then ...

'You seem to think that if we have enough data we can fix the housing crisis. And it's nonsense. Data's just a medium. It's just representative of a process, but it does nothing in and of itself.'

'But if we have a highly detailed picture of what is happening, then we can start to work out what to do.'

'Data show outcomes. The number of units, levels of vacancy and homelessness, rent prices, whatever. But it's radically altering our political economy that will make a real difference to housing. Evidence-based solutions applied in the same political economy will change nothing.'

Boring data work

'This job's like painting the Forth Bridge; we just do the same boring work all day, every day. It literally never ends.'

'Hey, sometimes we get to go to meetings.'

'Where all we talk about is our usual work.'

'And occasionally we get to run reports and summarize stuff.'

'Yeah, summarize the data we spend all day scanning and typing into the system.'

'But we get to use the scanner.'

'I mean, there must be a way of automating all of this so we could do something more useful.'

'Unless that makes us redundant.'

'Time for a game of solitaire.'

The data cuts both ways

‘The data can cut both ways. You think that data can drive change. That if we just have the right numbers, we can derive new policy that will address issues.’

‘Evidence-informed policy.’

‘Except the data can also be a shield and block change. If the numbers show some kind of progress or success, then they can become a kind of Praetorian guard against all forms of critique. The state doesn’t even need to listen because the data are telling them everything is fine.’

‘But ...’

‘And if a set of numbers aren’t helpful, then they’ll find a set that are.’

Torturing the data

‘What’s that phrase: if you torture the data enough it’ll confess. This dataset’s been through the mangle so that it’ll say what’s desired.’

‘They’ve doctored the numbers?’

‘Yes, but not directly. The numbers are the same. What’s important here is the method of analysis. They’ve combined four sets of variables to create a composite index. The sneaky bit is how they’ve attributed weights for each variable in the index, giving a low weight to what I’d consider the most important factor.’

‘So that gets downplayed?’

‘It’s basically reduced so it has no effect. And magically a key issue seemingly disappears.’

I'm not joking

'You've been moaning for the past half-an-hour about the data. It's like a mantra here; that your housing data are partial and patchy and weak quality. There's actually a lot of data. It's spread across a lot of agencies, but it's there. And it's pretty good quality and it's comparable across the whole country. You should try compiling housing data across Europe. You're not bad.'

'We're not good either. Or as good as we should be.'

'You're way better than average. Data can always be improved – gaps filled, quality increased – but there's a decent bedrock here. Seriously. I'm not joking.'

Echo chamber of misinformation

'Have you read that report I sent you?'

'On the development pipeline?'

'Yeah, that one. Have you seen the sources they used to back-up their argument? They're nearly all secondary. Rather than going and getting the source data and doing their own analysis, they've quoted op ed pieces in the Irish Times or Indo, or industry reports.'

'Really? Let me just find it and ...'

'As if op eds and industry analysis are neutral sources of information about a highly politicised issue. People make the data say all kinds of things. The report is like an echo chamber of misinformation.'

Data hidden in rhetoric

‘The thing you have to remember is that you’re going to struggle to make local politicians data literate unless they’re already data-minded. Some of them are really interested in data, but mostly they’re not. They’re interested in what their constituents think.’

‘But local area plans and county development plans have to be evidence-informed. You need to be able to ...’

‘Get re-elected. You need to put yourself in the head of a politician. You need to turn the data into a narrative they can sell. Don’t just give them tables, figures and maps, provide them with data hidden in rhetoric.’

Too much overhead

'Hey, Denny, did you see that tender for the Department of Housing?'

'The answer's "no".'

'But ...'

'We're not doing another project where we're going have to generate the data to be able to answer their questions. If there's a dataset available, or they're going to produce the data and then pass it on, then fine. But we don't have the time to be messing about capturing and processing data.'

'But ...'

'And it will be a lot of messing when we've other contracts to deliver. And hiring, training and supervising someone for a short-term contract is too much overhead.'

Diminishing returns

‘Are you still working on geocoding the data?’

‘It’s a never-ending job. I’ve got most of them done, but the remainder are going to take a while. It’s a shitty database.’

‘Just draw a line under it. You’ve already spent more time on it than they’re paying us for.’

‘I’d like to hit a 90 percent match.’

‘Before you started playing with it, it was at 68 percent. The coverage is now way better, especially for rural counties.’

‘I know but ...’

‘Don’t worry, nobody will care. The last ten percent isn’t going to change the overall pattern of data.’

Scavenger

'I kind of see my work as being a detective. Or an explorer or scavenger. All three.' Claire laughed.

'I kind of hunt down as many sources as I can. Then I dig around in them and see what data they have and what I can do with them. There's more data available than people think there is, but it might not be seen as data.'

'What is it then?'

'Take commercial real estate reports. They're like the wild west of data. They're full of a scattergun of facts and figures and I mine them out to construct something usable.'

Performative analytics

'It was just homework assigned by the Department, that's the way the local authorities treated it.'

'The HNDA?'

'Yeah. They were given a spreadsheet and asked to run the analysis and hand it back. Except they hired consultancies to do it because it was too challenging to do it themselves.'

'They cheated on their homework?'

'You know what I mean. They got it done because they had to. Box ticked. Exam passed. A cynic would say that it was designed to give the appearance of doing something constructive about housing, but then the next request arrives and its quickly forgotten.'

Silences are where the real story is

‘Okay, I agree, it’s an interesting report. But it’s based on the usual suite of indicators. The data they want us to see. But try thinking about it in a different way. What’s missing? What numbers don’t they want you to know about?’

‘You think they’re deliberately hiding things from us?’

‘I’d bet my landlord’s assets on it. It’s the gaps that’re interesting. Why’re there no data on those? Why did some stats get discontinued? What did they reveal? Who’s benefitting from us not knowing? The silences are where the real story is. And we have a lot of silences.’

Fighting fire with fire

‘Look, I know things are more nuanced. I get that. But at the same time, I’m not misrepresenting things. I am stating publicly available official statistics. And I’m not being overly simplistic.’

‘But, if you know ...’

‘I’m fighting fire with fire. I’m doing exactly what the lobbyists and market actors are doing. They boldly state this is the way the world is and I’m doing the same. My agenda is to change things. To improve housing for people. I use the stats to do that. And you can’t be wishy-washy doing that. You have to be bold and confident.’

Reconstructing the record

‘How’re you getting on?’

‘Okay. I’ve managed to back to 1991, but the digital records stop in 2011. For the rest I’ve had to go into the library stacks, unearth the paper files, take photos of the tables, go back to the office, type them in, then aggregate them up. It can be a couple of days work for a couple of cells in an Excel sheet.’

‘But you have some kind of time series of homeless data for the past 30 years?’

‘Yeah, cobbled together from different sources. Anything earlier than that will, I think, be guesstimates at best.’

Helping you help me

'You're going to FOI the data at their request?'

'That's not quite right. I want the data and they want to release it, because they know it can be put to good use, but they don't want to get into trouble by publishing it without permission and they're not sure the mandarins would approve them doing so.'

'But if they receive a FOI they'll be compelled to published it?'

'Exactly. And the FOI is their safety net. They can legitimately justify why they've released the dataset. And they're not being caught off-guard by the request as we've discussed it beforehand.'

Mixing categories

'Have you seen the new social housing figures? Something feels off. I think what they've done is recategorized HAP properties as social housing.'

'Really? In which case social housing's been redefined as units owned and rented by local authorities, plus households receiving rent subsidy?'

'Plus RAS and god knows what else.'

'How does that make sense?'

'It makes total sense if you're trying to pull the wool over peoples' eyes. Now, if the government claim they've increased the supply of social housing, you've no idea whether they've built new units or they're contributing towards the rent of privately rented properties.'

You're envious of us?

'Seriously? You're envious of us? Our data is patchy, or non-existent, or less than ideal.'

'You have it better than you think you do. Try to do housing research elsewhere. It can difficult to access data, or if there is data it lacks timeliness or spatial coverage. And the chances of being able to talk to the data producer or get special cuts of data are zero. You're actually quite spoilt.'

'Well, I don't feel spoilt and even if our data is better than elsewhere, we're here, we're studying here, and the data is often insufficient to answer critical questions.'

Face value

'Where did you get this build cost figure?'

'SCSI.'

'And you just accepted it and plugged it into our model?'

'It's the only estimated cost I could find. They provide a breakdown of projected costs related to build and land costs, siteworks, professional fees, levies, VAT.'

'So, it's a generic figure?'

'I guess so. What' wrong? You seem annoyed.'

'We need site specific estimates, preferably based on proven known costs for similar developments in the same area. Do we even know how their figures were calculated? Their underlying data?'

'No, but ...'

'No buts. We need to do it again.'

Missing piece

‘That’s fine, but you’ll only be a bit better off than holding up your finger to the wind.’

‘All I’m asking you to do is create a basic development model.’

‘And all I’m saying is, it’ll definitely be basic. You’ve four main parts to any model: land, materials, labour and profit. The last three are known entities, but we’ve pretty much no idea with respect to land values and price.’

‘But surely you can estimate them?’

‘Based on what? There’s no land price register. And it varies by location, type and zoning. It’s a huge piece missing from any model.’

One-in-six

'So, all the claims that landlords were fleeing the market and we needed incentives to keep them from leaving is nonsense? We've been subsidising ...'

'That's not quite right, small investors with one or two properties have been exiting ...'

'That's not what these figures show. The RTB says there were 246,000 private rented tenancies. The census says 330,000 households are renting. That's a gap of 84,000.'

'The CSO thinks about 50,000 of those are informal lettings; no rent, low rent, sub-lets, granny flats.'

'We're really meant to believe one-in-six rentals are informal and shouldn't be registered with the RTB?'

Vanishing data tools

‘Do you have that land availability map that Simon Coveney published in 2017?’

‘I don’t think so. It was an online interactive map.’

‘They’ve taken it down. Can you believe it? The link now redirects to Housing for All on gov.ie.’

‘Like the rest of the Rebuilding Ireland website. It’s like the policy never existed.’

‘I thought I’d compare it to the 2022 map produced by the LDA.’

‘That’s probably why it’s disappeared.’

‘So much for it being a vital step in the new State Housing Land Management Strategy as part of the NPF. And what ever happened to that?’

Comparing like with like

'Did you factor in the regs at the time?'

'Did I do what?'

'Factor in the different stamp duty rates effective within years.'

'I don't think so. No.'

'Then you need to do that. They're continually changing the stamp duty rules and cutoff thresholds. It makes a difference to full acquisition cost.'

'It's not going to make that much difference though, is it?'

'Not if you're not bothered by accuracy. Or whether the figures make sense when you're comparing across years or countries.'

'I need to do it again?'

'Yes.'

'But where I even find the rates for different years?'

Commercial black hole

'Is that public data?'

'What, footfall?'

'Yeah.'

'No. We buy that. Also planning and flooding data. Traffic data. A few others. We also scrape data, mainly relating to commercial premises. Clients also provide data. Banks, insurance and retail companies about loans, properties, customers, suppliers.'

'But we couldn't access it through you?'

'God, no. The client data is their data. We just make sense of it for them. The bought data, we might be able to give you derived insights, but not the data.'

'But none of this data is used by the state?'

'This stuff is a blackhole to them.'

Wasted labour

'You've just cleaned that database up as well?'

'I'll not give you my view of that steaming pile of crap.'

'It took us forever. And it's probably still got issues.'

'The same. There has to be a more sensible way of sharing data. We've cleaned it. You have. And probably at least twenty other organisations. That's a lot of replicated, wasted labour.'

'We need a repository of cleaned data.'

'Exactly. A site where you can access versions of cleaned data with documentation of the process used. And you can pick which one's most suitable or you most trust.'

Economies of scale

‘Rather than each local authority hiring a consultant to build what will effectively be the same data tool, just build one tool, populate it with all the data and cookie-cut it into 31 pieces.’

‘It’s not that simple. Each local authority wants its own tool and to be in charge of its development.’

‘But they would have their own tool and data. It is pretty much the same cost to create one tool for everyone as it is for one LA. That’s 30 unnecessary contracts.’

‘This isn’t about the tool, or the data, it’s about local autonomy versus centralised control.’

All you're doing

'What are the chances of getting it a full update by the end of the week, Janet? I have to brief the Minister.'

'Only if you give me twenty more hands.'

'Really? All you're doing is pulling the data together.'

'All I'm doing? It's not like I'm just plucking a book off a shelf. It all has to be assembled, cleaned, wrangled, validated and parsed into the dashboard.'

'Surely you have a standardised process for that?'

'I do. But this is mostly being done by hand. It's a slow, tedious process.'

'Can't we automate it?'

'Not without a system overhaul.'

Better than good enough

'The data's good enough, okay? Stop fretting.'

'And you're sure the data and analysis is accurate?'

'Look, a few years ago, you'd have done this by hand, it would have taken forever, and you'd still only have a ten percent sample. That sample would be accurate, but you'd have been estimating for the entire population from that. I've processed the entire document set, extracting and analysing the relevant data, in just a few minutes. Yes, it has some mis-readings, omissions and false positives, but we can tweak the code and iterate.'

'But is it accurate?'

'It's better than good enough.'

Inconvenient truth

‘What you’re asking for, we used to do. People forget that. We used to be pretty good with housing data.’

‘Really? When?’

‘Before the financial crash. There used to be an annual housing statistics bulletin that was quite comprehensive and the Department used to publish loads of spreadsheets.’

‘And then?’

‘Housing data became an inconvenient truth. Very inconvenient.’

‘And they stopped publishing ...’

‘Stopped publishing. Stopped collecting. Withdrew what was already published. Try finding one of those bulletins.’

‘That’s what we’re trying to fix.’

‘And it’s great. My point is, you’re not the first people that’ve tried to do that.’

Selling our own data back to us

‘The state doesn’t fully comprehend the idea of open data. They’re agencies still selling state data. It’s data that’s been generated through taxpayers contributions, but they sell it back to us.’

‘You’re talking about?’

‘PRAI for property records, Geodirectory for building data. Geodirectory’s owned by two state bodies, An Post and Tailte Éireann. The post person records the data for each building – location, build date, occupancy, type, number of floors. Why isn’t that open data? Why isn’t Eircode?’

‘They’ll say they’ve added value to the data.’

‘Pffff. They need to do that for state users anyway. We’ve paid for that.’

Nowhere near as easy as you think it is

‘You’re trying to make a distinction between administrative and operational data. They’re the same thing. The admin data are generated through an operational system. The operation of bureaucracy.’

‘It doesn’t really matter what we call it, our operational data are generated to deliver a service, not for producing statistics. They’re not designed to answer your generalised questions, but to ensure things work as they should, that people are treated correctly.’

‘And I’m saying they can be repurposed.’

‘Honestly, it’s nowhere near as easy as you think it is. You’re asking for a full system redesign.’

‘Surely ...’

‘No, you can’t.’

Even though you work here

'We do have that data, which means I should ...'

'Hang-on. I can't share that data with you.'

'What, why not?'

'Because it's individual household level data. GDPR and SI243 means we can only share it with the homeowner's consent. And I'd also need sign-off from the senior manager team.'

'Seriously? Even though I work here?'

'Do you think anyone in the organisation should be able to see your smart meter data?'

'We need it for our building energy model to estimate the upgrades needed.'

'I'd say you're going to have to make do with area generalised or synthetic data.'

One-off survey

'It was a one-off survey?'

'Yeah, we got some money from the Housing Agency to do it.'

'And you're not going to do it again?'

'We've no plans at present. It was quite a bit of ...'

'That's a shame. The findings were really useful. We refer to your report all the time. It would be a great if it was an annual survey. The problem with one-offs is you can't track change.'

'I agree, but it's better than nothing, and I'd say there's little hope of an annual survey given the cost.'

'The information would be worth it though.'

Pseudo-science

'You set the asking price at 410K?'

'I think that's realistic. It's a seller's market.'

'And that's based on? Key data and a formula?'

'Kind of. It's more ...'

'Is there a model or some science behind it?'

'Not really. I forgot, you're a data scientist.'

'Well, I was just wondering.'

'No, it's fine. It's based on taking a look at the property register for similar properties, and discussing it with other agents. It's kind of evidence-informed mixed with sales experience.'

'So, more than a hunch?'

'For sure. We're constantly tracking trends in property sales.'

'Kind of like a pseudo-science?'

Blinded by industry assumptions

'You're fitting the data to your existing preconceptions and approach, rather than seeing what else it might suggest.'

'The data are what they are. That's how much it costs to build apartments, which is what the government wants in order to create density. The only way that can happen is with subsidies.'

'That's the only thing those data suggest to you?'

'There aren't any other solutions!'

'How about drastically reducing land costs by using state land? Or switching to high density, terraced duplexes that are cheaper to build? Or shifting population growth to other cities? You're blinded by industry assumptions.'

Skin in the game

‘Look, it’s great that they produce that data, but it’s just based on their client base. And they’re just one organisation out of many with a particular kind of tenant. So, what does it really tell us?’

‘The data’s unrepresentative?’

‘It’s useful but you need to interpret it with care.’

‘And that’s an issue with civil society actors?’

‘With all actors. Private industry’s the same. It’s selective and it often suits their policy agenda. That’s why we need independent, official statistics.’

‘But the state’s also an actor.’

‘But they have the coverage and the CSO’s little skin in the game.’

Forecasts and guestimates

'Are we even going to make last year's total?' the Minister asked.

'I don't know. It might be a little bit less unless there's a surge of completions in the next three months.'

'Yet a couple of weeks ago you let me announce that housing completions would top 40,000 for the year. The Taoiseach even boasted about it to the media.'

'That was your special advisor ...'

'I don't care who it was! I'm being accused of deliberately misleading the Dáil.'

'The figures in the first part of the year ...'

'Were just that. I need accurate forecasts, not guestimates!'

Lost opportunities

'You just joined us from the private sector?'

'As a senior data analyst.'

'First impressions?'

'Honestly?'

'Of course.'

'It's all a bit of a mess. We should have the best view of what's going on. Through applications and regulations the state controls the majority of what happens and it captures masses of data, but it's all trapped in a multitude of siloed management systems and databases. What data I can access lacks decent resolution.'

'So, not impressed!'

'If it helps, I'm impressed that we manage to get as much insight as we do when we could do so much more.'

Sub-targets

'We need to be doing more than just saying we need 50 to 60,000 new dwellings a year. It's 50,000 of what that matters.'

'I don't care what they are so long as they add up to more than the target.'

'That's not going to fix things. We need targets for affordable units, social housing, student accommodation, build-to-rent, family homes, sheltered housing, and downsizing. And they need to be set across the country.'

'You're crazy if you think I'm going to set sub-category and area targets. Not meeting one target is bad enough; not meet half-a-dozen will see me sacked!'

Factually incorrect

'The Irish Times are seeking a statement in response to the Minister's comments.'

'Which ones?'

'That the figures on rent, sales and market activity are factually incorrect.'

'Seriously? The man who manipulates and spins the figures to try to hide the scale of the problem is complaining that we're factually incorrect? Talk about pot and kettle.'

'What do I tell them?'

'Tell them the truth; that we're quoting that banking report. Though God knows where they got their figures.'

'I'm assuming the banks.'

'It doesn't matter. Or whether they're correct. What matters is that they show how bad things are.'

What's that per unit?

'Is this remediation scheme going to apply to us?'

'I don't know, dear. It just says up to 100,000 buildings built between 1991 and 2013.'

'Well, that us then.'

'Let's hope so. They're creating a fund of €2.5 billion. What's that? €25,000 per unit? No way that's going to cover fire safety, structural and water ingress issues.'

'100,000 buildings, you said. There could be fifty apartments in a block.'

'Right. So, maybe five to ten grand. How can they not know how many buildings or units it applies to?'

'But it'll fix our ...'

'I wouldn't bank on it, dear.'

Making up numbers

‘Where is she getting these numbers from?’

‘I think she’s just making them up.’

‘Seriously?’

‘It’s an election debate. She’ll say whatever she thinks will sway voters.’

‘But it’s total nonsense!’

‘We both know that but ...’

‘Jesus. She’s making him look like a fool. Go and talk to the producer.’

‘He’s not going to do anything. It’s good radio. Plus, it doesn’t matter anyway. They’ll not have to live with whatever nonsense they’re spouting; it’ll die in any coalition negotiations for the programme for government.’

‘Except they won’t be in any negotiations.’

‘In which case, it definitely won’t matter.’

Questioning the model

'They're saying house prices are over-valued by ten percent and there's a risk of a "painful correction".'

'And you believe them when there's a chronic shortage of supply?'

'The worry's that people are carrying too much mortgage debt; a slight shock to the economy and ...'

'They're basing this on what?'

'Modelling of expected house prices given demographic and economic factors – population, income, access to credit, interest rates.'

'So, modelled data?'

'Real data used in a model.'

'So, basically a guestimate.'

'Not really, no.'

'But not a hard fact either.'

'No.'

'Then we'll defend our position by questioning the model.'

Data hoarder

'I'm pretty certain they must have the data.
That's all they do – collect, collect, collect.
Though they've no idea what to do with all of it
beyond its core use. It's such a waste.'

'They don't share it with anyone?'

'They don't even share it with other units in their
own organisation.'

'Are you sure?'

'I met one of their data analysts at a workshop.
She said they operate like a magnet; data flows
in, nothing flows out.'

'And you've approached them?'

'And been greeted with silence.'

'Can't they be compelled to share?'

Mike laughed. 'Yeah, right; like that'll happen.'

Stability matters

'You don't seem to understand that stability matters, especially when producing official statistics. Yes, the world is always changing and what you're dealing with is flux, but how we measure and make sense of that needs a stable base: fixed measures into fixed units.'

'And we can't maintain that?'

'Not when we're constantly introducing new measures and discontinuing others, altering existing methods and tweaking area boundaries. It's becomes impossible to do long-run analysis and short-term fixes are tricky.'

'So, you're saying you're opposed to the proposed changes?'

'Yes! It'll just mess things up further. Which I guess is the purpose?'

Market intelligence

'So, you're tracking users?'

'Of course. It's useful intelligence. We produce a number of metrics about customers: what properties they're viewing, how many times they repeat view, their preferences, their click-throughs, how often they log-on and spent online. We also have metrics on properties and agents. Plus data from other sources. The CSO, RTB, whatever.'

'And you use it for?'

'Market research and to try to convert clicks into sales.'

'And you sell the data?'

'Not yet. We've been looking into it, but we're not sure there's enough of a market.'

'But you would?'

'Why wouldn't we? We're a business.'

Out of sight

'It's criminal really. Every single planning application is chocked full of data, a huge amount of which is codified in their management systems. The assessment process is also captured. They have masses of individual application data, and what does the public get? A handful of planning stats aggregated to local authority scale.'

'David ...'

'Unless you're willing to buy it from a couple of companies who scrape out the eplanning systems and track it all at building level.'

'Just take a breath and calm down.'

'We're in deep in a housing crisis and key data is locked out of sight!'

Background reform

'I'm just saying that your critique doesn't reveal the whole picture. We are working to fix and improve things. I know that mightn't seem the case from outside the system, but we're gradually addressing many data issues.'

'Then why not publicise that?'

'Because that's not how the civil service works. Our work goes on in the background and the only announcements we make are made or approved by the Minister.'

'Then why doesn't she say something?'

'Because where's the political capital in that? She has to admit that things were a mess and needed fixing. Only new initiatives are trumpeted.'

Embrace the absurdity

'What if we just showed how absurd the situation is?' Mia asked. 'Used their data back against them?'

'Everybody already knows how absurd it is.' Claire sighed. 'Housing's a running joke. And who cares about the data? This is beyond data. It's all around us.'

'Nothing we've done so far has made a blind bit of difference. Nobody cares if we do another bloody march or hold up placards outside the Dáil. We're a broken record.'

'So, we give up?'

'We mock them. Lampoon them. Shame them. We embrace the absurdity. Pantomime. Carnival. Satire.'

'With data?'

'Lies and damn statistics.'

Staying inside the tent

‘Just tell them the consequences of not upgrading the data system.’

‘They know that already – it’ll stay the way it is.’

‘I mean the consequence for tenants. For people who are homeless.’

‘You can never express anything like that in one of these meetings, Val. The minute you do, you’re dismissed as a social justice wonk.’

‘Seriously? There are people who ...’

‘Seriously. They’ll stop inviting me. It’s better to focus on efficiency, return on investment, and doing more with less.’

‘So, ignoring what the systems are meant to do?’

‘Yes, if it’ll will actually help the systems do that.’

Only in theory

'Oh for a system that could tell us what the hell's going on? In real-time and in detail.'

'That's never going to happen, Cleo,' Tom said.

'Never.'

'You do know that you're the people in charge of making that happen, right?' Jack asked. 'That's literally your job.'

Cleo laughed, nervously. 'And it's an impossible task.'

'Well, not impossible,' Tom said, 'but not exactly straightforward.'

'But ...'

'Nobody really wants it fixed, Jack. Not really.'

'Everybody wants it fixed.'

'Everybody who isn't going to get burned by the data. Or pay for the fixes.'

'So it is solvable?'

'Only in theory.'

Off-the-record

'Aoife, can we chat off-the-record?'

'About?'

'The FOI you've submitted to the Department.'

'Is there a problem with it?'

'No, no.'

'So, I'll get the data?'

'That's what I wanted to chat about. We'll give you the data on two conditions.'

'And they are?'

'You don't share it with anyone else and you withdraw the FOI.'

'I only FOIed it because you wouldn't share it with me.'

'And now we will. But we don't want the FOI on the register as it might catch the attention of a journalist. We don't mind helping out an academic we know, but ...'

Public, but not public

'I finally got an response to our query about making that dataset open data.'

'And?'

““While we have not purposely published the data for public consumption, it becomes public through ministerial updates on PQs”.’ Except it has never become public because the Minister has never been asked about the data.’

'So, the only way it'll become public is if we can persuade a TD to formally ask the Minister a question?'

'It seems that way.'

'They're trying to claim they're not hiding the data, while making it all but impossible to access it!'

'So much for open data by default.'

Manufactured ignorance

‘You’re saying we manufacture ignorance?’

‘Mass produce it. We all know there are things we don’t know. We don’t measure them. And we often know why. We don’t want to know the answer, or at least other people to know. They are known unknowns.

‘But we also have unknown unknowns; things we don’t know we do not know. Hidden ignorance. They only become obvious when something goes wrong or someone starts asking questions.

‘Then there are unknown knowns. Things we think we know but have no concrete evidence. Anecdotes. Gut feeling. Tacit knowledge.’

‘But ...’

‘And we ignore our ignorance.’

Do bears shit in the woods?

'This isn't just an oversight, Rachel. And it's not because they can't do it. Or can't afford it. It's deliberate. If there's no data collected about it, then they can't be held to account.'

'You really think it's deliberate?'

'Do bears shit in the woods?'

'Even if you don't have evidence they do, you still know that's the case.'

'But only through conjecture. Which a politician or lawyer will easily deflect or shred. The lack of data is an ethics issue. It's a justice issue.'

'Don't you think that's over-egging it?'

'Tell that to the people who've been illegally evicted.'

AI slop

‘If there was ever an issue to fix with AI it’s planning. It’s rule-based. It’s mostly been digitalised. It’s inefficient and contested. It’s ripe for an AI overhaul.’

‘You think the solution to a crappy system is more shitification?’

‘From applying AI? There’s a mountain of existing training data in the form of previous planning applications. There’s a defined rule-set.’

‘There’s a mountain of partial, unstandardised, crappy data, and the decisions are riddled with subjective choices and politics.’

‘You can use AI to fix that.’

‘Garbage in leads to garbage out. The fix is governance, standards, resources, not AI slop.’

Less-than-open data

'Most state data is not open data. And most of the data made available doesn't meet core principles relating to file format, data quality, discoverability, metadata, address stability.'

'But you can access it?'

'If you can find it. And you have to take it as it comes. Or you can see the data mapped or graphed but not download the data itself. Or you need special permission. And datasets can just disappear or move to a new web address.'

'That's still better than no data at all.'

'But a long way short of the government promise of open by default.'

Data is money

'Do you know why we're a market leader? Data. Data equals knowledge. We know exactly what's going on in the market. Our data analytics team uses all the data we collect and those generated by others, statistically analyses it, and runs it through models. We know how to price things to ensure a sale and to maximize profit. We know what properties to invest in. We know how to minimize our costs. And we know how to package up this knowledge to sell to others. Others might not have realised it yet, but data is money; a form of capital.'

To AI or not

'You know that AI can produce useful tools, right?'

'I know it produces hallucinations.'

'They can be fixed. Saving ninety five percent of time on a task, and only losing five percent to check and fix, still saves ninety percent. And it cuts out subjective judgement.'

'I don't know. It's being rushed in without any testing. Or staff training. Or oversight. There's a scandal waiting to happen. Another Priory Hall.'

'Which can happen again with the existing system. With none of the gains in efficiencies.'

'But with our data shared with god's knows who, doing god's know what with it.'

Omission

'What the surest way of hiding something?'

'To make sure it's behind a secure lock and key.'

'Or?'

'To place it plain sight?'

'Is that how you hide something?'

'I'd put it somewhere nobody will find it.'

'But it would still exist.'

'You're saying the best way to hide something is to destroy it?'

'I'm saying it's to not create it in the first place. If you don't collect data about something then ... That's why there's no state data about this. The only way we can prove our case is produce our own evidence. Which they know is difficult.'

Data sculptures

'Are you okay?'

'What? Yes. I'm just trying to make sense of this.'

'It's a data physicalisation. Digital data turned into a physical form. In this case a sculpture.'

'So, this ...'

'Each layer is sized to the number of houses built each year. The blue wire surrounding each layer is the government target for that year. The red wire's what housing commentators said was needed. The further the wires are from the layer, the greater the shortfall. The sculptures next to it are average rent and house prices.'

'Those shapes are bad, right?'

'That's the scale of the problem.'

It practically built itself

‘What do you think?’ Steve motioned at the interactive dashboard.

‘It’s pretty cool.’

‘It took three days to build.’

‘Three days?’

‘Using some AI tools. They’ve pulled in all the government data on planning, scraped all the planning application files from the local authority planning sites and ingested them. They’ve converted all that into graphs and maps. Using the chatbot function you can ask it about any application, or data and trends across applications.’

‘And that only took three days?’

‘The site practically built itself. And it harvests and builds the database. There’s no excuse not to know any more.’

Can't fight disinformation with guestimates

'Who is supplying these idiots with data? Or are they just making it up?'

'It's all over X. And Grokipedia. It's total garbage.'

'It's being quoted in the media.'

'But with some pushback. And most are distancing themselves from the so-called facts while seeming to be sympathetic to the issue.'

'You'd never know we're a nation of emigrants.'

'They claim their worried about ...'

'How much social housing is occupied by non-nationals?'

'Hardly any. My guess is it's a much smaller percentage share of the total number living here.'

'We can't fight disinformation with guestimates. Get me the actual figure.'

For inside the system only

'You're moaning about our homeless data? I don't get it. It's probably the best data in the OECD. Every agency that receives state funds has to use the PASS system. It's fine-grained, standardised data.'

'That nobody except you can access.'

'It's highly sensitive data.'

'That could be classified, aggregated and shared.'

'We don't have permission for that from those we collected the data from.'

'And you don't need it if it is aggregated data. You're not sharing anyone's information.'

'Even the aggregated data ...'

'No, it couldn't. We supposedly have excellent data, yet nobody can do anything useful with it.'

A hidden treasure trove

‘Explain it to me again.’

‘It’s a little, hidden trove of 151 datasets. All from state sources. You can’t find it with a Google search. You have to know it’s there.’

‘But it’s not open data?’

‘I’d say it’s accidentally public data. Some of it is available elsewhere, but not all of it. It’s what they’ve assembled to build their own internal tools.’

‘And you found it ...’

‘They have an open mapping tool. I wanted to find the data it was pulling in, so I used the browser network panel and traced back up the file structure. And voilà.’

You've no idea how good you've got it

'Honestly, you guys keep moaning about the data, but I'm telling you, it's amongst the best in Europe.'

'You're joking, right?'

'No. My job is to create pan-European datasets. You try sourcing housing data for East Europe. Not a chance for just about everything. Italy is patchy. There can be several years delay in France. Their latest housing completion data is for seven years ago. Spain publishes data, but nobody believes it. You're actually spoiled with how much accessible data you have compared to other countries.'

'Now I know you're joking.'

'Seriously, you've no idea how good you've got it.'

Which is correct?

'Take this site,' Gerry said, pointing to an area of derelict land. 'It used to have a four houses, two converted into offices, and three shops on it. If you look at Google Maps they're still there. OpenStreetMap shows a single space. The Eircode map has five address points. Land Direct shows it is nine plots, each of which used to be separately owned. They remain separate parcels, but in practice they're now consolidated. The planning files and the National Planning Application Database show different site boundaries. And it all varies historically. Which is correct? This is replicated across Dublin.'

A hot mess

‘Look, I know the data’s a hot mess. You know it is. But most people don’t and they take it at face value. Which means they produce analysis that simply doesn’t reflect the actual situation. Which they then translate into policy or decisions that makes things worse, not better. We should share our cleaned up version. It’s not perfect, but it’s a damn sight better than the official dataset.’

‘But we then become responsible for its upkeep and any issues or mis-uses. Can we persuade them to fix it?’

‘They don’t care if it’s good data, just that it’s available.’

Data solar system

'We're to place the things we know most about at the centre?'

'And those we know least about at the edge.'

'And we're to group things into planets?'

'Yeah. So, like planet rent. Planet planning. We're creating a data solar system.'

'It's like being back at school. Colour stickers and marker pens.'

'They're using a metaphorical approach to try to surface things that wouldn't normally get said or is hidden.'

'Still seems a bit childish.'

'Children often see things we don't. Or see them in a different way. Just go with the flow. I'm putting planet demographics in the middle.'

Limited access

'Are those new dashboards and their underlying data accessible? I can't find them on your website.'

'They're just internal tools at present.'

'But you're sharing them inside the system?'

'They're available to county and city managers and other senior local authority staff.'

'But not the rank and file?'

'Not without authorisation.'

'How about the Housing Agency or Land Development Authority? The Office of the Planning Regulator or CSO?'

'Not yet.'

'And they're all as good as Trabis? Which I've seen.'

'They're similar.'

'So, there's detailed data on housing, planning and construction, but nobody knows you have it or its unavailable?'

Well-crafted prompts

‘All of our analysis is now AI-generated. We write a prompt to produce a descriptive analysis of a variable – say, the number and rate of renting – at the local, regional and national scale. Then we check that over and tell it to do the same for all the other variables we’re interested in. The models now are really good, especially if you craft the prompt. We hardly have to tweak the outputs. What it produces is what the councillors or planners want – a high-level overview that compares across places in a standardised way. It’s saving us a ton of time.’

Misalignments

'Do we have good data on unactivated planning permissions?'

'We have data, but it's hard to compile. What gets extracted for the CSO is planning permissions. Even that takes a lot of work as the systems are not designed to run data reports. It all has to be checked by hand. And we don't track permissions downstream; it's difficult and time consuming.'

'Can you not just cross-reference with commencements?'

'In theory, yes, but in practice, no. The systems are using different IDs and there might've been plan modifications. It takes a bit of effort to align; even then it's messy.'

Do we really want to know?

'We don't collect what I want to know about. Culture, identity, place. The experience of living in a community.'

'Can you really measure that? It's all subjective. There's no hard facts. It's all intangibles.'

'Is it though? There's a whole consumer surveying industry built on measuring attitudes, values, opinions and behaviour. Running political polls. Why don't we have that kind of data for neighbourhoods?'

Malcolm shrugged. 'We don't have the in-house capacity, nor the budget to hire new staff.'

'But it would be worth it to know.'

'Do we really want to know if we can't do anything in response?'

Maddening

'The numbers just make me so mad. I'm sitting here all day, every day, processing homeless data. At one level it's tedious work. But it's also heartbreaking. These are real people. Families who, through no fault of their own, find themselves without a home. And everything we're doing is only having a marginal effect. People are becoming homeless faster than we can source new accommodation. And we do move a lot of people out of emergency accommodation every month. I'd hate to think what the cumulative number is of people who've been homeless over the past few years. It's maddening.'

An Irish thing

'I think it's a bit of an Irish thing. We don't like the idea of people collecting data about us. Knowing about us. I guess it's a legacy of colonialism. Of using data to exercise power. And that hangs over not just what citizens do, but also the state. We tend to collect the bare minimum; enough to fulfil a certain task, but no more. Or we make do with what we've got. Public servants are citizens after all; they're enrolled in the same systems.'

'So, the data regime's cultural?'

'And social. I imagine it's similar in other postcolonial states.'

Fiefdoms and siloes

'My dream is to get everything in the organisation into a single GIS that's well designed and organised. Make something that's more than the sum of its parts. But that's not an easy nut to crack because of all the different structures and systems, as well as all the personalities and teams. The place is full of fiefdom's and silos. Mixing my metaphors, but you get the picture. Everyone wants to do it their way and nobody wants to share. They'd all benefit from a single, comprehensive, standardised system, but nobody wants to give up control of their own mess.'

Aha! moments

'I work with the data everyday so I'm kind of inured to its power; you know, it's power to reveal things. But that's not the case for most people. They're often surprised by our maps, graphs or infographics. When I show the data at townhall meetings for the local area plans people are always fascinated and shocked by some of the data. It shows them what's really going on; the scale of an issue.'

'If they're a local don't they already ...'

'They might have an idea that the issue exists, but little else. The data leads to 'aha!' moments.'

Rabbit holes

'You can really go down a rabbit hole with this, searching for records. Loads of properties lack official documents. Even the department who's in charge of housing can't tell me who owns the building they occupy. Nor can anyone else. It could be them. Could also be Britian since they built and then occupied the Customs House. It could be someone else. It's the same with train stations, museums, barracks, monuments, public parks. Basically anything historic. Who owns these sites? Nobody knows. Which means my dataset isn't complete. Instead it's riddled with holes and uncertainty. And it's driving me mad!'

Data won't fix things

'You can have all the data in the world and it won't fix the problem. We already know what the problem is. We're not building enough units. Rent and purchase prices are too high. Too many people are becoming homeless.'

'So, we should stop trying to get more and better data?'

'Don't get me wrong; that's not what I'm saying. We absolutely need better data. But better data won't magically fix things. It just helps illuminate what's going on. What'll fix things is active intervention. The reality is, data doesn't lead to action. What does is fear of losing power.'

Money and power talks?

‘But is planning really data-driven? It’s political, no? The councillors and the developers have got it all worked out.’

‘Anyone who says the data isn’t front and centre doesn’t know what they’re talking about. We use data all the time; to inform land-use zoning, to make decisions on planning permissions, to formulate and monitor policy. Sure a specific development might have political support, but if the numbers don’t add up, it isn’t happening. Someone will take a judicial review.’

‘Come-on, Mike. Let’s be realistic. Money and power talks.’

‘And evidence speaks truth to power.’

‘Always?’

‘Nearly always. Ninety nine percent.’

We need a tool for that

'You don't need a tool to tell you what the problems in a neighbourhood are, or how they've developed over time. Everyone knows that.'

Audrey sighed. 'But you do need a tool if you want to model what'll happen in the future and to run scenarios. What'll occur if we build this mix of housing? Or don't create a balanced community? Or don't provide sufficient recreation facilities?'

'Do you though? Isn't kind of obvious?'

'Then why are things how they are? A standardised tool is faster and more robust and realistic. And you can understand its estimates. Unlike armchair predictions.'

Predicting markets

'It's a sure bet though, isn't it? Demand is exceeding supply. You build it and they'll come.'

'That's only true for houses if a buyer can afford it. Which a bank decides, not them. For offices and retail, demand is more uncertain. Work from home, economic outlook ...'

'But ...'

'Despite what you think, developing a site is a risky business. You've this giant financial calculation to perform; a million costs to predict – land, materials, services, transport, labour, energy, insurance – all of which can change. Once finished, will it sell? If our data and forecasts are wrong, we go bust.'

Build it and they will come?

'You've zoned this area for retail?'

'We're re-activating its zoning from 2006.'

'And you think they'll be sufficient demand?'

'The councillors seem to think ...'

'What's the existing retail vacancy rate?'

'There's no data on that.'

'But the shopping centre here,' Sean pointed at the map, 'is half empty.'

'It's on the other side of the town.'

'And on the main street?'

'There's a few.'

'Yet we're zoning for more retail development.'

'The plan is to create economic growth.'

'By emulating the over-development of pre-crash? Just because there's no hard data doesn't mean we can just ignore what's plainly obvious.'

Necessary not just useful

'I agree, it would be great to know more about housing conditions, but there's no existing source for that.'

'So, let's conduct a new annual survey.'

'Have you any idea how expensive and time-consuming it is to run a representative, nationwide survey? What would be the pay-off?'

'We'd know about the material conditions of our housing stock.'

'And? What action is going to be facilitated by it? It'd be better if the data were the exhaust from an inspection that has consequence.'

'But ...'

'If you want that data then you're going to have to convince people it's necessary, not just useful.'

Sound but limited

'Can't we just compel the real estate agents to give us their data? They're compiling it already.'

'Technically, yes, using the Statistics Act. Practically, no.'

'But they're publishing it already in their market reports.'

'Yes, but their data generally only covers Dublin, and more specifically Dublin 2 and 4. Sometimes Cork, Galway and Limerick. And their data is based on records of just six companies specialising in the prime office market. Most of the country and the full range of office space is missing.'

'So their reports are ...'

'Aimed at their audience. It's probably sound data, but its limited.'

Abandoned, derelict, ruin?

It's one thing to introduce a new tax on dereliction; it's another to implement it. It's going to be unworkable.'

'Surely we just need to identify all the derelict buildings. Local authorities already have registers.'

'Which woefully under-record properties. And what is dereliction? We don't have a tightly defined legal description. The Derelict Sites Act just says properties that look neglected or unpleasant or in a ruinous or dangerous condition. That's all subjective.'

'Surely it's just somewhere that's abandoned.'

'But if it's habitable, then it's vacant not derelict.'

'So habitability.'

'And how do you judge that? People live in squats.'

Always a crisis, never solutions

‘The housing crisis isn’t a recent event. We’ve had a housing crisis in the Liberties since before the state was even formed! Appalling conditions and exploitative landlords pre-independence. Overcrowded slums in 1920s and 30s. Poor sanitation in the 1940s and 50s. Housing affordability in the 1960s and 70s. The stigmatisation of public housing in the 1980s. Speculative development, a property bubble and gentrification during the Celtic Tiger boom. Then the crash, negative equity, mortgage arrears and abandoned urban regeneration. Now unaffordable rents, short-term lets, homelessness, lack of amenities.’

Pat slammed his hand on the table.

‘Always a crisis, never solutions.’

Tip of an iceberg

'These few open datasets are just the tip of an iceberg. Think about all the organisations that must be collecting some kind of housing data relating to stock, land, households, prices, tenants, development finance, energy, construction, homelessness, whatever. All the services and IT systems they're running.'

'Does anyone actually know?'

'I doubt it. If there's an inventory no-one I know knows about it and that includes the Department of Housing.'

'But surely someone has ...'

'You think so?'

'But ...'

'All that data and we're basing most of our policy decisions on a handful of indicators and ignoring the rest.'

Hardcopy

'You're serious? The only way we can make an application is by mail?'

'If it's a very large development, we also ask you to create a website and upload all the documents there.'

'You're the national planning appeals body.'

'Believe me, we'd love to go paperless, but legally we need everything as hardcopy.'

'Legally?'

'It's in the legislation and until that's changed ...'

'That's madness.'

'We even print out all emails and add them to the file.'

'No wonder you move at a snail's pace.'

'That's under-staffing and people chancing their arm with developments they know don't meet the rules.'

Retrofit

'This area here, 40 percent of the office space is vacant.'

Jane moved her finger across the map.

'This area it's 32 percent. This area 17 percent. This area 25 percent.'

'Where are you getting these figures?'

'I walked and observed. I compiled all the 'to rent' and 'for sale' adverts. I spoke to agents. If there's no official data, you make it yourself.'

'Look, I agree that there's empty space available ...'

'Tens of thousands of square metres of space.'

'But very little of it is grade A space that meets current client demands.'

'Then retrofit. Don't build more!'

Housing oracle

'Imagine you have a complete set of detailed housing and planning data at your fingertips. Data on existing stock, how it is used and managed. Information on housing need and demand. Housing experiences and wellbeing. Environment sustainability. Infrastructure. Community and sense of place. Market activity and housing finance. The future pipeline of strategic plans, planning applications and construction. Interactive graphs and maps. An entire oracle of housing. That's what we want to create.'

'Sounds great.'

'And we can deliver some of it. But the rest requires a vast improvement in available data. That's what we need you to fix, Minister.'

Science, testing and negotiation

'If we release a new statistic on planning refusals we'll also be asked for the reason. You'll need to capture that.'

'So, you want us create a taxonomy of refusals?'

'A comprehensive one; there should be no 'other' category.'

'That might take a bit of work; to build then road-test against a sizable sample of application outcomes.'

'And you'll also need a rule-set for selecting a primary reason. Many applications are refused for more than one reason.'

'That's going to be more tricky as it'll need consensus.'

'A new official statistic is always a mix of science, testing and negotiation.'

Outsourcing

'We set the terms of reference and oversee the work, but everything is outsourced.'

'So the council is producing a plan for what is effectively a whole new town, but it's all being undertaken by consultants?'

'We don't have the capacity or the skillset. There's one lead, coordinating company, who are subcontracting to seven others. One is doing the strategic planning, another the urban design. Others are doing data work, the demographic and transport modelling, the environmental impact assessment, infrastructural demands.'

'This is basically public planning undertaken by private interests.'

'Working to our brief.'

'But possibly with their own agenda?'

Doormats

Sadie held up the doormat. Printed on it was '522,486 adults living at home with parents'.

'They're for the election.'

'Eh?'

'You put one outside your door for when a politician canvasses.'

Sadie picked up another.

'14,159 people homeless in May 2024.'

'They're designed to start a conversation about the housing crisis. What they're going to do to solve it.'

'Ah.'

'Do you want one? It's free if you'll place it on your doorstep.'

'I'd need to talk to them? The politicians?'

'Only if you want to.'

'They're all slippery bastards.'

'They'll at least know what issues influence your vote.'

Other registers

'You need to broaden your thinking of data beyond numbers. Beyond representation. As facts. Think of them through other registers.'

'Other registers?' Mike frowned confusion.

'What are data? Are they discursive? Material, semiotic, constructed, performative? Are they technical, political, cultural or social? Do they exist independently of the world or are inseparable from it?'

'You've lost me, Jane.'

'What if we were to understand or express data as a poem or as dance?'

'Data as a dance?'

'What if we were to sonify it or transform it into a colour palette? What would be the relationship between medium and message?'

False impression

'That just doesn't feel right to me.' Killian pointed at the screen. 'There's no way that's the right pattern. For Fingal.'

'That's what's in the spreadsheet.'

'What happens if you alter the category bands? Swap from five to eight classes.'

'Hold on a sec. There.'

'That looks a bit better. How about spatial scale. Try shifting from LEA to EDs.'

The map on the screen refreshed.

'That's a bit more like it. The uneven distribution is starting to show. Now small areas.'

'Give it a moment.'

'It's actually pretty concentrated. The variation was being washed out giving a false impression.'

Trawl and extract

'You're probably going to need to look at 300 plus refusal decision letters across several local authorities to get a full sample of outcomes.'

'Actually, we've already trawled the eplanning sites for all 31 local authorities recording the application status for 137,000 cases from 2023 on.'

'You could automate that work?'

'It took a bit of fiddling with the scraping code to do the different websites, but yeah. And for 2024, there were 3,282 applications refused. Next we're going to capture all those decision letter PDFs then extract the relevant text.'

'You're working on a full sample?'

'That's the plan.'

Fix the data, fix the crisis?

'You actually think that if we fixed the data issues, we'll fix the housing crisis?'

'I think that it'll help us to better understand what's going on.'

'We know what's going on! Rents are too high, house prices are too expensive, not enough units are being built, the chosen solution is being left to the market. Yet, the unfettered market is the key problem.'

'If we want evidence-based debate and policy, then we need sound evidence.'

'This is to do with politics, ideology and capital. Data is just a useful prop for support or an inconvenient truth to be dismissed.'

Data sovereignty

‘And where exactly are the data stored?’

‘The cloud.’

‘Yeah, but where?’

‘Who knows?’

‘But shouldn’t you have some idea? They’re sensitive personal records.’

‘I don’t think it matters, does it?’

‘It might if you’re worried about legal jurisdiction and control. If they’re on a server in Ireland they’re subject to our laws. But if they’re stored elsewhere ...’ Barry shrugged. ‘Or how are they protected as they travel through other states?’

‘Surely Irish state data is governed by our laws?’

‘Even if it’s not located in our state? Why does the German government insist its data’s stored in Germany?’

Quicker than you can do one

'There's no detailed data on disputes between landlords and tenants. So, we've been downloading the RTB determination orders and tribunal reports, then combing through them.'

'By hand?'

'Recording the details in a spreadsheet.'

'But there must be thousands.'

'Over 19,000 for 2015 to 2024.'

'That's going to take you forever.'

'Probably. But we seeing some patterns. Landlords who routinely evict tenants. The number of illegal evictions each month.'

'Why not use machine learning to do it? I could write you a script.'

'Seriously?'

'It'll need ground-truthing and fine-tuning, but it'll do all of them quicker than you can do one.'

There's no way

'There's no way they're going to publish that data unless they're forced to.'

'But it'll help us understand what's going on.'

'Which is why they don't want it in the public domain. It'll be used as a weapon against them.'

'You make it sound like they don't want to fix the housing crisis.'

'I think they genuinely believe they're trying to do that. It's just that they're making little progress.'

'They're making it worse! So, give us the data. Some decent analysis has got to help, right?'

'But GDPR. Constitutional property rights. Resource constraints.'

'Pfff. This is an emergency!'

Bureaucratic mess

'Unless you've worked in the public sector you've no idea how difficult it is to get things done. Take the BCMS. Setting that up involved three core groups – a practitioner advisory group, a steering group and a high-level oversight group. Hanging off of those were a training sub-group, a legislative sub-group, a planning sub-committee, a land-use and transportation sub-committee, a project management office. And I'm probably missing a couple. Each involves several government departments and agencies.'

'Is there no way to cut through all that bureaucratic mess?'

'And risk the fall out of it failing due to exclusions and errors?'

It's all about facts

'There's none of this data politics business from a legal perspective. It's all about facts. Whether someone owns a property, where a boundary line sits, what is stated in a freehold or leasehold agreement.'

'But surely there're ambiguities? That's why disputes exist in the first place.'

'But they're disputes about facts. Their legitimacy and interpretation. But not questioning whether they are a fact or the logic or method of their production.'

'If you're questioning their ...'

'There's no politics. It's technical and legal, not cultural and sociological. It's about establishing truth. If all facts are relative, then all is lost.'

Means to an end

For you, the data's important. For most workers, it's just something they have to do as part of their job. It's a means to an end. Or an additional job on top of everything else. It doesn't have to be perfect, it just has to do what's needed.'

'What's needed is high quality data.'

'The data's generally fine for the task it's being collect for; just not for compiling official statistics.'

'But ...'

'The kind of data you want means redesigning and standardising IT systems and tasks across the whole country. Is that data really worth that kind of investment?'

Glass half-empty

‘Oh, flip,’ Tom rolled his eyes. ‘Part of this data strategy is to implement a data maturity framework.’

‘And?’

‘Guess who’s going to have oversee that?’

‘And what’s the problem? Surely ...’

‘The problem is we’re barely a little past immature. There’s not a chance of us making mature any time soon.’

‘You just need to demonstrate reasonable progress towards the goals set.’

‘It’s a set up for failure. A data immaturity framework and we’d be fine.’

Jane laughed.

‘The glass isn’t always half empty.’

‘The glass is mostly empty. Getting it to half-full is going to take a miracle.’

Mindless work

'Tom. Hey, Tom.'

Clara waited a moment then yelled 'TOM!'

Tom looked up from his screen and plucked a bud from his ear. 'What?'

'How much longer?'

'I don't know. Two hours? Maybe three? I've maybe 50 reports left to enter.'

'Okay, good. Sandra's waiting.'

'Then she'll have to wait.'

'Well, can you speed up a little?'

'I'm happy for someone else to take over. It's just typing in data. It's boring, mindless work. Ask Jimmy, he at least enjoys reading the cases.'

'Jimmy makes mistakes. It needs to be done right.'

'Which means it takes time. Lots of time.'

Data police

'Tara's driving me nuts!' Aoife sat on the park bench. 'She's like the data police. Always checking my work as if I'm a school kid.'

'Those forms need to be spot on.'

'I know. I've been told like fifty times. She actually shadowed me this morning. Watched how I entered the data when I was interviewing mortgage applicants. Then she nitpicked what I'd done.'

'And had you done it correctly?'

Aoife pouted.

'So, you made some mistakes.'

'One. I made one. And only because she was looking over my shoulder, stressing me out.'

'One could be a very expensive error.'

Does the data add up?

'Are we going ahead or not?'

'I think it's touch and go.'

'Poppycock,' Greg said.

'Why touch and go, Conor?'

'I'm just not sure the data adds up. We'd need property prices to rise by more than seven percent a year given current construction, materials and financing costs to make a marginal profit. If those rise, and they probably will, then ... Plus those buying would need a household salary greater than 120K and if you were buying to let then the rent costs ...'

'The market needs supply,' Greg countered.

'They'll be buyers.'

'Would you buy at that cost?'

A polished turd is still a turd

'Have you seen the state of that data? It's a mess. It's like landfill. A jumble of garbage.'

'It's not that bad. It's comprehensive. Every ...'

'Seriously? Have you tried working with it? We're spending days cleaning and validating it. I've worked on dozens of projects. I've witnessed everything. People forgetting to do stuff. Deliberately omitting things. Cutting corners. Adding errors. Mis-typing. Mis-coding. No standardisation. No double checking. Moving on and not bothering to train the new person. Not giving a ... But this is the worst.'

'But once you've cleaned it ...'

'A polished turd is still a turd.'

This is how it works

‘Go and ask Kenny, the GIS guy, for the data. Not the manager, okay. Kenny.’

‘And if I meet the manager?’

‘Then say nothing. His first instinct will be to say “no”.’

‘But we need the data to do what they want.’

‘He’ll be more worried about letting the data slip out of their control. Kenny knows what we’re going to do with it.’

‘But surely ...’

‘Look, this is how it works. We’ll show the manager the final product. He’ll be thrilled; boast about it to other units. But nothing would ever happen if it were left to managers.’

Won't or can't?

Mick turned away from his screen. 'They said no.'

'What?'

'They won't give us the data we wanted at the small area level. They don't do special runs at that scale for disclosure reasons. If the numbers are too low, with local knowledge you'd be able to identify households.'

'But we were asking for all the sub-categories to be collapsed into one. There shouldn't be disclosure issues.'

'Probably. But you'd only know by doing it.'

'So, it's more they won't do it than they can't?'

'Well, I guess they don't want to set a precedent that'll lead to more requests.'

Seeing sense

'You're taking the local authority to court?'

'They're limiting our right to trade.'

'But surely going to court will delay things further?'

'Someone needs to do something! We're in the middle of a housing crisis and their policy is to set targets that are ceilings not floors. They're capping the number of homes that can be built.'

'But ...'

'On the basis of out-of-date data and shitty models. The land is serviced. Greystones is on the rail line. They should be encouraging us to build not holding us back!'

'But court?'

'It's the only way to make them see sense.'

Make data come alive

'I'm telling you, maps are the answer. If you want things to adapt and change then you have to be provocative. You have to nudge people to accept what you're proposing. And that has to be done gently and not in a hostile way. Which means showing them data in a compelling way. That means maps.'

'Why not just provide a convincing argument?'

'You have to show not just tell. People understand maps, especially since the smartphone. Every app has a map for communicating information. Plus they're interactive. You can play with them. Query them. They make data come alive.'

A botched workaround

'You can't just work round a glitch, John. You might get rid of an operational problem, but you're creating data problems.'

'It's the system, it's ...'

'You've a bunch of entries where the date of install is missing or set to some random date in 2013!'

'I thought ...'

'These errors mess up our reporting on the awarding of upgrade grants and their completion. But more importantly, whether someone can draw down the money on completion. If a grant was awarded in 2013, they're past the 12 month window. If there's no date, it never happened.'

'Sorry, I'll fix them.'

Yes as a default

'Sir, it's about the survey.'

'It's Dave, not Sir. What about the survey?'

'I think we should review some of the questions.'

'Why?'

'My hunch is our sample are providing answers that they think we want to receive. Or they're guessing because they don't know the answer.'

'Such as?'

'Take question eight, 'is the site located in a high radon area.' Most people living in Dublin answer 'yes', even though it's a low risk area. Which means much of the data are incorrect.'

'They tick 'yes' as a default?'

'That's my assumption.'

'And your solution?'

'Add a 'don't know' category.'

Upskilling by osmosis

'You can't expect people to be good at working with data when you provide little beyond basic training and no supports. If data's central to their work they need on-going training.'

'We've provided training courses, but few people ever signed up for them.'

'That's because you're expecting your staff to do them without reducing their load to do so. In fact you've layered all the new data-related requirements on top of what they were already doing. You want more for less without showing and enabling people how to do more.'

'That's not ...'

'Fair? But expecting upskilling by osmosis is?

Face value

'I can't believe they've swallowed that crap.'

'Look, just calm down. Vested interests are always going to promote their interests. But that doesn't mean that what they're saying isn't true.'

'Isn't true? It's bullshit. And there's no way to verify their data or modelling.'

'It looked pretty plausible to me based on my experience.'

'You're on their side now?'

'I'm not anyone's side. I'm saying that they do have data the rest of us don't. Which gives them better market intelligence.'

'And we just accept that at face value?'

'No. But we also can't just dismiss it out of hand.'

Smoke and mirrors

'I don't believe a word of it. It's all smoke and mirrors.'

'Just calm down, Lorna. They're entitled to make their pitch.'

'They're going to ruin this area with that development. Wreck an ecosystem and blight a community.'

'They're also going to bring a lot of jobs. And local businesses will be suppliers.'

'Did you look at their numbers? They're whatever the creative accountancy of demographics and labour markets is.'

'Lying with statistics.'

'Exactly!'

'Cherry-picked data going into fudged scenario models.'

'Then provide your own counter-data. Do it transparently. Let people see what they've done.'

'Yes. Good. Let's do that!'

Discrepancies

'I know you don't like the new system and you want to continue to use your spreadsheets, but that has to stop.'

'But you can't trust that new system. I know what I'm doing the old way.'

'But it's leading to too many discrepancies. Errors are more difficult to make in the new system.'

'I'm having to use my spreadsheets to validate it. See, it's undercounting cases. 458 in my spreadsheet, 454 in the new system.'

'Let me have a look.'

Five minutes later, Mark turned from the screen.

'You've duplicates in your spreadsheet.'

'What?'

'Use the new system, Margaret.'

Datawashing

'Is there anything on the open data portal we could use?'

'There's plenty of stuff, but it's mostly useless. I mean, it's like a joke, right? All the crap they've uploaded. Low value datasets of little practical use. A smattering of high value stuff that's a mess. Or duplicates that have better metadata and documentation elsewhere. Little wonder that there's barely any downloads.'

'I take it you're not a fan!'

'I'm a big fan of open data, but that site's basically datawashing.'

'Datawashing?'

'Pretending that you're taking open data seriously, when it's really for show. It's a form of hoodwinking.'

Strategy seeking evidence

'This evidence-informed strategy is more like strategy seeking evidence. Everything about it has been arse-about backwards. I feel like a total fraud.'

'Welcome to the civil service where our job is to make the government's policy a reality.'

'Their policy is that policies should be evidence-informed!'

'Informed, not evidence-based.'

'Which is a license to ignore the evidence for political gain. And once the strategy's launched, then the job will be to identify data that'll help measure its progress. Which actually means finding data that demonstrates it's been a success. Whether it has been or not.'

'Cynics are always fast learners.'

Better than nothing at all

'But if you use Google Street View you'd be able to see the derelict properties.'

'Hmmm. It's a good idea in theory, but I'm don't know if it'll work in practice.'

'But it has potential, right?'

'Some of their imagery is pretty old and they've not updated it recently.'

'But they've re-done towns. How about a pilot test? One town.'

'It's not the towns I'm worried about. It's the rural areas. A lot of derelict properties are going to be obscured.'

'So use their satellite imagery. The two together.'

'It's going to be sketchy.'

'It'll better than nothing at all.'

A critical look back

'We need to step back and try to see the bigger picture. We've never done a proper review.'

'Everything seems to be fine.'

'But is fine good enough? Could it be better? What could we tweak? What might be a useful new innovation? What do others think about what we're doing?'

'I don't hear any complaints.'

'Complaints and feedback are not the same thing. Taking a critical look back at what we've doing and learning from it; that has to be a good thing, right?'

Paul shrugged.

'That was a rhetorical question.'

'Let me guess, I'm to lead this review?'

Following best practice

'Don't worry about it. We're not going to be starting from scratch.'

'We are though, aren't we?'

'Well, yeah, but loads of others have already done it, so we can learn from them; follow a tried tested approach and avoid the mistakes they made.'

'We both know that setting up a new data system is going to be a nightmare. And the technical stuff is only going to part of the challenge.'

'Which is why we need to establish best practice. I'm telling you, this isn't going to be as bad as you think it'll be.'

'Except, it will.'

Local allocation without local modelling

‘We’re allowing you to adjust the scenarios by tweaking the weights in the models. For example, you can adjust household formation figures or migration flows based on local knowledge or policy choices.’

‘And it predicts how many housing units we need to build?’

Claire nodded. ‘It estimates housing targets under different scenarios based on all data we’ve pre-loaded.’

‘At local level?’

‘County level.’

‘But not for towns?’

‘A number of datasets are only released at local authority scale. It’s up to you how you allocate the target figures to towns.’

‘Local allocation without local modelling.’

‘But better than no modelling.’

What can we learn from our own data?

'We have a massive of information on planning – all the planning applications and submissions, our assessments of them, the material we gather for strategic plans. But we do nothing with it. We don't look at trends in applications or see if there're patterns in our decisions. We don't ...'

'Okay, okay, I get your point. What are you proposing?'

'We set up a planning intelligence unit. See what we can learn from our own data. Feed it back into ...'

'We don't have the resources for that.'

'But ...'

'The system isn't broken, is it?'

'But it could be better.'

It'll never work

'You have data. The local authorities have some. Government departments. State agencies. Local businesses. Civil society organisations. What if all that data was pooled rather than sitting in silos?'

'It's not that simple. GDPR. Commercial interests. Security.'

'But everyone gains from sharing the data. They'll have way more intelligence. And privacy and other concerns can be dealt with.'

'But who'll run it? Business won't trust government to do it properly. The state isn't handing personal data to a private enterprise. NGOs trust neither.'

'Then create a co-owned data trust and implement strong data governance.'

'This is Ireland. It'll never work.'

Data will ...

'You're obsessed with data. Data this, data that. Data, data, data. Data isn't going to fix this housing crisis.'

'But more and better data will help us make greater sense of what's going on. It will provide a stronger evidence base for policy and decision-making. It will provide the fuel for running realistic scenarios and better predicting the future. It will let us better administer how we manage and provide housing. It will ...'

'Okay, okay, I get your point.'

'Data by itself won't fix anything. But used as part of a process ...'

'You're preaching to the newly converted!'