UNIVERSITY OF LIVERPOOL | Heseltine Institute for Public Policy, Practice and Place

*A Position Statement to accompany the publication of The Good Practices Reference Guide*


# *A Position Statement: Building a Data Ecosystem in Liverpool City Region to Unlock the Value of Big (Local) Data*

By Mark Boyle and Sue Jarvis, Heseltine Institute for Public Policy, Practice and Place,
Alex Singleton, Geographic Data Science Lab, University of Liverpool
and Aileen Jones, Liverpool City Region Combined Authority

# 1. PURPOSE OF THIS POSITION STATEMENT

As elsewhere, civic leaders (CLs)[1] in Liverpool City Region (LCR), including and in particular Liverpool City Region Combined Authority (LCRCA) and its six constituent local authorities, are alert to the possibilities presented by computing power and data-driven solutions to critical economic, social and environmental challenges, and to the reality that at the centre of any high-performing 'Smart City' Region[2] must be a high-performing data ecosystem. Accordingly, CLs in LCR are now striving to develop such an ecosystem to enable stakeholders to more effectively exploit the latent value inherent within local datasets, and datasets with local expression. The extant energy, strengths and assets of the City Region are cause for optimism – but the practice of sharing data is still nascent, and significant obstacles will need to be navigated if further progress is to be made.

**The strategic importance of computerised urban data analytics in Liverpool City Region**

Computerised data analytics are recognised key progenitors of growth in central LCR strategic documents (transport, skills, innovation, industrial strategy, spatial planning, regeneration, infrastructure, climate, housing, health, education), and if harnessed effectively provide the opportunity to improve both the productivity and prosperity of the City Region. The LCR Local Industrial Strategy to be published in 2020 has an explicit focus on 'Tech for Good. Moreover, the Metro Mayor has articulated his intention to make LCRCA a data-driven organisation that utilises data to inform and improve public policies and services. Local authorities are also making significant moves to conceptualise, audit and improve their data infrastructures, and Liverpool City Council (LCC) in particular is taking stock of the values it wishes to use to calibrate its data management practices.

The University of Liverpool's *Heseltine Institute for Public Policy, Practice and Place* is committed to supporting LCR's CLs in building a high-performing data ecosystem, capable of rivalling international peers. In March 2020, the Institute will host an international symposium entitled *Building Smart Cities with Citizens and for Public Good*. To inform this symposium, international experts based at BABLE UG and Fraunhofer FOKUS in Germany have undertaken a review of international good practices – mainly drawn from European cities – in the development of urban data ecosystems, and in a Heseltine Institute-sponsored *Good Practices Reference Guide* have profiled paradigmatic exemplars.

To complement this Guide, the Institute has published this parallel *Position Statement* to highlight the lessons extracted from the international good practices studied which might inform CLs in LCR. We are conscious that many local institutions, businesses and civil society groups have much to say on this topic; there is no shortage of creative thinking and potential future imaginaries[3]. Ultimately, it will be for CLs to gather and synthesise good ideas, determine local priorities, and secure local consensus on what a smarter LCR might look like and who will be responsible for building it. Our contribution is therefore strictly exploratory and confined to furnishing the local conversation with evidence, food for thought, and perhaps even a little inspiration.

**In this Position Statement we:**

- Clarify what we mean by 'Smart City' and define associated key terms
- Collate and articulate what we consider to be the fundamental challenges
- Provide a concise health check of the status of computerised urban data analytics and data-driven solutions in LCR
- Suggest a roadmap for redevising data sharing arrangements in LCR
- Pose 'consultation questions' to clarify next steps.

1 We use this phrase to refer to elected officials and those holding office, senior leadership teams in the LCRCA and in constituent local authorities, chief officers at the Local Enterprise Partnership, leaders of other public, private and third sector anchor institutions, and local citizen-influencers.

2 According to the UK Geospatial Data Commission better sharing of geospatial data alone could unlock £11 billion of untapped growth.

3 Although not an exhaustive list, we note here in particular the influence of Iain Buchan, Simon Maskell, Ann Williams, Rosemary Kay, John Whaling, James Noakes, Joanne Phoenix, Lee Omar, Colin Sinclair, Joseph Spencer, and Brian Bishop on our own thinking.

# 2. THE SMART CITY: KEY DEFINITIONS

The phrase 'Smart City' lacks a clear and universally agreed definition, and its use often causes confusion and uncertainty. In this Position Statement, we adopt the following working definition:

**Smart City**: a city in which networked computing power and computerised data analytics are embedded in foundational operational systems and in the fabric of everyday life, to enable institutions, infrastructure, services (including municipal administration, education, healthcare, public safety, real estate, transportation, and utilities) and individual citizens to benefit from increased instrumentation, interconnection and intelligence (sometimes abbreviated to In3)[4].

A series of other common definitions then follow:

- **Instrumentation**: the capability to generate and capture live real-time data through the use of sensors, meters, appliances, and personal devices
- **Interconnection**: the capacity to circulate, pool, exchange, integrate, and communicate this data, through digitalisation and technologies such as IoT, 5G, radio-frequency identification (RFID), data platforms, and interfaces
- **Intelligent**: the application of data science (statistics, modelling, optimisation, visualisation) and computing power (AI/autonomous systems/machine learning, and cognitive, quantum and ubiquitous computing) to enable Big Data to make decisions, and/or assist humans to do so
- **Big Data**: a label applied variously to data which is: huge in volume, consisting of terabytes or petabytes; high in velocity, being created in or near real-time; diverse in variety, being structured and unstructured in nature; exhaustive in scope, striving to capture entire populations or systems (n = all); fine-grained in resolution and uniquely indexical in identification; relational in nature, containing common fields that enable the conjoining of different datasets; and flexible, holding the traits of extensionality (can add new fields easily) and scalability (can expand in size rapidly)[5]
- **Urban data ecosystem**: the presence, scale, quality, and alignment of instrumentation, interconnection, and intelligence in any given city
- **High-performing urban data ecosystems**: data ecosystems which maximise the efficacy – or technicity – of instrumentation, interconnection and intelligence (In3)
- **Technicity**: the productive power of technology to achieve intended outcomes.

4 This definition draws upon IBM's depiction of smart urbanism and In3.
5 Characteristics of Big Data per Rob Kitchin's ERC Programmable City project: http://progcity.maynoothuniversity.ie.

# 3. UNLOCKING THE VALUE OF URBAN DATA: FRAMING THE FUNDAMENTAL CHALLENGE

In the tradition of Kondratiev, Kuznets, and Schumpeter, economic historians have long sought to identify waves and cycles of technological disruption and economic advancement. Today, there is emerging an increasingly influential claim that a fourth industrial revolution (also known as a cyber-physical revolution, Industry 4.0, 4IR, and/or Society 5.0) is imminent, predicated upon computing power, digital connectivity, and a data revolution.

As part of a wider suite of interventions designed to prepare their cities for the fourth industrial revolution, civic leaders (CLs) in metropolitan centres across the world are increasingly leveraging computerised data analytics to help solve their most pressing economic, social and environmental challenges[6]. At the heart of the 'Smart City' is Big Data, provocatively titled 'the new oil'. It is anticipated that cities with the largest, best archived, and most open data 'lakes', 'warehouses', 'repositories', and 'platforms' will enjoy significant competitive advantage.

**Data: the New Oil?**

Data are a vital primary raw material for the Smart City; just as oil shortages constitute a serious threat to an oil-fired urban economy, any scarcity of data, particularly well-curated data, represents an existential threat to a data-fired urban economy.

Few cities are presently in a position to fully harness the data revolution; most are only now actively working to build fit-for-purpose data infrastructures. A heterogeneous geography of capability is emerging, both internationally and within the UK. To unlock the full value of Big Data, CLs and their institutions and other significant data owners will need to radically increase accessibility for public, private and third sector stakeholders – and in particular their software coders – to big *local* data[7]. But convening data keepers and securing consensus for new institutional data sharing mechanisms presents profound logistical, ethical, legal, commercial, and political challenges.

In particular, the origins and development of smart technology within a framework of what Shoshana Zuboff[8] calls 'surveillance capitalism' have given rise to technology that is not only configured primarily to serve the interests of commercial data harvesters, but which is also substantially – and manifestly – under-regulated. For Zuboff, surveillance capitalism is driving computerised data analytics, and over-determining the terms of its application and trajectory of its development.

Whilst a powerful coercive force, the appropriation of computerised urban data analytics by surveillance capitalism is not inevitable. In fact, problems arise not because of smart technology, per se, but because this technology is being enabled and constrained by a very particular politico-institutional dynamic – a new mode of capital accumulation whose business model is the extraction of value from amassed linked personal data with little juridical, regulatory or ethical oversight. Computerised data-driven solutions are essentially benign. Everything depends upon the political constitution of data markets, and the architecture of data ownership and sharing arrangements – specifically, how these arrangements are designed, regulated and governed, and whether they command a social licence.

**The future is ours to make**

Whether or not computing power and the data revolution will lead to a better tomorrow will depend upon the social and political choices we make today.

We believe that there is an incontestable argument for radically scaling and accelerating the innovation and adoption of computerised data-driven solutions and unleashing the powerful data analytic capacities of public bodies, private sector market actors, and third sector social enterprises. But equally, we believe that critical market failures indicate that Smart City experiments have accelerated ahead of ethics, law, public policy, purpose, regulation, and governance, and that there is legitimate concern that burgeoning computerised data analytic solutions are lacking in social responsibility and failing to deliver for the public good. There is a need to temper the rush towards a 'smart' future, and to take more seriously the potential threats posed by AI and Big Data and the importance of comprehensively attending to what some have called 'tech and data science for public good'. In our view, both innovation and regulation are not incommensurable; indeed, a clear ethical, legal, and stable data sharing compact will likely spur investment and inspire confidence in computerised data-driven solutions.

**The fundamental challenge for architects of the Smart City**

How can municipal and civic leaders design, build and govern urban data ecosystems that enable public, private, third sector actors and citizens to exploit more fully the powerful data revolution in a way which is democratic, ethical, underpinned by a social licence, and which maintains ongoing stakeholder – and public – trust?

**Put more simply:**

How can city leaders and municipal authorities increase access to and enable better harvesting of urban Big Data whilst maintaining buy-in from stakeholders and citizens?

6 See Batty, M., 2018. *Inventing Future Cities*. Cambridge, Massachusetts: MIT Press.
7 See Singleton, A., Spielman, S. & Folch, D., 2017. *Urban Analytics*. London: SAGE Publishing.
8 Zuboff, S., 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. London: Profile Books.

In this Position Statement, we contend that if we are to ensure the use of smart technology and Big Data is maximised for the public good, there will be a need to place under scrutiny the prevailing structures of data ownership and stewardship, to gain a better understanding of how smart/digital technologies and data sharing markets might be rendered subservient to national and local democratic institutions and processes, and how they will be configured, regulated, ethically governed, and incentivised such that they serve the commonweal.

CLs in many cities are now actively building innovative institutions to unlock the value of urban data by increasing access to data but on terms that are sustainable and ethical. Varying in legal status, constitution, anatomy, autonomy, membership, and modus operandi, a wide variety of institutional forms and mechanisms designed to improve data curation and sharing are now emerging[9]:

- *Public sector to public sector data sharing*. Stewards of publicly owned data work to defragment the data ecosystem by creating central data reference architectures, common and Open Data Standards, and Open Interfaces to enhance interoperability. They assume responsibility for creating a trustworthy data environment and determining which other public sector bodies have access to datasets, under what conditions, and for what purposes.
- *Public sector to private sector data sharing*. The absence of data sharing arrangements between the public sector and for-profit market actors contributes to market failure, untapped datasets, and sub-optimal solutions to critical policy and public problems. Stewards of publicly owned data also adjudicate on which public sector datasets should be published as Open Data, which should not, and the levels of access that should be granted to other stakeholders – including private sector actors – to these non-Open datasets and on what terms. Robust new business models are being developed to incentivise market actor contributions to the project of unlocking the value of local datasets by enabling the commoditisation and monetisation of data by the private sector, but on terms which are ethical and for public good.
- *Private sector to public sector data sharing.* New sources and forms of private sector Big Data are starting to be incorporated into public data repositories. We need to sustain the UK's internationally leading investments in data philanthropy that establish the provenance, create adequate documentation, and develop the infrastructure for safe curation and analysis of post-competitive private sector data[10]. Motives for sharing data include fiscal benefits, access to analytical capacity, and civic duty. An imperative for the public sector is to avoid vendor lock-in, which renders public bodies reliant upon particular data providers for critical products and services and unable engage other vendors for fear of penalties.
- *Public sector regulation of data and data analytic markets*. Nationally, governments are now catching up with smart technological innovations and working to redevise data and data analytic marketplaces. But CLs have levers to pull too. Some are now refocusing thinking, away from narrow economic principles and towards securing a social licence for the innovation and adoption of computerised data analytics, by reframing markets as social and political constructions and promoting collaborative and participative market design. Here, and against the backdrop of global concerns over growing social inequalities and climate change, CLs might initiate seismic shifts in thinking within the corporate sector, from a singular focus upon profit and maximising shareholder value (MSV), to wider economic logics, defined by ethical and 'stakeholder capitalism' and gestured to in ideas such as the triple bottom line[11].

**From shareholder capitalism to stakeholder capitalism**

**Davos Manifesto 2020: The Universal Purpose of a Company in the Fourth Industrial Revolution**

Reflecting growing interest in commutative justice and ethical capitalism, to coincide with its 50th anniversary, the World Economic Forum launched a new Davos Manifesto in January 2020, setting out ethical principles to guide companies in the age of the Fourth Industrial Revolution. An old mid-twentieth century model – stakeholder capitalism – of a company's purpose was dusted off and given new life for the twenty-first century.

According to Klaus Schwab, Founder and Executive Chairman of the World Economic Forum (2019), "The purpose of a company is to engage all its stakeholders in shared and sustained value creation. In creating such value, a company serves not only its shareholders, but all its stakeholders – employees, customers, suppliers, local communities and society at large. The best way to understand and harmonize the divergent interests of all stakeholders is through a shared commitment to policies and decisions that strengthen the long-term prosperity of a company."

According to the Manifesto, companies should now be judged according to the extent to which they participate in 'shared value creation', measured by 'environmental, social, and governance' (ESG) goals, in addition to standard financial metrics.

**Regulating urban Big Data markets**

In contributing to the design of data and data analytics markets, CLs will need to:

a) Acquire expertise in areas such as data property rights, managing externalities, and countering monopolies. Standard tools such as Game Theory and Agent-Based Modelling can be applied to ascertain how particular market configurations might impact company strategies
b) Understand better how to make markets more inclusive by opening them up to the concerns of those who are let down by them: what would make markets 'better' from the perspective of those failed or excluded by them, and what actions could be taken to open up the markets in question to their concerns? CLs will need to build institutions to enable a wide range of concerned stakeholders to engage with and challenge a market's failures, to question its logic and design, and to make markets more collaborative and open to democratic debate and social or political praxis[12].

9 Examples include data licences, warehouses, platforms, trusts, repositories, cooperatives, portals, juries, partnerships, assemblies, Community Interest Companies and panels.
10 The UK Consumer Data Research Centre (CDRC), based at the University of Liverpool, is a world-leading example of how data philanthropy can be operationalised at scale for the public good: see https://www.cdrc.ac.uk.
11 See the Davos Manifesto 2020: https://www.weforum.org/the-davos-manifesto.
12 This observation derives from insights produced in the ERC MISFIRES project led by Susi Geiger at UCD: https://misfires.ucd.ie.

# 4. A CONCISE HEALTH CHECK OF THE STATE OF COMPUTERISED URBAN ANALYTICS IN LCR

Building a smarter LCR, in the political-economical-institutional context of a City Region recovering from de-industrialisation and undergoing urban regeneration and renewal, presents its own unique challenges and opportunities. The *Good Practices Reference Guide* identifies four fundamental action areas and twelve key themes as a starting point for an aspiring Smart LCR to address. Stakeholders in LCR may find it beneficial to use the framework to think about the progress being made locally in each of the action areas and with respect to each of the key issues. More importantly, they might use the framework to consider the extent to which there is and can be alignment across the entire data ecosystem. Below, we use the framework to undertake a concise health check of the state of computerised urban data analytics and data-driven solutions in LCR.

## Action Area 1
## GROUNDWORK AND PREPARATION

The groundwork and preparation required when starting to organise a data ecosystem (*rationale for data plans, stakeholder mapping, data audits*)

**Commentary**

Whilst at various points local authorities and other stakeholders have sought to craft a new data vision for their constituent areas, sustained follow-up and lasting impact has been insufficient. The project of building a smarter LCR appears to have experienced a number of faltering starts, and thus far has made only intermittent progress. Owing to the historical **lack of an overall data vision or strategy**, a decade of biting austerity and far-reaching cuts to local authority budgets, and the embryonic status of LCRCA, LCR's data landscape at present is best described as fragmented:

a) Public, private and third sector organisations have insufficient knowledge of which other organisations are collecting data and of what nature

b) Datasets are currently owned by a wide variety of institutions and scattered across numerous sectors and jurisdictions

c) Data are being generated and curated in different formats without an agreed LCR reference architecture and according to different standards and protocols, and are therefore lacking interoperability

d) Data are too often sealed within legally, commercially, and technically closed systems.

As a consequence, public and private datasets in and around LCR are being insufficiently shared, and the full economic, social, and environmental value of local data remains underexploited.

However, LCR now has a real opportunity to accelerate its smart, digital and data transition, develop and scale its data ecosystem, and implement smart and data-driven solutions to pressing economic, social and environmental problems. The necessary energy, impetus and political will to build a smarter LCR and a high-performing LCR data ecosystem is already and flourishing. CLs in LCR are cconvinced, capable and committed to the project. They are acutely conscious of the Smart City agenda and fully awake to the importance of high-performing data infrastructures; there are good links with world-class industrial partners and commerce, SMEs are displaying ever-evolving leadership abilities, and capital investments and pilot projects are establishing significant (if fledgling) infrastructures. **A historic moment of opportunity beckons**.

## Action Area 2
## GOVERNANCE, MANAGEMENT, ETHICS AND REGULATION

The governance, management, ethics, and regulation arrangements necessary to set the basic framework *(governance and management, governing for public good: building Smart Cities with and for citizens, governing data ethically)*

**Commentary**

No one can 'own' a Smart City, but every Smart City needs to have a recognised organisation with sufficient intensity of focus and convening power to ensure that resources, capacities and infrastructure are in place, and to drive change. Such an organisation has yet to be conceived in LCR. Given the ever wider expansion of urbanisation processes into hinterlands and functional areas of city cores, and the establishment of a City Region scale of governance and coordination, it would seem imperative that a data ecosystem be developed at the LCR scale, and LCRCA would appear best-placed to lead or at least nurture this project. However, this institution is still relatively new and is only now beginning to garner influence. Moreover, a decade of austerity has taken its toll. Public finances are severely limited, and no institution, including the LCRCA, currently has adequate resources and scope to champion a local computation, digital and data transformation. Further investment and capacity-building will be essential. Mechanisms are needed to ensure that innovation and business opportunities are maximised, ethics are upheld, and the voices of citizens are heard. CLs might usefully identify resources and skillsets within achor institutions, which if aligned and/or combined could cultivate sufficient capacity to devise and oversee the implementation of an LCR data strategy. Meanwhile, the new, pioneering SIF-funded health and social care-focused **Liverpool Civic Data Cooperative** (CDC) provides the single most important addition to the data governance landscape, and next we propose a vision which has the concept of the CDC at its heart.

## Action Area 3
# TECHNICAL INFRASTRUCTURES AND CHALLENGES

The technical infrastructures and challenges which are fundamental for an integrated ecosystem *(building Open Data ecosystems and fostering interoperability, investing in data infrastructure: hardware and software, data security, data visualisation*)

**Commentary**

LCR has a number of strategic technical-related assets which afford significant competitive advantage and upon which it should build**:**

- **Deployment of 5G networks**: a consortium comprising Liverpool City Council's Adult Social Care Department, the NHS, university researchers, local SMEs, and a UK 5G technology vendor are testing the use of 5G in the City Region, along with artificial intelligence, virtual reality and IoT, with the aim of reducing the digital divide and enabling better health and social care in communities
- **A supercomputer dedicated to industrial R&D**: the Hartree Centre/Daresbury hosts one of the most powerful supercomputers in the UK, and works in cooperation with STFC's UK research and development centre to enhance the data analytic capacities of local actors, including SMEs
- **Digital skills**: local government and industry-led initiatives are addressing some of the skills gaps (basic and advanced) that exist in the City Region; for example, the NHS has several programmes for training its staff and patients to use digital tools
- **Sensor City**: this initiative is a joint venture between the University of Liverpool and Liverpool John Moores University, and offers technical expertise and business support to academic and industry enterprises to promote the use of sensor solutions
- **LCR Activate**: this support and funding programme is designed exclusively for helping digital and creative businesses in the City Region to grow. It provides practical, hands-on support and funding to help digital, creative and createch businesses develop using emerging technologies, such as AI, machine learning, virtual and augmented realities, Big and Open Data, high-performance computing, and cloud and cognitive computing
- **Consumer Data Research Centre**: led by the University of Liverpool, this centre functions as a consumer data access broker for academic and other stakeholder groups conducting research for the public good
- **Materials Innovation Factory**: this University of Liverpool and Unilever facility is a world-leader in robotics- and computer-aided materials discovery and design
- **Knowledge Quarter**: this £1 billion, 450 acre innovation district located in Liverpool city centre incorporates world-class innovation clusters in the areas of health/antimicrobial resistance/infectious disease, materials chemistry, and high-performance and cognitive computing
- **LCR Backhaul Network (The Digital Spine)**: an investment in a 260kms fibre optic network connecting all six local LCR local authority areas, the Hartree Supercomputer, and the GTT fibre optic cable
- **AI Solutions Hub**: the STFC Hartree Centre, and IBM Research intend to expand their existing partnership, which applies world-leading AI and high-performance computing to address industrial challenges, as part of the UK's Industrial Strategy
- **Liverpool Civic Data Cooperative (CDC)**: a SIF-funded project designed to increase wider access to health and social care date with the active consent of local citizens.
- **Transport for the North**: the LCR Combined Authority works alongside Merseytravel to improve the delivery of transport services. For instance, real-time information about bus routes and service efficiency is now offered to users in the City Region, and via the use of sensor technology, traffic lights are enabled to give priority to buses when roads are congested to improve punctuality. In their efforts to improve public transportation in the Northern region, the Combined Authority and Merseytravel are planning to implement a single travelcard to enable users to traverse the entire region. To this end, they will share data about timelines, timetables, and planning with all the Northern regional transport authorities
- **Virtual Engineering Centre**: a leading technology innovation centre for engineering development using virtual engineering technologies and expertise
- **Strong market actors**: the region hosts major tech companies such as Atos, EPAM, Unilever and The Very Group, and has the UK's largest SME-led e-health cluster, a nascent IoT cluster, and a growing number of tech enterprises and SMEs in the digital and creative sectors.

Securing long-term funding to extend these initiatives and scale pilot projects will be critical. Aligning these assets through a strategic data plan will be required to enhance the technicity of each.

Crucially, LCR lacks a common technical reference architecture and collectively agreed and shared data standards and protocols. The lack of a searchable repository for data and the absence of a common Digital Object Identifier (DOI) protocol impedes the use of the most powerful data sharing practices. The University of Liverpool's Consumer Data Research Centre comes closest to offering something akin to an LCR Open Data Platform[13]. As a consequence, the **lack of interoperability** between datasets remains a fundamental problem and significantly hampers the capacity of computerised data analytics to provide effective solutions to LCR's chief economic, social and environmental problems.

13 https://data.cdrc.ac.uk/geodata-packs.

Action Area 4

## RESOURCES, FINANCE, ECONOMICS

Resources, finance and economics questions which need to be asked of data-driven ecosystems *(financing models and procurement, cost–benefit analysis)*

**Commentary**

The challenge of paying for Smart City projects derives from the novelty and risks associated with investing in proto-technology; investor confidence is often reduced by the absence of a demonstrable 'proof of concept' ledger and a lack of clarity over how to monetise the benefits of a project. To date, LCR has leveraged finance from EU Structural Funds, public sector strategic and capital funds, and innovative public and private partnerships. But there is clearly a need for fresh thinking on financing models to enable the scaling up and rolling out of pilot projects and further interventions. If it is to secure sufficient finance to fund its Smart City objectives, LCR will need to develop **sustainable business models** which prioritise the interests of the public purse and the importance of investing in projects which serve the public good and the concerned citizen, whilst remedying market failure by providing sufficient incentives for trade and commerce actors to enter the data and data analytic markets.

# 5. UNLOCKING THE VALUE OF BIG (LOCAL) DATA IN LCR: A VISION AND ROADMAP

We conclude by proposing a potential vision and roadmap for the development of a high-performing LCR data ecosystem. This closing section is by necessity exploratory and designed to be thought-provoking. Only through local dialogue, led by CLs and including stakeholders and citizens, can a consensus be reached on a sustainable path forward. A clear and agreed vision and roadmap will provide a strong orientation for the entire data ecosystem building project, informing all subsequent decisions, focusing effort, rationing scarce resources, and maximising impact for any given spend. A strong sense of mission will act as a progenitor of new bespoke data policy ideas, interventions and practices.

**A vision for a high-performing LCR data ecosystem**

In searching for an overarching vision, CLs in LCR must begin by asking a number of critical questions:

- Why should we build a Smart City?
- Which vision, values and principles should underpin our efforts to develop a high-performing data ecosystem?
- What can we do with smart technology that we could not do with simple technology?
- What might a citizen-centred *and* entrepreneurial Smart LCR look like, and how might we build it?
- How seriously should public concerns over surveillance capitalism be taken?
- What does it mean to speak about 'tech for good' and 'tech for public good' in the context of Smart Cities?
- How might CLs in LCR create a Smart City at City Region level, in collaboration with citizens and for citizens, dedicated to the cause of enabling public, private and third sector actors to capitalise on Big Data whilst also serving the public good?

These questions defy easy resolution and have yet to be fully debated by CLs in LCR. Nevertheless, we detect a consistent orientation in local thinking, which if cultivated and more clearly articulated could underpin a future vision. There is no appetite among CLs to fall prey to the kinds of techno-fetishism, naïve boosterism, or overblown hyperbole that too often accompanies talk of Smart Cities. CLs are keen to engage computerised data-driven solutions only insofar as they prove helpful in addressing real and pressing local, economic, social, and environmental challenges. For CLs, purpose must always be prioritised over technology, and the latter must always be subordinate to the former. The object is not to become a Smart City per se, but to deploy the most effective computing power and data-driven solutions to address problems and challenges, and only if these solutions prove superior to others.

**For a challenge-based approach**

We believe that an emerging consensus around this fundamental point of departure opens up the possibility of adopting a **challenge-based approach** to building a smarter LCR.

Initial activity could be focused around a small number of agreed high-priority challenges. For example, under the headline challenge of 'Supporting Inclusive and Clean Growth', attention might first be given to building data ecosystems fit for purpose for tier one challenges, with learning then applied to building effective instruments of data governance for tier two challenges.

| **Headline Challenge: Supporting Inclusive and Clean Growth** | |
|---|---|
| Tier 1 challenges | Improving health and reducing health inequalities, fostering inclusive economic growth, decarbonising the economy, and becoming a net-zero carbon City Region by 2040 |

| | |
|---|---|
| Tier 2 challenges | Improving economic productivity, maximising the opportunities associated with Brexit, addressing mental health, promoting service reform, tackling air pollution, resolving homelessness, regenerating town centres, improving education outcomes, creating a circular economy, enabling SMEs to internationalise, enhancing the culture and heritage offer, reducing crime and creating safer communities |

In our view, the most significant innovation in LCR's data governance to date is the pioneering Liverpool Civic Data Cooperative (CDC). A SIF-funded project with a focus on health and social care, the CDC represents a partnership of eight NHS Trusts, Liverpool Clinical Commissioning Group, and four higher education institutions: the University of Liverpool, Liverpool School of Tropical Medicine, Liverpool John Moores University, and Edge Hill University. It is working to develop an integrated data and digital innovation facility which supports collaboration between health tech partners and provides secure access to relevant data, while cementing trust from the public in how data is used. Accountability in data governance is safeguarded by elected members, and thus secured for the public within a 'diameter of trust'. Trustworthy national infrastructure is used wherever possible. Local communities work with NHS analysts, data scientists and health technology engineers to find new ways of improving healthcare and wellbeing, while citizen juries adjudicate on which datasets might be shared and which kept confidential. When it serves the public good, the CDC will also open public sector datasets to private market actors with particular computer data analytic capacities.

**Propagating LCR Civic Data Cooperative by extension or emulation**

We believe that the CDC model has the capacity to become a central governance mechanism in a new LCR data ecosystem by extending the model beyond health and social care, or replicating it by establishing CDCs for each priority challenge area.

All other activities should be aligned to ensure that each challenge-based CDC has the necessary support to succeed:

- Stakeholder mapping and data audits should be undertaken to identify which actors, data owners and datasets will prove most useful in addressing each challenge
- Technical support and investment in ICT infrastructure should be refracted to give priority to each CDC, including the scaling of Sensor City, the rolling out of 5G and the Digital Spine, digital industrialisation, skills strategies and business supports, access to and support for the Hartree/Daresbury Supercomputer, and the AI Solutions Lab
- A layered series of repositories for each challenge should be created, promoting Open Access to appropriate public and private datasets, with graduated, structured and secure access to other datasets depending upon the identity of the user and the intended purpose of the data analyst
- The principle of interoperability should be applied as deeply and widely as existing rules allow. CDCs should work to create a common reference architecture for LCR, agreed data protocols and standards, Open Data Interfaces (application programming interfaces – APIs), and an LCR Open Data Portal. A premium should be placed upon data cyber-security. CDCs might capitalise upon the governance framework and facilities provided by the Consumer Data Research Centre
- A robust business model for each CDC should be developed. Innovative new public sector, private sector, and public–private partnership (PPP) business models and financial instruments should be devised in which both the public and private sectors share the risks and rewards of computerised data analytic investments in proportion to their investments
- Cost–benefit analyses for each CDC should include positive and negative externalities and contingent valuation, and be based upon a total economic value (TEV) model.

Although self-determination and autonomy would be the key tenets, an LCR CDC (or federated series of CDCs) should be guided and aligned overall by an umbrella body. This body would be responsible for overseeing the development of a high-performing LCR data ecosystem. Given the paucity of resources and the as dearth of broad-ranging capacity in any single municipal and anchor institutions, and the quality of expertise distributed across stakeholders, CLs would need to pool local capabilities and co-create and co-resource an inter-sectoral and inter-institutional LCR Data Analytics Network (DAN). This coalition could assume responsibility for:

- Codifying a challenge-based approach to the acceleration of the innovation and adoption of computerised data analytics for public good in LCR
- Identifying one or two further critical challenges upon which to focus initial activity, beyond health, health inequalities and social care
- Scaling and applying the LCR Civic Data Cooperative by extension or replication to these other challenges
- Providing oversight of a challenge-based Smart LCR strategy and support for an LCR CDC or series of CDCs
- Helping CDCs build robust and sustainable business models
- Ensuring the scaling of existing pilot projects, securing future infrastructure investments and refracting support and and through CDCs
- Developing an overarching LCR reference architecture, and increasing dataset interoperability within and between CDCs
- Promoting collaborative and participative market design by CDCs, identifying and deploying local policy levers to address market failures in key challenge areas.

Finally, we see a role for an independent, Higher Education Institution (HEI)-led, inter-disciplinary, and inter-sectoral 'ideas catapult' to generate and pioneer innovative ideas and solutions with respect to data sharing arrangements and the redevising of local data markets.

**Ideas catapult for a new data stakeholder capitalism**: thought leaders in this catapult would enjoy the freedom to be creative and inventive in devising business models to enable public sector datasets to be opened up to private businesses in ways that produce and allocate shared value fairly. It would collaborate with CDCs to:

- Develop imaginative ways of affording private enterprise access to public datasets on terms which suit both parties[14], through judicious allocation of capital funds, setting terms for Public–Private Partnerships (PPPs), spatial planning and land use regulation, providing business supports, promoting a skills agenda and procurement policy, and setting ethical AI and data charters
- Create and enact new models of collaborative and participative market design to tackle long-standing market failures (such as insufficient incentive structures, unaccountable corporate monopolies, unconscious bias in algorithms and a lack of transparency, large-scale job displacement, and digital inequalities and addictions) to foster private investment, privacy and security of personal data, and transparency and safety in AI decision making
- Brief the LCR Smart Data Coalition and CDCs regarding national data strategies and regulatory policies which bear on local market failures, including the Department of Media, Digital and Culture's (DMDC's) new Centre for Data Ethics and Innovation and ongoing preparation of a National Data Strategy, the UK Industrial Strategy, the UK Government's Office for Artificial Intelligence, the Open Data Institute, the Cabinet Office's Geospatial Data Commission and new UK Geospatial Data Strategy, the Ada Lovelace Institute, the Alan Turing Institute, Lords Select Committtee on AI, and the Digital Strategy for Parliament.

**A roadmap for a challenge-based approach to building a smarter LCR**

Bringing together the ideas set out in our vision statement, we conclude by proposing a roadmap to guide future action:

**Short term:**

a) Civic leaders in LCR to convene and participate in a 'highest level' Data Summit, perhaps titled 'Unlocking the Value of Big (Local) Data in LCR' to signal decisive intensification in effort to build a high-performing LCR data ecosystem and to undertake a health check of the existing data landscape

b) Civic leaders in LCR to build sufficient institutional capacity to scale LCR's computerised urban analytics ambitions, by pooling distributed resources, assets and expertise, and co-creating and co-resourcing a new LCR Data Analytics Network (DAN)

c) Civic leaders and Data Analytics Network to agree a challenge-based approach to unlocking the value of Big (Local) Data

d) Civic leaders and Data Analytics Network to agree a limited number of priority challenge areas for initial action, in addition to the existing focus on health and social care – suggested challenges are decarbonising LCR and achieving net-zero carbon emissions by 2040, and promoting inclusive economic growth and full(er) employment.

**Medium term:**

e) The current LCR Civic Data Cooperative (CDC) to focus upon increasing access to health and social care data, either by extension (to other challenges) or emulation (replication and customisation for other challenges) to become the key institutional structure to oversee the unlocking of local datasets and datasets with local expression

f) Each Civic Data Cooperative to develop a robust and sustainable business model

g) Civic leaders, Data Analytics Network, skills programmes and Civic Data Cooperatives to collaborate to ensure existing and future technical assets, resources, capacities, infrastructures, and investments are refracted to support Cooperatives with delivering powerful data-driven solutions to priority challenges

h) HEI anchor institutions to utilise the Liverpool Good Business Festival to test the idea of creating an Ideas Catapult to pioneer new data sharing arrangements and new business models for data sharing based upon the concept of stakeholder capitalism. If feasible, such a catapult should be built and led by an inter-disciplinary and inter-institutional research team. It should develop and deliver an AI innovation diffusion programme to LCR businesses.

**Long term:**

i) Civic leaders, Data Analytics Network, and Civic Data Cooperatives to work to co-create an LCR reference architecture and to maximise interoperability within and between Cooperative datasets. Within municipal authorities and anchor institutions, data generation practices to be gradually and incrementally aligned over time with this reference architecture

j) A virtual cycle to be cultivated in which the data analytics network develops a strategic framework for the advancement of computerised data analytics in LCR and nurtures Civic Data Cooperatives. Civic Data Cooperatives to drive forward the search for data-driven solutions to critical LCR economic, social and environmental problems. A research-led Ideas Catapult to ensure that Data Analytics Network and Civic Data Cooperatives are exposed to the latest thinking and learning on innovative governance mechanisms for data sharing, and collaborative and participative market design.

14 A good example is Red Ninja's work on emergency vehicles prioritisation https://www.gov.uk/government/case-studies/red-ninjas-smart-tech-clears-the-road-for-ambulance-crews.

# 6. CONSULTATION QUESTIONS TO CLARIFY NEXT STEPS

We invite responses to all or any of these questions to inform our thinking on – what next?

1. Should public sector datasets be opened and shared widely with public, private and third sector organisations and citizens?
2. How can civic leaders in LCR increase access to and enable better harvesting of local data? What limits should be placed on data sharing?
3. How can smart technology help to solve the key economic, social and environmental challenges LCR faces? Which challenges is it best suited to address? Can you give a practical example of how it might help?
4. Does the 'health check' outlined in this Position Statement provide an accurate assessment of the current status of LCR's data ecosystem? Are there any key initiatives, assets, or future plans which are missing from this health check?
5. Which international good practices identified in the Good Practices Reference Guide have greatest relevance to LCR and why? Which cities should we be learning from, and with respect to which problems/solutions?
6. What ethical issues most concern you about organisations sharing data more widely?
7. Is a highest level 'stock taking' and 'future planning' LCR Data Summit needed? Who should be responsible for convening such a Data Summit, and which civic leaders and anchor institutions should be invited to attend?
8. Does LCR have the expertise and capacity to build a Smart City Region? What are the gaps or points of weakness?
9. How might distributed expertise and capacity across LCR's public, private and third sector bodies be identified and harnessed? What might prevent effective collaboration?
10. Should any single institution take responsibility for the LCR Smart City agenda? If so, who should this be?
11. Is the idea of an inter-sectoral and inter-institutional Data Analytic Network a good one? Is it likely to be effective?
12. Should the model of the Civic Data Cooperative become the cornerstone of LCR's approach to data sharing? What alternative models of data sharing exist and how might they be helpful?
13. Does LCR need a single City Region Open Data platform or portal?
14. How might dataset interoperability be fostered across LCR datasets? What would increased interoperability enable you and/or your organisation to do?
15. In what ways are AI, computerised data analytics and data-driven solutions failing people in LCR? Who should be responsible for rectifying these failures, and what actions should they take?

Responses can be sent to:
Heseltine Institute for Public Policy, Practice and Place: heseltine@liverpool.ac.uk.

## ABOUT THE HESELTINE INSTITUTE FOR PUBLIC POLICY, PRACTICE AND PLACE

The Heseltine Institute for Public Policy, Practice and Place is a non-partisan, internationally recognised University of Liverpool research institute, bringing together expertise from across the University and policy communities to co-create, impact upon and influence public policies for tomorrow's cities. We do this by:

- Undertaking, funding and publishing fundamental and applied research
- Providing thought leadership, consultancy and advice
- Convening and hosting events, including conferences, policy provocations, workshops and seminars
- Building and strengthening academic–practitioner networks
- Developing capacity-building and providing training courses
- Providing an evidence base to have impact on public policy.

This Heseltine Institute Position Statement, the accompanying Good Practices Reference Guide, Briefing Notes on key case studies examined herein, and a special international symposium held in March 2020 represent the Heseltine Institute's contribution to unleashing the potential presented by Big Data in the Liverpool City Region whilst preserving and championing 'tech and data management for public good'. Through our activities and outputs in 2020, we hope to contribute to the clarification of what a citizen-centred Smart Liverpool City Region might look like, and how we might build it.

Outputs can be downloaded from the Heseltine Institute website: **www.liverpool.ac.uk/heseltine-institute.**

To cite this Position Statement:
**Boyle, M., Jarvis, S. Singleton, A. & Jones, A., 2020. *A Position Statement: Building a Data Ecosystem in Liverpool City Region to Unlock the Value of Big (Local) Data.* Liverpool: Heseltine Institute, University of Liverpool.**

Further details about the Heseltine Institute can be obtained at:

Heseltine Institute for Public Policy, Practice and Place
University of Liverpool
1-7 Abercromby Square
Liverpool
L69 7WY

Web: www.liverpool.ac.uk/heseltine-institute
Email: heseltine@liverpool.ac.uk
Tel: +44 (0)151 795 2485

Acknowledgements
Thanks to Mark McKellier for copy-setting and Fran Elkington with preparation of the final manuscript.

Copies of the report can be accessed at:

www.liverpool.ac.uk/heseltine-institute

THE UNIVERSITY IS A MEMBER OF THE ELITE
RUSSELL GROUP OF RESEARCH-LED UK UNIVERSITIES

PRODUCTION: UNIVERSITY MARKETING COMMUNICATIONS
PLEASE RECYCLE THIS DOCUMENT