ORIGINAL RESEARCH

Real Estate Data Marketplace

Philip Treleaven[1] · Jeremy Barnett[1,2] · Andrew Knight[3] · Will Serrano[1]

Abstract

Real estate financing, construction, and management are being revolutionised by the data science and infrastructure technologies of Artificial Intelligence (AI), Internet of Things (IoT), Big Data, Digital Object Identifiers (DOI) and Blockchain. The key to automation and integration is a Real Estate Data Marketplace across the lifespan of finance, planning, construction, regulation, management, even refurbishment and demolition. Integration of digital 4.0 and industry 4.0. Real estate is becoming increasingly digital, automated and integrated; embracing connected-buildings, wearable personal devices, integrated transportation systems, buildings and infrastructure data management and smart cities. The key to promoting the use of data-driven techniques in the real estate sector is the introduction of a Data Marketplace that promotes validation of data and tokenisation to create value for the owners of the data, Blockchain security and governance for users of the data, the application of Artificial Intelligence for data analytics and ethics for governance. The technical capability exists today for a complete end-to-end reshaping of the real estate market based on data and ethics. This is data-driven real estate (cf. Google, Amazon, Alibaba, Tencent). The contributions of this review paper are fourfold: (a) to provide a vision of a fully digital real estate marketplace; (b) to brief the industry on digital technology and also stimulate engagement with the computer science community; (c) to encourage debate amongst professionals, academics and politicians of the benefits of full automation and integration of real estate; (d) present ethics for the Real Estate Data Marketplace based on lawfulness, fairness and transparency.

Keywords Data Marketplace · Blockchain · Smart contracts · Real estate

1 Introduction

The data science 'tsunami' led by the FANGS (Facebook, Amazon, Netflix and Google) is radically changing all areas of life including work, leisure, business and education. In many areas, professionals say they do not know what their industry will look like in the next 10 years. In real estate,

✉ Will Serrano
 w.serrano@ucl.ac.uk

 Philip Treleaven
 p.treleaven@ucl.ac.uk

 Jeremy Barnett
 jeremy.barnett@ucl.ac.uk;
 jeremy.barnett@resilience-partners.co.uk

 Andrew Knight
 aknight@rics.org

1 University College London, London, England

2 Resilience Partners, London, England

3 Royal Institution of Chartered Surveyors, London, England

the key technologies—namely Big Data, Internet of Things, Artificial Intelligence, behavioural/predictive analytics, and Blockchain technologies, offer the potential to transform the way in which the entire sector operates. It is the combination of data analytics with a distributed ledger, smart contract and tokenisation technologies that can transform how real estate assets are designed, operated managed and transacted.

By the end of 2020, more data were generated than in the past 5000 years combined and this will rise tenfold in less than a decade [1]. Trading data will boost innovation and create new revenue streams of data that otherwise remain untapped. Although there have been significant reductions in the cost of data collection, the cost of selling verified and relevant data has become expensive. Security is also a key factor if organisations are to exchange confidential information. Most companies underutilise most of the IoT data that are currently collected [2]. For example, one oil rig with 30,000 sensors examines only 1% of the data collected.

Rifkind described IoT as being made up of a Communications Internet, an Energy Internet and a Logistics Internet

that work together in a single operating system [3]. Since Rifkind described his vision of the use of IoT to develop 'the collaborative commons and the eclipse of capitalism, the market has been overtaken by two disparate movements described above, Big Data and the growth of cryptocurrencies. Rifkind's view of the single operating system however remains valid if it is forged together with a fourth Internet, the Internet of data, which will allow a true digital marketplace to be created.

The most effective use of data is by the creation of a Data Marketplace, which is a platform that connects providers and consumers of data sets, and data streams, which ensure high quality, consistency and security. Data suppliers can authorise the marketplace to licence their information on their behalf and ensure compliance with strict terms and conditions. Consumers can also offer to rent or sell their own data back to the marketplace. Third parties can offer tools to analyse or make predictions from the data. The complexity of building such a marketplace should not be underestimated. Microsoft retired the Azure DataMarket which was a cloud-based Master Data Solution that attempted to create a trusted view of critical business data after 7 years. Fundamental roadblocks preventing full potential have been identified as:

- Most data are unstructured;
- There is no consensus around methodologies for structure;
- Pricing models for data are complex and depend on a number of variables including the accuracy, extent, content, age and degree of commercial sensitivity.

1.1 Real Estate Data Marketplace

In a marketplace, a true market for the data will be created as providers gain expertise in packaging and managing their data in a manner that will attract interested buyers. Data can be aggregated, limited for use by non-competitors, synthetic or anonymous. Sellers will be able to set constraints around the purpose for which data have been provided, the length of time that it can be accessed, and the quality of the data by smart contracts which can be written to govern access to the data (Fig. 1).

There are several barriers that will have to be overcome if this has become widespread in commercial real estate. Many market participants such us landlords or investors benefit from the status quo around the opaque nature of data about the exclusive rents that have been achieved, which is one of the key metrics used in the valuation process. These participants use commercial sensitivity as a reason not to publish that data and are often in a position to take advantage of what is affect a private market, using information asymmetry to acquire mispriced assets.

Computational Statistics - computationally intensive statistical methods.

AI Algorithms - mimicking a new form of human learning, reasoning, knowledge, and decision-making

- Knowledge or rule-based systems
- Evolutionary algorithms
- Machine learning

Fig. 1 Algorithm taxonomy [7]

The use of a fast-scalable infrastructure that supports micropayments by the use of Blockchain is required to support the security, transparency and ethics of such a marketplace. Examples of such an infrastructure are provided by the DX Network (https://dx.network/), which is an open market project based on Ethereum and a data broker (https://databroker.online/). The use of such a Distributed Ledger Technology (DLT) network or that proposed by IOTA could record not only the validity of the data but the purpose for which it was purchased and subsequent use, in order to ensure that the owner is properly remunerated for the transaction.

1.2 Real estate technology

AI technologies are already being used in real estate to make sophisticated and accurate forecasts. Startups such as SkylineeAI (https://www.skyline.ai/) have built large datasets with traditional data points such as rent, vacancy and capitalisation rates, with alternative data points such as geospatial information, mobile usage and satellite images. Constancy in deal transparency is a key benefit, as the amount of input data and complexity accumulates AI is seen as the tool to interact and manage Data Marketplaces. AI is now being used to pinpoint trends and direct investors towards postcodes where values are higher, bringing unprecedented speed and insight to market evaluation [4]. The insurance sector is poised for dramatic change as data analysis can improve risk identification, business interruption and improved health and safety/best practice compliance [5].

Major automation programmes are underway across the real estate market, notably finance, building information modelling (BIM), computer-aided design, and autonomous machines:

- *Development finance* FinTech and aptly named MadTech (confluence of marketing, advertising and technology) centres on computer programs and other technology used to automate financial services. In real estate, an example is investment platforms for developments.

- *Building information modelling (BIM)* BIM involves an intelligent 3D model-based process that gives architecture, engineering, and construction (AEC) professionals the insight and tools to more efficiently plan, design, construct, and manage buildings and infrastructure [6].
- *Digital twin* A digital representation of an entity using deep tech to create living simulation models. By bridging the physical and the virtual world, data are transmitted seamlessly to allow the virtual entity to exist simultaneously with the physical entity. Digital twins are currently being developed to work at various levels of granularity from a single asset, through development, portfolio, city, region and nation.
- *Computer-aided design (CAD)* Software is used by architects, planners, engineers, and others to create precision technical specifications and three-dimensional (3D) models.
- *Industry 4.0* The automation, integration and data exchange in manufacturing technologies.
- *Autonomous machines* Use of autonomous robots; self-driving machines used to transport materials across the worksite and to automate construction.
- *RegTech* Automating collection and use of statutory compliance data.
- *PropTech* The use of data from construction, asset management, facilities management and real estate transactions.
- *ConTech* Bringing technological innovation to the way construction is planned, designed and built. Early versions include 3D simulations and modern methods of construction.
- *PlanTech* Using a 360° view of digital transformation in the built environment to promote innovation in land use and transport planning and the development of sustainable urban neighbourhoods.
- *Algorithmic dispute resolution* The use of Artificial Intelligence (AI)-based technology for automated dispute resolution in real estate and construction; prior to resorting to the civil justice system.

1.3 Real estate market

Real estate is becoming an ever broader asset class with the traditional definitions of residential, office, retail and industrial being extended to include various 'alternatives' classes such as logistics, self-storage, data centres, co-working as well as some forms of infrastructure and other utilities. In many cases, assets are being built and/or re-developed as explicitly mixed-used combining office, retail, residential, etc. Whilst owner-occupied residential properties remain a huge asset class in many jurisdictions, significantly larger in value than commercial real estate, various forms of rental tenure offered by both private and institutional investors represent a significant and growing sector with the American multi-family model increasingly being developed in other jurisdictions as buy-to-let. Demographic trends have also driven the various forms of senior living developments that continue to attract investors.

Regarding the so-called real estate lifespan, this includes finance, planning, regulation, design, construction, management, and even dispute resolution, refurbishment and demolition. In transforming real estate, as with other business sectors, disruptive digital technologies group into Data, Infrastructure Technologies and Robotics.

2 Data science

The real estate marketplace is undergoing the same 'revolution' that has transformed financial services, retail, healthcare, and regulation. Big Data—the collection in real-time of huge quantities of increasingly heterogeneous data; financial, business, economic, social media and alternative data. AI analytics—the use of sophisticated self-programming machine learning models. Digital marketplaces are online platforms providing an automated and integrate 'ecosystem', linking all stakeholders.

In reviewing the data science technologies transforming real estate, we divide them into data technologies, analytics technologies and infrastructure technologies.

The data technologies include:

- *Big Data* The collection, storage and processing of increasingly huge and heterogeneous data sets that facilitate the real-time management of, for example, people, energy, buildings and infrastructure.
- *Data models* A common data model for encoding real estate data.
- *Markup* A computer notation that uses tags to define data elements within a file. The two types are (a) markup languages such as comma-separated Values (CSV), extensible markup language (XML) and JavaScript object notation (JSON); and (b) data exchange formats enabling translation, such as fast healthcare interoperability resources (FHIR).

The analytics technologies include:

- *Algorithms and Artificial Intelligence (AI)* Algorithms for analysing data cover three broad domains: computational statistics (e.g. Monte Carlo methods), AI (e.g. artificial neural networks), and complex systems (e.g. agent-based systems).
- *Federated learning* A machine learning setting where the goal is to train a high-quality model with training

data distributed over a large number of distributed data sources, to preserve privacy.

- *Analytics applications* Applications include forecasting, optimisations, production and robotics; as well as natural language processing (NLP)—the analysis and synthesis of natural language and speech; and sentiment analysis—using NLP, statistics, or machine learning methods to extract, identify, or characterise the sentiment content of text or speech.

The infrastructure technologies include:

- *Internet of Things (IoT)* The inter-networking of 'smart' physical devices, vehicles, buildings, etc., that enable these objects to collect, exchange and automate data.
- *APPs and chatbots* An application that performs a specific task, especially as downloaded by a user to a mobile device or computer; increasingly using natural language for interaction.
- *Digital object identifiers (DOI)* A DOI is an identifier or handle, potentially persistent, used to identify objects uniquely, standardised by an international body (DOI, 2015; DOI, 2020).
- *Computable contracts* Computer-readable, understandable and executable legal rules.
- *Tokenisation* The process of issuing a blockchain token (generally refers to a 'Security Token' which is tied to an underlying asset or currency and requires regulatory compliance to ensure trust) that digitally represents a real tradeable asset.
- *Blockchain technologies* Including Distributed Ledger Technology (DLT), distributed databases that secure, validates and processes transactional data; and Smart contracts, a self-executing contract with the terms of the agreement between buyer and seller directly written into lines of code.
- *Edge technologies* Computing is done at or near the source of the data, instead of relying on the cloud at one of a dozen data centres to do all the work.

2.1 Data technologies

As discussed, the major trend across all sectors is the collection, storage and processing of increasingly huge and heterogeneous data sets; historic and real-time. Data include financial, business, economic, social media, and what is referred to as alternative data. This facilitates the real-time management of, for example, people, energy, buildings and infrastructure.

This 'data-driven' trend is mirroring what has already occurred in financial services (e.g. capital markets) [7] and is underway in retail and healthcare. To make increasingly

comprehensive data sets available for analytics, real estate will need common data models and markup.

2.2 Analytics technologies

The terms algorithm, Artificial Intelligence (AI) and machine learning are used interchangeably. However, data science algorithms cover three broad domains: computational statistics (e.g. Monte Carlo methods), Artificial Intelligence (e.g. artificial neural networks), and complex systems (e.g. agent-based systems) [7].

2.2.1 Computational statistics

Computational statistics models refer to computationally intensive statistical methods including resampling methods (e.g. bootstrap and cross-validation), Monte Carlo methods, kernel density estimation and other semi- and non-parametric methods and generalised additive models.

2.2.2 AI and machine learning

AI algorithms are a continuum of epistemological models spans three main communities:

- Knowledge-based or heuristic algorithms—where knowledge is explicitly represented as ontologies or IF–THEN rules rather than implicitly via code (e.g. rule-based);
- Evolutionary or metaheuristics algorithms—a family of algorithms for global optimisation inspired by biological evolution (e.g. genetic algorithms, genetic programming, etc.); and
- Machine learning algorithms—a type of AI program with the ability to learn without explicit programming and can change when exposed to new data; mainly comprising supervised, unsupervised, and reinforcement learning.

2.2.3 Complex systems

Lastly, a complex system is any system featuring a large number of interacting components (e.g. agents, processes, etc.) whose aggregate activity is nonlinear (not derivable from the summations of the activity of individual components). Examples include cellular automata, agent-based models, network-based models, and multi-agent systems.

Three important analytics application areas are:

- Natural language processing—the application of computational techniques to the analysis and synthesis of natural language and speech;
- Sentiment analysis—the process of computationally identifying and categorising opinions expressed in a piece of text used for behavioural analysis;

- Behavioural/predictive analytics—the analysis of large and varied data sets to uncover hidden patterns, unknown correlations, preferences, etc., to help make informed decisions.

2.3 Infrastructure technologies

For the marketplace, the Blockchain + DOI layers support information management both for a single building and across an infrastructure. The federated learning layer support privacy-preserving analytics across distributed data sets. The computable regulations/contracts support automation by encoding rules on the operation of the infrastructure. Finally, the top layer can even support mobile apps for an Open Real Estate initiative (cf. UK Open Banking). Figure 2 shows a technology stack for a real estate marketplace infrastructure, integrating digital and physical elements.

2.4 Robotics

Robots and drones are becoming an increasingly familiar aspect of life and work in the UK, playing a role in areas ranging from emergency services to oil and gas and construction. They can be used for surveys and to transfer data quickly from places that are hard to reach and can be used to manage costs, control risks, increase safety and influence outcomes. PWC predict that by 2030, there will be 76,000 drones in UK airspace and predict substantial cost efficiency and productivity benefits by this increase in adoption [8].

Researchers at Harvard's WYSS Institute for Biologically Inspired Engineering [9] have built small construction robotics programmed to work together as a swarm-based on the habits of termites. The four-wheeled robots can build brick-like walls by lifting each brick, climbing the wall and laying the brick in an open spot. They have sensors to detect the presence of other robots and rules for not getting in each other's way.

The advantages of using this technology are:

- Building topology optimised structures that are difficult to manufacture.
- Design can be 'gamified' to find the best and most beautiful structures to be conceived
- Can operate in austere environments and could be useful responses to natural and manmade disasters.

Recent examples of the use of this technology are buildrone.co, which uses flying robots designed to detect and repair leaks in pipelines. Fibrebots is a digital fabrication platform which uses a combination of cooperative robotic manufacturing with sophisticated material architectures. The swarm of robots are designed to wind fiberglass filament around themselves to create high-strength tubular structures which can rapidly create architectural structures [10].

Fig. 2 Real estate technology stack
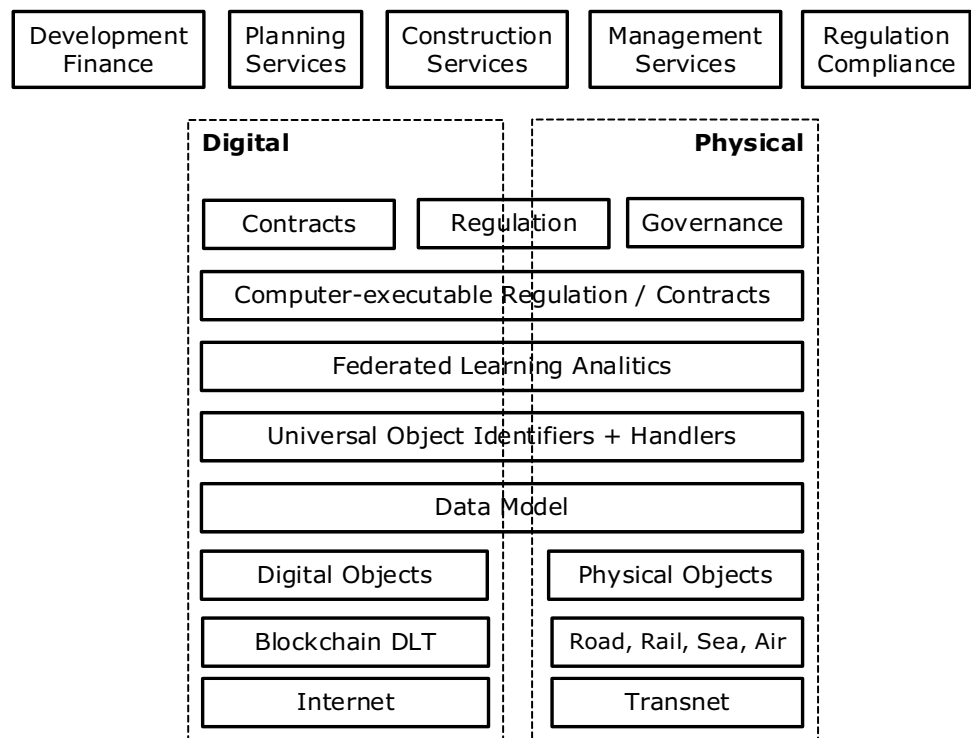
The flying of drones is highly regulated in the UK where the Civil Aviation Authority have released Dronecode, which emphasises that flight should be below 400 feet, within sight of the operator and never near aircraft [11].

3 PropTech market

The use of emerging technologies has resulted in the development of specific technologies, now referred to as 'Deep Tech'. Deep Tech refers to fundamental breakthroughs in science and engineering that have the potential to profoundly impact industries and people's lives. This forms a matrix of application areas/sectors such as: FinTech/InsureTech, RegTech, HealthTech, PropTech, LawTech, EdTech, FoodTech/AgriTech, BioTech, SafetyTech; even MadTech (marketing and advertising technology) and key technologies, such as data science.

PropTech and real estate FinTech is a fast-moving sector within a conservative industry, which is rooted in traditional business practices. Real estate is the largest asset type and one of the last to embrace innovation:

- Real estate FinTech: describes a wider field of technology-based platforms which facilitate the trading of real estate ownership. The assets can be buildings, shares or funds invested in this real estate, debt finance and ownership rights. Other platforms that focus on the 'built environment', involving architects, engineers, construction firms and facilities managers are describes as ConTech (Construction Tech) [12].
- PropTech: Baum defines PropTech as being about information, transactions and management. PropTech is about management and operation of buildings, and the use of digital passports to improve speed and efficiency of transactions, in what has been identified as Know Your Asset KYA [13].
- Prop RegTech: building research establishment environmental assessment method (BREEAM) was first published by the Building Research Establishment (BRE) in 1990. It is the world's longest established method of assessing, rating and certifying the sustainability of buildings. The Hackitt report following the Grenfell Tower Enquiry has proposed a new regulatory framework and the introduction of a new Regulator with strict enforcement powers for those who cannot demonstrate compliance [15].

The global real estate market is worth $217 trillion according to Savills 2016 and the World Bank, of which only 25% relates to commercial property. Annual real estate trading has averaged $636 billion annually since 2007 and reached £900 billion in 2015 [13]. This is a large illiquid asset class which has the potential to reduce transaction costs and illiquidity. The assets can be buildings, shares or funds invested in this real estate, debt finance and ownership rights. The digital platforms have the potential to provide diverse information to a host of interested parties, from the existing owners and occupiers to prospective purchasers.

The global construction sector is the second-largest industrial sector valued at $8.6 Trillion p.a. [14], but is described as being dysfunctional [15] with construction disputes running at 10% [16]. Tech investment suddenly boomed for the construction industry in 2019. The first 6 months of the year saw a 30% increase in investment compared to the entirety of 2018, with a total of more than $1 billion spent on new technologies [17].

Although the use of 'Deep Tech' for regulatory compliance is in its infancy, public concern around the safe design of residential buildings following the Grenfell Tower disaster has resulted in calls for a new approach to Regulatory Compliance. The current RegTech market is based upon high compliance costs and the growing demand for regulatory intelligence in the financial services sector, where rising penalties for non-compliance have fueled growth, together with a demand for real-time compliance analysis for regulators and businesses. It is anticipated that similar growth will occur in real estate for reasons set out below.

4 Digitalisation of real estate assets

The digitisation of assets is now gaining the interest of institutional investors around the world. Germany has launched the Boerse Stuttgart Digital Exchange (BSDX) becoming first regulated exchange for digital assets which will eventually allow European retail and institutional investors direct access to digital assets and provide flexible and relatively low-cost trading. Closely followed by SIX Digital Exchange (SDX) in Switzerland where a prototype exchange has been delayed due to regulatory and legal concerns. In the UK, the IPSX is the world's first regulated securities exchange dealing with both Initial Public Offering (IPO) and secondary market trading. Placetech publicised the floatation of the IPSX's first asset, The Mailbox in Birmingham, which is a large mixed-use development including hotels, retail, shops and restaurants [18]. 20|30 have built a Blockchain equity issuance platform that was accepted into the (Financial Conduct Authority) FCA Sandbox in 2018 which aims to enable companies to raise capital by issuing equity tokens.

It is important to recognise that in considering the impact of data-driven real estate that two concepts need to be addressed:

- *Digitisation* This is the conversion of analogue information into digital form. Much of the information that will be considered below for Property Reg Tech is currently not collected, but exists in emails, phone conversations, meeting notes and correspondence.
- *Digitalisation* This is the actual process of the technologically induced change by using digital technologies to change a business model.

Companies and organisations must transform their cultures and data architectures to adapt to the new technologies, new types of data new governance and compliance regulations if the vision set out in this paper is to come to fruition. Due to the multiple formats that critical information is now generated for the construction and real estate sectors, the authors believe this can best be effected in a DLT form infrastructure layer. Data that are arranged in either manner can be used to allow tracking of title, identification of risks, eradicate title fraud and improve land surveying accuracy. It can also be used in conjunction with smart contracts to automate processes as set out below.

4.1 Tokenisation

Tokenisation refers to both the fractionalisation of property rights or it can refer to the digital representation of asset ownership. The former Chairman of Nasdaq, Robert Greifield, predicted in 2017 that by 2022, all stocks and shares will be tokenised [19]. Tokenisation of real estate assets will permit increased liquidity in the real estate capital markets and raises serious issues about the funding and ownership of infrastructure assets. The 2020 report of the Said Business school into tokenisation of real estate [13] concluded that, although tokenisation offers exciting possibilities for the real estate investment market, there is a danger that innovation will be set back if attention is focused solely on digital fractionalisation of single assets when funds and debt may offer more attractive opportunities in the short term.

A token can be a store of value, a unit of payment, a unit of account, used as an input to a smart contract and carry a payload. There are at least 61 definitions of a token [20]. Tokens can be unique or identical, active or passive, open or closed, fungible or not, transferable or not, and exist on a substrate such as Ethereum. The ERC-20 token has emerged as the technical standard for all smart contracts on the Ethereum Blockchain for token implementation [21]. This paper proposes a novel use for tokens, which is as a method of capturing and trading individual data sets as a payload to a tradeable ERC20 compliant token.

4.2 Know your asset

The existing purchase and sale transaction process for commercial real estate is complex and involves multiple steps. Identification and closing a transaction can be expensive and time consuming for both parties due to:

- Inefficient and expensive property search processes;
- Time-consuming paper-driven predominately offline due diligence processes to assess: financial, environmental and legal issues which affect price;
- High title and related costs;
- Slow expensive and opaque financing mechanisms and payments.

In their Smart Buildings Report [22], concluded that the mantra 'Location Location Location' had been replaced by 'Location, Information, Analytics' as the effects of IoT enabled building management systems improved user experience using techniques such as the Information Value Loop, a consulting term for continual improvement. Deloitte has refined their view, now describing 'Location, Experience, Analytics' defining 'Smart Buildings' where mobile apps are created to create personalised tenant/end-user experiences. IoT can be used to minimise risk and increase investment potential by prioritising smarter systems for energy, security, parking and predictive maintenance.

Data governance following the General Data Protection Regulation (GDPR) restrictions in Europe is assuming more importance as companies seek to capture large unstructured data sets that are being generated.

Royal Institute of Chartered Surveyors (RICS) has developed a data standard to support both the definition of common terms as a real estate taxonomy and to allow digital data transfers using XML, JSON and CSV. The data standard provides full support for due diligence across land and property including cost across the full lifecycle, measurement and property/land attributes, valuation, legal, cadastral and geospatial data, tenure, lease/contract terms, etc. The Data Standard is freely available in the public domain and has been developed to support the various written standards currently used around the world such as International Valuation Standards (IVS) for valuation, International Land Measurement Standard (ILMS) for overall due diligence, International Property Measurement Standard (IPMS) for property measurement and International Construction Measurement Standards (ICMS) for lifecycle costing. The Data Standard supports the concept of data provenance to allow the source, recency, reliance, etc., of data to be captured.

5 Automation of the real estate marketplace

It is relatively easy to speculate on a future data-driven real estate sector since there is already considerable discussion and work underway on Blockchain finance platforms for real estate [23], IoT, 5G [24] and Smart Cities [25]. On this article, we look at the impact of the data technologies on the different areas of real estate.

Blockchain is an ideal tool for those who wish to demonstrate compliance with new regulatory regimes such as fire safety post Grenfell Tower as follows:

- As a state transition machine, it can be used in conjunction with a gateway system of project management
- Permit regulators to monitor remote locations;
- As a tool to verify building control;
- Incentivise behaviours through tokenisation;
- Provide a trusted, transparent and independent store of information for use by all parties to the construction of a building, and also the owners, occupiers and investors in real estate assets.

5.1 Finance

There are a number of Real Estate FinTech companies who are concentrating on this emerging sector:

- StreetWire, a real estate Blockchain services company, is working to innovate and streamline the process and management of transactions using Blockchain real estate company
- Insurami operate a deposit guarantee platform which helps landlords and tenants
- Midasium has created software for independent landlords and property managers to manage the cash flow of their property portfolios

5.2 Facilities management

Data and infrastructure technologies such as IoT, Artificial Intelligence and Blockchain will allow real estate assets to be managed digitally in real-time and automated. IoT has major implications when one considers that every device is likely to have an Internet connection and software to fully manage its operation. Facilities management (FM) involves a number of disciplines to ensure functionality, comfort, safety and efficiency of the built environment by integrating people, place, process and technology. The European standard for FM agreements (EN 15221) defines FM as 'the integration of processes within an organisation to maintain and develop the agreed services which support and improve the effectiveness of its primary activities'. ISO 41011:2017 sets the purpose as 'improving the quality of life of people, and the productivity of the core business'.

5.3 Design

Building information model (BIM) software and digital twin is a dynamic digital representation of a physical real estate asset, which provides the elements and dynamics of how an IoT device operates and lives throughout its lifecycle. Digital simulation models are created to update and change as the physical state of the underlying asset changes.

The Centre for Digital Built Britain (CDBB) is a partnership between the Department of Business, Energy and Industrial Strategy and the University of Cambridge to understand how the construction and infrastructure sectors could use a digital approach to better design, build, operate and integrate the built environment. The National Infrastructure Commission's report [26] recommended the development of a digital twin of Britain's infrastructure. The CDBB has published the first output of its Digital Framework Task Group, The Gemini Principles, to guide the national digital twin and the information framework that will enable it [27].

Algorithmic or computational design uses sets of instructions to perform certain tasks such as generating digital models and can be used to design structures that could not be drawn in the conventional sense by an architect. In addition, designers can present dozens of alternative solutions which would be impossible using traditional methods. BIM, parametric and machine learning can be used from product and interaction design, to large scale urban complexes and masterplans [28]. One technique is to share geometric data across several design models to create an aggregate model from discrete parts, so when the architect changes the original geometry, the data are automatically updated whenever it has been referenced. Further, algorithmic design facilitates digital fabrication so that manufacture can take place off site.

The use of algorithmic design is of particular use with Flex, which is the ability to change the use of the building from its initial design, a movement that has seen tremendous growth in recent months. One example is the project residential project where users will have the option to change the room configuration, creating their own living spaces [29].

5.4 Regulation and compliance

With the encoding of regulation as computable (computer-executable) specifications, it will be possible to automate compliance checking. In addition, when building regulations change it should be possible to identify real estate that is no longer compliant and needs upgrading. IoT Platform as a Service (PaaS) collects and process sensor data at scale to manage the building and device lifecycle from onboarding through to decommissioning [30]. In the view of the authors,

the introduction of a new regulator coupled with new obligations upon all those involved in the design construction management and operation of buildings (the 'duty holders'), coupled with the increased use of turnover based sanctions for regulatory breach, will prove the most fertile ground for the adoption of the emerging technologies.

- Environmental monitoring: BREEAM (Building Research Establishment Environmental Assessment Method) is the world's leading sustainability method for master planning projects, infrastructure and buildings. SimStock allows the automatic creation of dynamic thermal simulation models to analyse retrofitting systems, test for potential integration of renewable technologies, evaluate energy storage potential and map buildings in dense urban areas for air quality, particle pollution, and future climate condition [31]. The London Buildings Stock Model (LBSM) [32], commissioned by the Greater London Authority to tackle fuel poverty and improve the energy efficiency of the capital's housing stock contains data on every domestic and non-domestic building within the M25.
- Decommissioning: decommissioning of assets including nuclear facilities includes activities such as planning, physical and radiological characterisation, facility and site decontamination, dismantling, and materials management. The International Atomic Energy Agency (IAEA) maintains resources such as databases, which provide information on nuclear facility characteristics and inventories of spent fuel and radioactive waste.

Fire safety: The Hackitt Review into Grenfell Tower [15] concluded that the current system of building regulations and fire safety is not fit for purpose. The Grenfell Tower Inquiry: Phase 1 report made a number of recommendations for how owners and managers of high rise residential buildings should collect and retain data. On 20th July 2020 the draft Fire Safety Bill was published which set out the government's proposals for reform which include a new Building Safety Regulator which will be housed within the Health and Safety Executive. In the proposed forward, the Secretary of State for Housing, Communities and Local Government said,

'Industry should also prepare for the substantial changes that are coming. Many in the industry have already been working to improve standards and have shaped and influenced this Bill. Now they must accelerate this work to improve industry-wide culture, competence and compliance. The new Building Safety Regulator will work closely with the industry to drive a change in culture that prioritises residents and ensures their safety.'

It is also clear that regulators around the world are turning to substantial turnover based sanctions which it is thought will encourage better behaviour by the introduction of massive financial penalties.

5.5 Verification

Although the main benefits of using Blockchain generally considered to be trust and automation, it is of critical importance that any use of this technology in the built environment should include a robust system of verification of the material that is stored or linked to a Blockchain record, which later becomes the authoritative record of events. The truth of a fact or event is never absolute, but what is important is the quality of the verification that has been used to establish the fact, so that the user can make his or her own value judgment as to the weight that is attached to the information.

Verification can be achieved in a number of ways. A simple multi-stage sign off is often utilised in crypto currency by the Multisig which is a specific type of digital signature where two or more users to sign off transactions or documents.

RICS, and other professional bodies, provide various standards around valuation, due diligence, construction costing, etc., which require their registered members to produce data and provide professional reliance on its efficacy. In this way, a degree of assurance, provenance and verification can be obtained by having a clear understanding of the source of data used in any marketplace or individual transaction/process of due diligence. Currently, Facilities Management companies use enterprise asset management software such as IBM Maximo Asset Management to report on buildings, vehicles, fire extinguishers, maintenance schedules and workflows to manage the assets. These old legacy systems cannot be relied upon to verify the accuracy of the underlying information and are therefore of limited value to regulators and insurers who wish to rely upon information generated by third parties for statutory and other compliance.

For example, the legal system uses concepts of 'more likely than not' and 'beyond reasonable doubt' to establish a benchmark for reliance by courts and juries of evidence that is produced. Courts and tribunals grapple with the concept of 'reliability', which is never absolute, but when viewed objectively and subjectively, usually produces an acceptable standard. A simple technique that can be borrowed from the criminal law is the little-used concept of corroboration. Evidence of 'unreliable' witnesses has traditionally been accepted by juries when there is independent support for the proposition. It is also of critical importance when dealing with digital evidence to establish the provenance and audit trail, so that the accuracy of core information is verified in court, where the accuracy and detail of various drafts of the document are put under the forensic microscope.

5.6 Dispute resolution

Real estate and especially construction is notoriously litigious. Cases generally concern contractual obligations and compliance with those obligations. Regulatory oversight is increasing in respect of safety where it has become apparent that insufficient attention is often paid to continuing obligations to ensure, in particular compliance with fire safety best practice.

The use of DLT ledgers to create a digital passport of compliance data can encourage collaborative behaviour in forming an agreement about what data should be retained. It can also encourage good practice through what has been described as the 'Panopticon effect' described by the philosopher Jeremy Bentham, where a prison can be designed so that one guard can see into every cell, but no individual prisoner knows he is being watched (now made possible by CCTV).

Blockchain permits almost real-time settlement of trades which eliminates the risk of payment defaults and dispenses with the need for a counterparty, thus substantially reducing the cost of trades. An area already undergoing major innovation is Alternative Dispute Resolution (ADR), especially automated Online Dispute Resolution (ODR). In terms of LawTech [33], we broadly divide online dispute resolution into:

- *Consumer ODR* Use of technology by online suppliers and consumers.
- *Judicial ODR* Any hearing outside of the courtroom which uses technology.
- *Corporate ODR* Managing the resolution of global multi-stakeholder projects.

The core principle in a corporate ODR system or service is that parties engaged in a major commercial project should agree before the commencement of the project that any disputes will be referred to the system in the first instance. Algorithms search or a 'win win' using best alternative to a negotiated agreement (BATNAs) and worst alternative to a negotiated agreement (WATNAs) 'best/worst' alternative to a negotiated agreement and parties negotiating positions are scored by criteria that are established at the outset of the mediation.

Dew and Barnett designed an architecture for a regulator-led corporate ODR system, designed to resolve disputes around manufacturing and marketing of dangerous products [33] as shown on Fig. 3:

- This uses smart contracts to automate disclosure of critical documentation.
- Parties will be offered a 'win win' rather than risk the cost, delay and uncertainty of conventional litigation.

5.7 Why automate real estate

The current due diligence process of buying, selling and renting real estate is currently slow, expensive and favours market participants who have access to private data. This has led to a thinly traded, illiquid market for the asset class where price discovery, valuation, financial stability and investor participation is hindered. Whilst real estate investment trusts (REITs), commercial mortgage backed securities (CMBS), residential mortgage backed securities (RMBS) and various forms of real estate funds in many jurisdictions provide proxies to direct investment in real estate, albeit imperfect and not without significant risk both to individual investors and overall financial stability (see Global Financial Crisis), for retail investors, the size and heterogeneous nature of much real estate is a significant barrier to wider investment in the asset class. Improved availability of trusted data and the potential for fractionisation would undoubtedly support a more widely traded asset class with all the benefits we see in listed securities and increasingly in corporate bond markets as trading in these assets increasingly moves away from over the counter (OTC) onto exchanges with counterparty support. It should be recognised that real estate remains an essentially private market where the illiquidity premium, and the current information asymmetry, both favour existing market participants—the move to automate real estate will face behavioural barriers that dwarf the technical issues.

The advantages of data-driven real estate and digital automation are: (a) efficient fundraising and investment; (b) communication of documents and data amongst partners; (c) providing intelligent AI tools for supporting professionals; (d) automated checking of compliance and regulation; (e) automated dispute resolution; (f) efficient management of real estate; (g) providing additional services to users of the real estate via mobile apps, etc.

The disadvantages of digital automation: (a) require a higher level of skills and training from the stakeholders, and especially the construction sector; (b) raising the potential of system failure due to the introduction of sophisticated technology; (c) decline of companies who fail to invest in new technologies and automation.

6 Ethics in the Real Estate Data Marketplace

Ethics within the real estate comprises responsibilities to clients and customers, responsibilities to the public and responsibilities with colleagues within the same profession or professional institution. Real estate is the intermediary between buyers or sellers, investors or occupiers with competing interests against each other for rental or sale price. As responsible within this middle ground, ethics shall be driven
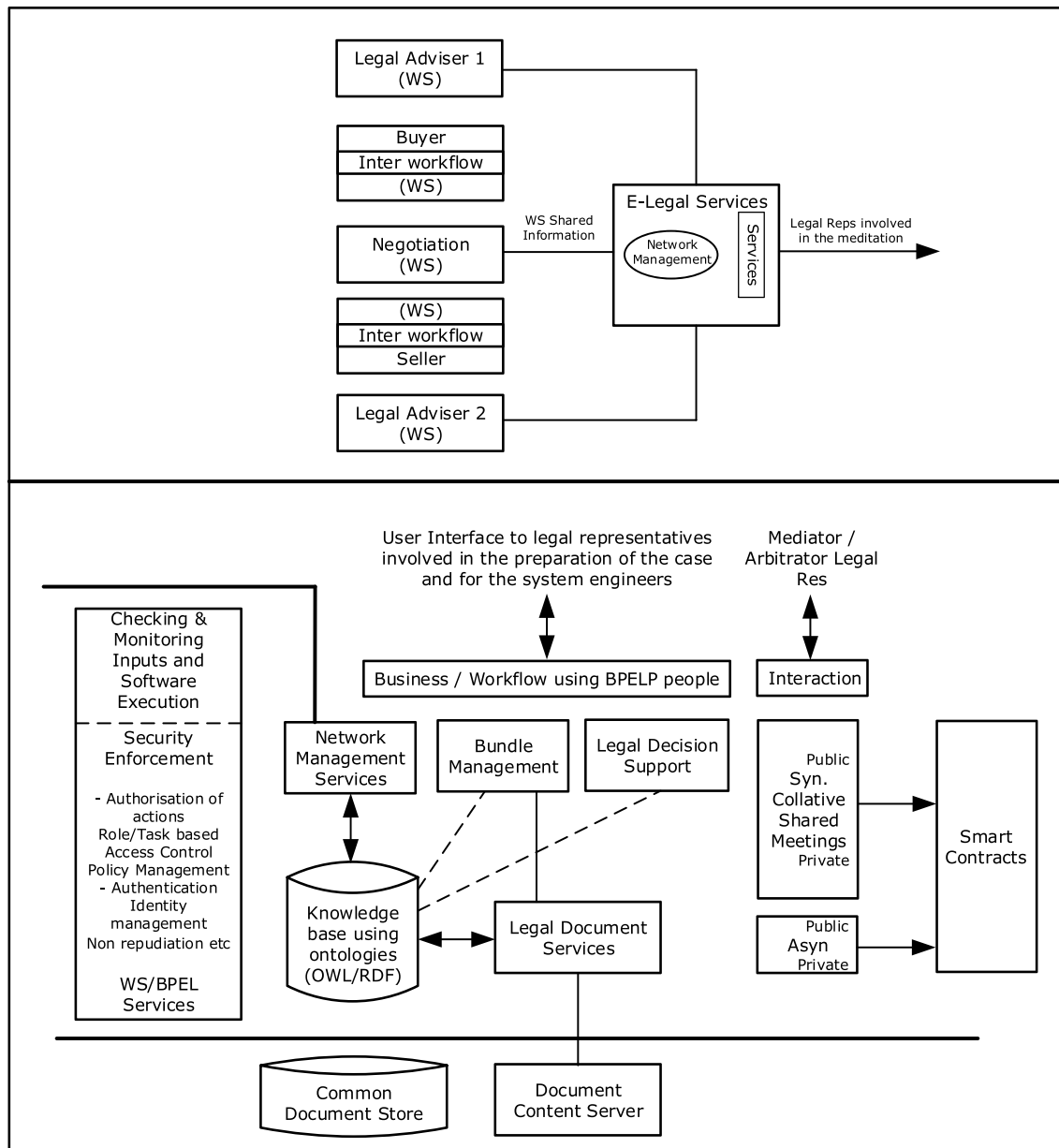
Fig. 3 Corporate ODR system [33]

and enforced by the real estate sector. Real estate professionals services shall not mislead clients or gain an unfair advantage over their customers. The given advice shall be objective, professional, unbiased by any personal motivation or potential benefit or advantage [34]. Real estate professionals shall avoid exaggeration, misrepresentation, or concealment of relevant facts relating to the property or the commercial transactions. This includes their honesty and truthfulness in their real estate communications where the true picture in their advertising, marketing, and other representations.

Like most professional bodies in the built environment, and indeed across professions generally, RICS sets standards and regulates its members and regulated firms not only with a set of technical standards such as valuation but just as importantly, a set of codes and ethical standards. With five high level principles around acting with integrity, always providing a high standard of service, acting in a way that promotes trust in the profession, treating others with respect and taking responsibility, RICS' ethics standards are underpinned by the International Ethical Standards (IES). In addition to member and member firms being expected to adhere to all appropriate laws, statutes, and other regulations in the jurisdictions in which they practise, they are also expected to implement these ethical considerations in every aspect of

their professional conduct. Technical standards also cover issues around data handling, record keeping which clearly impact their responsibilities around handing and assuring data.

A governance framework is key in any data pool or marketplace where participants provide or trade data. Governance shall control for what purposes data can be used; the rules and restrictions around sharing at different levels of granularity and around access right for different classes of market participants. The different methods to anonymised data shall provide aggregate information where privacy or confidentiality issues shall be mitigated. Governance shall also ensure the methodology used in data collection and provision by market participants is itself ethical and effective in quality and assurance terms.

Data Marketplaces ethics, or the ethical use of customer data, is based on the trust of the user in the system or the data custodian. Data Marketplaces shall be ethically driven from their creation towards a shared, inclusive and cross-functional responsibility [35]. Ethical requirements shall be embedded holistically within the technological approach. This includes the respect of the cognitive human ability to make their own decisions based on clear, factual, accurate or unmanipulated information. Customers shall be treated fairly with respect to their basics needs in terms of financial loss or privacy. Transparency in the way decisions are made by the Data Marketplace includes the intelligibility, explicability and verifiability of the conclusions and actions. Ethics shall be driven from the top of the Data Marketplace organisation to ensure is adopted across its different business, technical business functions and existing governance frameworks. Accountability shall resolve ambiguities and liabilities based on common agreement by all parties.

AI ethics in the Data Marketplace must be embedded within the design, development of the different decision processes and outcomes of AI algorithms [36]. Human requirements and features such as responsibility, transparency, audibility, incorruptibility, predictability also shall be applied to the AI algorithms that intend to replace human judgment. AI shall be designed to align with the rules and morals of human decisions based on right and wrong principles based on contextual information, experiences and memories. AI shall be designed for humans to easily perceive, detect and understand its decision logic that minimises bias and promote inclusive representation. Bias generated by bias shall not be embedded in the AI algorithms either in their training datasets. AI shall protect the marketplace user's data while preserving the users' rights over its access and storage. Users of the Data Marketplace shall be empowered with the control over their own private or generated data, interactions with other users and transactions within the Data Marketplace.

Data ethics within the Data Marketplace shall be defined within a framework that covers the management of the entire data lifecycle in terms of how data are collected, stored, shared and governed [37]. Data analytics in terms of how data are analysed shall consider the application of AI algorithms and machine learning. Data outcomes shall take into account the impact and actions generated from the insights of the data itself. The data ethics framework hierarchical layers shall include the ethical vision, the management infrastructure and architecture, the strategy to achieve the data ethics vision, data insights for efficient monitoring. Governance for the Data Marketplace shall comprise support, sponsorship, training and development about ethical data issues and the consequences around the use or misuse of data. Data recommendations within the Data Marketplace shall be relevant to the users needs, rather than biased by economic interest from the data owners or the Data Marketplace service provider [38]. Big data present challenges for human users as it cannot manually processed or assimilated. Algorithms that provide search, filter and reorder big data services shall adhere the same ethics as the data itself [39].

Ethics in cryptocurrency and smart contracts cover two different aspects. The first consideration is their position as an ethic enabler as they provide verification, transparency, authenticable and accountability in the stored information. A cashless society where all transactions are recorded supports the ethical use of the economic model and purpose of currencies. However, on the other hand, as cryptocurrencies provide inherent anonymity, they can be also be used in illicit transactions if they are unregulated. In addition, the value from cryptocurrency price is relative where the worth of these assets could become certainly insignificant due to hacking or the development of new cryptocurrency assets.

7 Data Marketplace for the real estate

This section proposes a Data Marketplace model for the real estate ecosystem to bring together its participants entirely into a single interface. The benefits that Data Marketplaces deliver to its different users have still not been implemented in the real estate. The proposed model covers the real estate from its creation and build through to its management and final decommissioning. The presented Data Marketplace follows the Royal Institute of British Architects (RIBA) Plan of Work Stages for clarity and industry alignment.

The Real Estate Data Marketplace is an online digital decentralised platform where different users such as data providers and data consumers are based on licenses, service levels agreements and governed by terms and conditions (Fig. 4). The Real Estate Data Marketplace autonomously incorporates new data in a structured and scalable format. This data is enhanced with metadata provided by Artificial Intelligence. Therefore, the proposed Real Estate Data Marketplace generates a valuable cycle of data usage where once

Fig. 4 Real Estate Data Marketplace framework
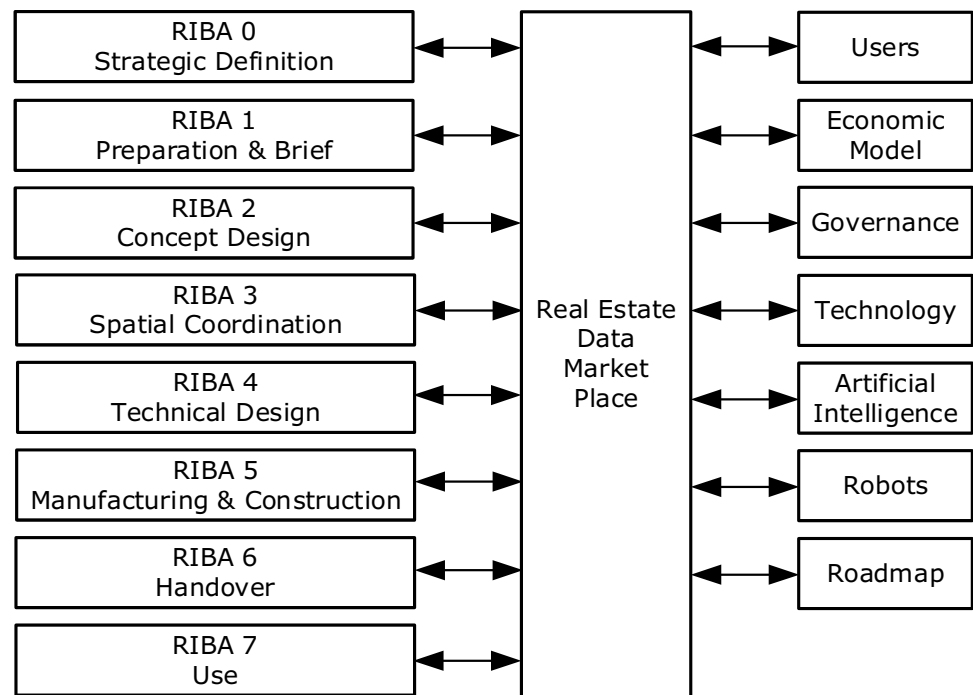
new data are securely collected, users are able to securely access to information.

The key functionality of the Data Marketplace is:

- Manages all data in a single decentralised platform: every information from its different sources is orchestrated and validated by the different Data Marketplace users via Blockchain and Distributed Ledger Technologies. This process enhances the quality and therefore value of the data.
- An automated data service catalogue of technical and real estate metadata that organises, describes and analyses data. This process is autonomously performed via Artificial Intelligence.
- Development of standards to define the metadata, for example, an abstracted lease can be understood with clear definitions of 'headline' versus 'effective' rent.
- Data and information visibility and traceability through the entire RIBA stages where origin, evolution, and meaning of data are tracked. Information is validated, verified, secured and governed with agreed service level agreements and service license agreements.
- The architecture is scalable in terms of the platform, users, data and services in a decentralised structure for high performance.

7.1 Data Marketplace users

The Real Data Marketplace users generate an ecosystem where users are data providers and data consumers. The intention of the Data Marketplace is to support the information flow between (1) different construction stages for the same project and (2) different project within the real estate. This approach enables design lessons learnt, optimises the construction delivery, reduces project risks and optimises the return or investment. The key value for the Data Marketplace is that unifies access to information therefore reduces time to find for information and to verify it via surveyors. Users from the Real Estate Data Marketplace cover the entire RIBA stages, although some users are more relevant at some RIBA stages than others (Table 1).

7.2 Data Marketplace economic model

The Data Marketplace business case is based on the trust of the data consumers information is validated and verified. Blockchain and Decentralised Ledger Technology (DTL) solutions build trust in every transaction where information is cryptographically validated by the ecosystem participants. The model defines trading between private entities on a private market when data are to remain confidential and also between public entities in a public market when data can be or must be open. The economic model of the Data Marketplace is a combination of revenue streams based on:
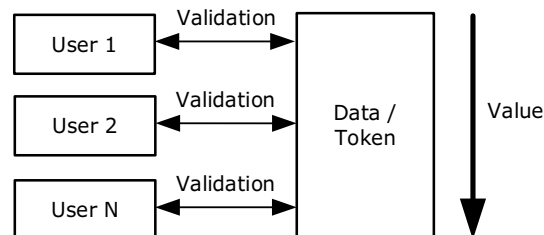
Table 1 Data Marketplace users

User	Responsible	Relevant RIBA stage
Investor	Capital of the project	0–1
Owner	Ownership of the Real Estate	0–1
Lawyer	Terms and Conditions and service license agreements are met for the Data Marketplace and the Design and Construction project	0–7
Real Estate Developer	Develop the Project such as budgets, Scope, Business Cases, planning permission, Feasibility studies, project brief	0–2
Insurer	Insure the project and the Real Estate	0–1
Project Manager	Management of the design, procurement and construction of the project. This includes the project programme	2–6
Designer	Design of the Real Estate according to relevant Standards. This includes Requirements, Risks, Project Execution plan, Health and Safety	2–4
Cost Consultants	Manage to the project cost	2–5
General Contractor	Deliver the construction, procurement, installation and commission of the project	5–6
Surveyor	Physical assessment the Real Estate including surveys for third parties like insurers	6–7
Agent	Commercialisation of the Real Estate	7
Maintainer	Maintenance of the Real Estate	7
Occupier	Use the building	7
Internet of Things	Capture digital information about the real estate via sensors. These include information from the different building systems	6–7
Artificial Intelligence	Optimise the Real Estate performance based on energy, utilisation, well and health being and fault detection and diagnosis	6–7

- Traditional licensing: companies purchase data and information based on tailored contracts. This method is largely used by insurers that require current and factual data to assess the real estate risk.
- On-demand: data consumers purchase data by access, time or size as a subscription model. This approach is used by Artificial Intelligence and the Internet of Things to retrieve information from the Data Marketplace. Also, tenants and landlords can use this approach to get information from the Real Estate and its usage.
- Collaboration: data providers are data consumers and vice versa. The collaborative model enables different stakeholder of the project, including designers and contractors to validate each other information. The more validations, the more valuable the Data Marketplace.

Digital channels for invoicing the collected payments and revenue from the different marketplace users.

7.3 Data Marketplace governance

The Data Marketplace governance covers the way in which the market operates and the relationships between market users and both other users and the market itself. Governance responsibilities include:

- Ensuring the neutrality regarding participants. Neutrality must extend to permitting unfettered access by future participants
- Ensuring data are structured in accordance with standards that will be adopted
- Enforcing and legally auditing data validation and verification
- Legally auditing AI algorithms (black-box) and software developers
- Maintaining the Data Marketplace
- Supervising terms and conditions: licensing conditions for data categories and related licensing agreement.
- Monitoring service level agreements (SLAs)
- Managing licensing models: use of information between proprietary data, data suppliers services
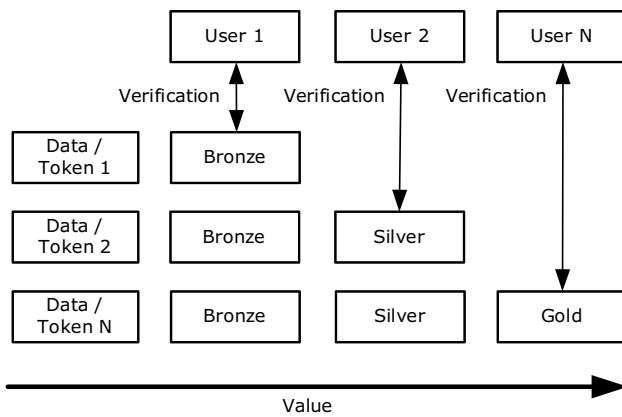

Fig. 5 Real Estate Data Market validation

Fig. 6 Real Estate Data Market verification

7.3.1 Validation

Validation services play a major part in the infrastructure as the degree of third party validation determines the value of any given data set at any particular moment in time. Tradability of data is enhanced by tokenisation of the actual data which are attached to the token as part of the payload to allow liquidity and transparency (Fig. 5).

7.3.2 Verification

Verification is provided by private companies such as Resilience Partners Ltd (RPL) who provide a 3-step verification architecture for their clients to determine required levels of security and professional consortia such as the Foundation for International Blockchain and Real Estate Expertise (FIBREE) and the Royal Institute of Chartered Surveyors (RICS). The level and extent of the verification applied to each data set determines the validity and hence the value of the data. Verification also includes the digital certification of the Data Marketplace users via Public Key Infrastructure or any other cryptographic method (Fig. 6).

7.3.3 Personal identifiable information

Although the marketplace is not designed to hold personally identifiable information (PII), consideration is required about the way in which sensitive information passes between the participants. Compliance with international standards such as ISO 27001 for data security and national legislation such as the General Data Protection Regulation 2018 (GDPR) is essential. Appropriate governance depends on the method used by participants to store their data, prior to transfer, using cloud-based storage services such as Azure Data Lake and Amazon S3, or a distributed file system such as Apache Hadoop.

One of the key challenges of a well-designed and effectively implemented Data Marketplace is how to secure information in the data lake, which requires consideration of not only the appropriate standards that apply such as Adopting Data Standards (ADS) and IoT cloud orchestration but also the implementation of network authentication protocols such as Kerberos and the data and encryption via symmetric keys, Advanced Encryption Standard, (AES-256) and public-key cryptography, Rivest–Shamir–Adleman (RSA-2048). Metadata and smart contracts are used to manage data access rights, capture data profiling results and commentary by data developers and end-users. The storage of policies that define production readiness and enforcement techniques can also be defined. It is therefore proposed that a regulatory sandbox is developed based on that operated by the FCA, which proves the underlying innovative technology in a live environment with real customers, which allows firms to better understand and fine-tune their business model [40].

7.4 Data Marketplace technology

The technology of the Real Estate Data Marketplace transforms the traditional data lake into a Data Marketplace where data are:

- Normalised following a data standard that defines the model, format and attributes of data.
- Aggregated into logical clusters or catalogues.
- Assured and verified according to consistency and quality checks for accountability.

The core technology architecture is based on Blockchain and Distributed Ledger Technologies that enables the verification, validation and purchase of data (Fig. 7). Every transaction is recorded and audited. In addition, data are protected against fraudsters and cybersecurity attacks. The core technology of the Data Marketplace also ensures that data services are purchased and invoiced according to terms and conditions and service licence agreements.
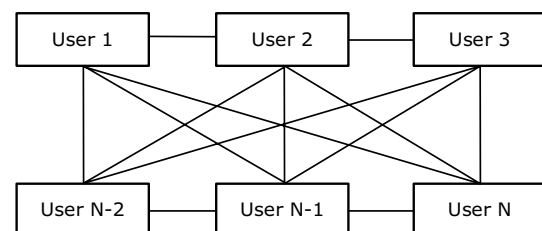

Fig. 7 Real Estate Marketplace Blockchain network

The Technology architecture is a decentralised, distributed and digital ledger in a peer-to-peer network. In detail:

- Token: real estate tradeable data. Tokens cover several information types such as real estate documentation and IoT real-time data. The characteristics of the token will depend on the nature of the task that the data are to be used for. The proposed Data Marketplace model proposes users to set up tokens if the marketplace does not meet their needs.
- Nodes: each user of the Real Estate Data Marketplace corresponds to a node of the peer-to-peer network.
- Blocks: batches of transactions are stored in tokens which are based mainly on two different data transactions: real estate documentation and real estate IoT real-time data.
- Decentralisation: information is stored by the different Real Estate Data Marketplace users. Users can store data in cloud-based storage servers
- Wallet: the use of cryptographic keys and hashes provide addresses to access the real estate information owned by the users
- Mining: the validation process is performed decentralised the relevant nodes that have the knowledge to validate the real estate data. IoT real-time data are validated automatically
- Transactions: the alteration of the data is recorded in transactions
- Proof of work: real estate nodes use different algorithms to verify and approve transactions. This also includes timestamps schemes to audit changes

7.5 Data Marketplace Artificial Intelligence

Artificial Intelligence in the Real Estate Data Marketplace is applied for:

- Management: the bulk of data validation and verification performed by a computer that automatically learns about the data in feature learning, anomaly detection and association rules.
- Commercial: machine learning can automatically cluster data to generate catalogues via several techniques such as support vector machines and decision trees. In addition, the search for data and metadata tools is also supporting by supervised and unsupervised learning.
- Analysis: recognition of trends for predictions, this can be performed by neural networks such as convolutional neural networks or long short-term memory that perform regression analysis about future values or fault detection and diagnosis.

7.5.1 Ethics

The implications of the use of Artificial Intelligence in the real estate sector are extensive and beyond the scope of this paper. Rights and obligations arise in real estate transactions which have wide-ranging effects on the legal rights of individuals, so care must be taken to ensure that any legal decision-making relating to property rights, payments and borrowing is not made automatically by a machine. Intelligent algorithms will increasingly require formal testing, possibly certification, regulation, and status in law. This is especially challenging for machine learning algorithms due to their prevalent 'black box' and evolving nature [41].

Recent concerns around public trust in the use of algorithms require testing and auditing around three basic concepts:

- *Lawfulness* Defined in terms of a requirement to 'identifying the purpose of a system and how this relates to law'.
- *Fairness* Defined in terms of the requirement to avoid 'discrimination and consideration of the impact of individuals' reasonable expectations.
- *Transparency* Contrastingly, while the importance of transparency is often asserted, it is not immediately defined.

The use of machine learning and Artificial Intelligence is in its infancy in the real estate sector, so ethical issues are not yet presenting to those who are planning pilot implementations. The key risk for AI learning algorithms is already introduced bias in the training data sets, therefore an ethical Data Marketplace must include theoretical best models created by computers (genetic algorithms, adversarial learning, reinforcement learning). The authors are aware however of plans to introduce systems to evaluate tenants ability to discharge lease obligations and data analysis of movement around a building which will give rise to a tension between privacy and the perceived 'greater good' of the community of users.

Research is required to consider ethical issues that may arise from future use cases. For example, one of the challenges in the post Grenfell Tower drive for building safety is the ability to provide current and accurate data to Fire Offices for use in the event of a disaster. That data could include information from CCTV and IoT which, when aggregated could demonstrate unlawful occupation of residents. This is critical data for the preparation and execution of an evacuation plan but potentially highly contentious material that the tenant may wish to protect.

7.6 Data Marketplace robots

Data Marketplace robots and drones play a key role in this proposed Data Marketplace. Robots are data producers and data consumers, which can automatically be involved in the validation and verification stages. Drones can survey the external areas, whereas robots can survey the internal areas, respectively, of the real estate to verify the design and construction.

7.7 Data Marketplace roadmap

The proposed Data Marketplace roadmap is divided into several stages:

- Stage 1: Develop a Real Estate Data Marketplace proof of concept or research project with a reduced number of real data and real users. The proof of concept chooses the different tokens, the Blockchain technology and the hosting options of cloud and premise.
- Stage 2: Develop a standardised data structure and relevant metadata. The data structure covers RIBA stages of documentation and IoT devices information.
- Stage 3: Develop the Data Marketplace governance model with a clear demarcation of responsibilities.
- Stage 4: Develop a Real Estate Data Marketplace in a real project.
- Stage 5: Develop the commercialisation of the Data Marketplace and platforms to collect revenue.

8 Case study

8.1 Construction disputes

A key area for the use of corporate ODR is the construction sector due to the tendency for litigation in construction disputes. Furthermore, the industry is evolving to the digital BIM which combined with disruptive digital technologies such as AI and especially Blockchain present an organic application for corporate ODR. With corporate ODR for construction, the consortia require signing up for ODR before the beginning of an infrastructure or building contract, via a fast method based on neutrality, transparency and fairness. Relevant parties could be presented a series of possible 'algorithmic' mediated solutions as an option to expensive and prolonged arbitration or litigation.

8.2 RPL and the DigitalDisruption@BREI consortium

RPL and the DigitalDisruption at the Bartlett Real Estate Institute (BREI) consortium are currently undertaking a case study with UCL Estates to model the regulatory information required for the statutory regulators on the £14 m construction PEARL project (Person–Environment–Activity Research Laboratory). This project is using a collaborative process to validate data for JCT stage gate 5 as it transitions to the Project Board. The pilot has been given access to stage gateway meetings which are held by the developer, attended by all members of the construction team and the facilities manager who will have responsibility to operate the building post hand-over. All design issues are discussed and approved, with full consideration being given to regulatory requirements around fire safety and environmental compliance. An important objective is for all parties to reach a consensus about the document set that is attached to each stage gateway and a record kept of approvals and discussions around changes to key documentation. Structuring data in this manner can facilitate obligation analysis and consistency in approach across all projects.

9 Conclusions

This review paper has proposed a vision and a model of a Data Marketplace for the real estate with a brief on the digital technology industry. The purpose of this article is also to stimulate engagement with the computer science community and to encourage debate amongst professionals, academics and politicians of the benefits of full automation and integration of real estate.

Next research stages will include the development of a Proof of Work Data Marketplace based on Blockchain and Ethereum Virtual Machine. ERC20 tokens will cover Smart Contracts based on total supply, balance, transfer, approve and allowance of data. The proof of work will be validated using data sets from research institutions who manage IoT sensor infrastructure to emulate an Industry 4.0 model.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

1. IOTA presents Data Marketplace. https://blog.iota.org/iota-data-marketplace-cb6be463ac7f. Accessed 07 Jan 2021
2. Creating a successful Internet of Things data marketplace. https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/creating-a-successful-internet-of-things-data-marketplace. Accessed 07 Jan 2021
3. Rifkin, J.: The Zero Marginal Cost Society. St. Martin's Publishing Group (2014)
4. Zipori, G.: The (data) science of the deal: how ai will transform commercial real estate. https://www.forbes.com/sites/forbesrealestatecouncil/2019/01/23/the-data-science-of-the-deal-how-ai-will-transform-commercial-real-estate/#2d6a48b84200 (2019). Accessed 07 Jan 2021
5. How Smart Buildings will benefit Commercial Property Insurers. https://www.cognizant.com/perspectives/how-smart-buildings-will-benefit-commercial-property-insurers (2017). Accessed 07 Jan 2021
6. Building information modelling. www.autodesk.com/solutions/bim. Accessed 07 Jan 2021
7. Treleaven, P., Barnett, J., Koshiyama, A.: Algorithms: the law and regulation. Computer **52**(2), 32–40 (2019)
8. Skies without limits, Drones taking the UK's economy to new heights. https://www.pwc.co.uk/intelligent-digital/drones/Drones-impact-on-the-UK-economy-FINAL.pdf. Accessed 07 Jan 2021
9. Collective construction with robot swarms. http://people.seas.harvard.edu/~jkwerfel/morpheng.pdf (2012). Accessed 07 Jan 2021
10. FIBERBOTS: design of a multi-agent, fiber composite digital fabrication system. https://www.media.mit.edu/projects/fiberbots/overview. Accessed 07 Jan 2021
11. The law and regulation of drones|AI lawyer. http://www.jeremybarnett.co.uk/data-protection-and-future-internet/the-law-and-regulation-of-drones-ai-lawyer. Accessed 07 Jan 2021
12. Baum, A.: Proptech 3.0 The Future of Real Estate Said Business School. University of Oxford (2017)
13. Baum, A.: Tokenisation: The Future of Real Estate. University of Oxford Research (2020)
14. ResearchMoz global construction industry guide. http://www.digitaljournal.com/pr/3293605 (2017). Accessed 07 Jan 2021
15. Building a Safer Future. Independent Review of Building Regulations and Fire Safety. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/707785/Building_a_Safer_Future_-_web.pdf (2018). Accessed 07 Jan 2021
16. Arcadis global construction disputes report. https://www.arcadis.com/en/united-states/our-perspectives/global-construction-disputes-report-2020/ (2020). Accessed 07 Jan 2021
17. The state of Construction Technology. https://www.us.jll.com/en/trends-and-insights/research/the-state-of-construction-technology (2020). Accessed 07 Jan 2021
18. IPSX. https://www.ipsx.com/about. Accessed 07 Jan 2021
19. Conference 'Tokenization a great opportunity for real estate & asset financing'. https://cms.law/en/nld/news-information/tokenized-assets-amsterdam-tokenization-great-opportunity-for-real-estate-asset-financing. Accessed 07 Jan 2021
20. Siegel, D.: The Token Handbook. Hacker Noon (2017)
21. ERC-20. https://eips.ethereum.org/EIPS/eip-20. Accessed 07 Jan 2021
22. Deloitte, Smart buildings, how IoT technology aims to add value for real estate companies. https://www2.deloitte.com/us/en/insights/focus/internet-of-things/iot-commercial-real-estate-intelligent-building-systems.html (2016). Accessed 07 Jan 2021
23. Serrano, W.: iBuilding: artificial intelligence in intelligent buildings. Adv. Comput. Intell. Syst. **1043**, 395–408 (2019)
24. Serrano, W.: The blockchain random neural network for cybersecure IoT and 5G infrastructure in smart cities. J. Netw. Comput. Appl. **175**, 102909 (2021)
25. Serrano, W.: Digital systems in smart city and infrastructure: digital as a service. Smart Cities **1**(1), 134–154 (2018)
26. Data for the public good. https://nic.org.uk/app/uploads/Data-for-the-Public-Good-NIC-Report.pdf. Accessed 07 Jan 2021
27. The Gemini principle. https://www.cdbb.cam.ac.uk/system/files/documents/TheGeminiPrinciples.pdf. Accessed 07 Jan 2021
28. Design computation. http://www.designcomputation.com/. Accessed 07 Jan 2021
29. Algorithmic design: an insight into T2.a Architects' process. https://www.designbuild-network.com/features/algorithmic-design. Accessed 07 Jan 2021
30. Dubai smart city platform. https://www.dubaipulse.gov.ae/iot. Accessed 07 Jan 2021
31. SimStock UCL Energy Institute. https://www.ucl.ac.uk/energy-models/models/simstock. Accessed 07 Jan 2021
32. London buildings stock model, the Bartlett faculty of the built environment, UCL. https://www.ucl.ac.uk/bartlett/energy/news/2018/may/greater-london-authority-commissions-ucl-develop-london-buildings-stock-model. Accessed 07 Jan 2021
33. Dew, P., Barnett, J.: e- Legal services: To speed up justice for B2B systems (2016). https://goughsq.co.uk/publication/e-legal-services-to-speed-up-justice-for-b2b-systems/
34. The code of ethics. https://www.nar.realtor/about-nar/governing-documents/the-code-of-ethics. Accessed 07 Jan 2021
35. Ethical tech: making ethics a priority in today's digital organization. https://www2.deloitte.com/content/dam/insights/us/articles/6289_ethical-tech/DI_DR26-Ethical-tech.pdf. Accessed 07 Jan 2021
36. Everyday ethics for AI. https://www.ibm.com/design/ai/ethics/everyday-ethics. Accessed 07 Jan 2021
37. Ethical use of customer data in a digital economy. https://assets.kpmg/content/dam/kpmg/uk/pdf/2019/04/ethical-use-of-customer-data.pdf. Accessed 07 Jan 2021
38. Serrano, W.: Neural networks in big data and web search. Data **4**(1), 7 (2019)
39. Serrano, W.: A big data intelligent search assistant based on the random neural network. Adv. Intell. Syst. Comput. **529**, 254–261 (2017)
40. FCA sandbox. https://www.fca.org.uk/firms/innovation/regulatory-sandbox. Accessed 07 Jan 2021
41. Barnett, J., Treleaven, P.: Algorithmic dispute resolution. Comput. J. **61**(3), 399–408 (2018)