Data as Economic Goods: Definitions, Properties, Challenges, Enabling Technologies for Future Data Markets

Demchenko, Y.; Los, W.; de Laat, C.

DOI
10.5281/zenodo.2483184

Publication date
2018

Document Version
Final published version

Published in
ITU Journal

License
CC BY-NC-ND

DATA AS ECONOMIC GOODS: DEFINITIONS, PROPERTIES, CHALLENGES, ENABLING TECHNOLOGIES FOR FUTURE DATA MARKETS

Yuri Demchenko, Wouter Los, Cees de Laat
System and Networking Lab, University of Amsterdam

Abstract *– The notion that data has value is commonly recognized. However, data value is different from that associated with consumable goods. There is a number of initiatives to create data markets and data exchange services. Existing business models of paid or commercial data (sets) services such as data archives are based on service subscription fees. However, emerging data-driven technologies and economies facilitate interest in making data a new economic value (data commoditization) and consequently identification of the new properties of data as economic goods. The following properties are leveraging FAIR data properties and defined as STREAM properties for industrial and commoditized data: sovereign, trusted, reusable, exchangeable, actionable, and measurable. Other properties to be considered and necessary for defining workable business and operational models are non-rival nature of data, data ownership, data quality, value, privacy, integrity, and provenance. The paper refers to other discussions and projects on defining data as consumable goods and market mechanisms that can be applied to data exchange, such as data markets, data exchange, and industrial data space.*

Keywords – Big data, big data infrastructure, data exchange, FAIR principles, data management, data markets, open data, STREAM data properties

1. INTRODUCTION

The emergence of a data-driven economy, powered by big data and cloud computing technologies [1, 2, 3], motivates research and exploitation of data market and data exchange models that can facilitate/enable effective data exchange as economic goods and application-specific integration while protecting personal data, data ownership and intellectual property rights (IPRs).

Companies, government bodies, academic institutions and citizens have access to, and are increasingly using more, data today than could ever be imagined a decade ago. Traditional data sources such as company databases and applications are now complemented by a variety of open data and social media data or sensors embedded in IoT devices, including mobile devices, smart meters, cars and industrial machines. Economy digitalization and a growing volume of data has created an entirely new market of big data technologies and services to help organizations capture and extract value from all the data. The revenue from big data technologies and services, however, is small compared to the value that is expected to result in sectors such as trade, manufacturing, finance and insurance, public administration, and health and social care that now have the tools at their disposal to make innovative uses of data to drive high-value business and societal outcomes [4]. Unleashing the full potential of data produced by the digital economy will require both the creation of infrastructure to facilitate data exchange and development of new market models and mechanisms.

The fact that data has value is commonly recognized. However, data value is different from that associated with consumable goods. There is a number of initiatives to create data markets and data exchange services. Existing business models of paid or commercial data (sets) services such as data archives are based rather on service subscription fees than measurable properties of data. The quality of data sets is in many cases assessed by an independent certification body or based on peer review by experts. Such models are useful for specific use cases, but they do not provide consistent models to make data an economic good and enable data commoditization (i.e. transforming data into an object of trade and exchange [5]). Interest to make best use of research data has resulted in the formulation of the FAIR (Findable – Accessible – Interoperable – Reusable) data principles [6] which are fully accepted by the international research community and widely supported by industry and business. However, emerging data-driven technologies and economies facilitate interests to making data a new economic value (data commoditization) and consequently identification of the new properties of data as economic goods. This paper leverages FAIR principles to define the proposed STREAM data properties for industrial and commoditized data: sovereign, trusted, reusable, exchangeable, actionable, and measurable. Other data properties important to enable workable/actionable business and operational models include data ownership, data quality, data value, privacy, integrity, and auditability.

This paper refers to past and ongoing discussions to define data as economic goods and market mechanisms that can be applied to data, such as data markets, data exchange, and industrial data space.

The remainder of the paper is organized as follows. Section 2 analyses the emerging data-driven economy and identifies data markets as important driving forces. Section 3 discusses gaps and challenges on the adoption of data markets. Section 4 introduces the proposed STREAM data properties and explain their importance in the context of data markets. Section 5 analyses data market concepts and defines the main functional components and mechanisms required to enable open data markets and data exchange. Finally, the conclusion section provides a summary and invites for future discussion.

2. DATA ECONOMY AND DATA MARKETS

The European data economy is growing at a fast pace and will continue to do so in the upcoming years. The overall value of the data economy grew from €247 billion in 2013 to almost reaching €300 billion in 2016. According to the estimates of the European Data Market study, the value of the data economy in 2016 was worth nearly 2% of the European GDP. By 2020, the data economy is expected to increase to €430 billion with an overall impact of 2.5% on the GDP in the baseline scenario, which is defined by a continuation of the positive but moderate growth trend of the economy [7].

A data-driven economy uses processes and products/services digitalization, it collects and processes large amounts of data that include both personal and non-personal data. For personal data, the recently enacted GDPR mandates multiple measures to protect personally identifiable data, in particular informed consent [8]; this should build further confidence and trust in digital technologies and build trusted environments for data exchange via the future data market [9]. In fact, research has shown that people tend to share more data when they trust the companies processing the data. The data economy tends to involve the maximum possible amount of data that may be produced by multiple sources, so creating a trusted environment for data exchange and trading is a key enabling factor for data markets and data-driven technologies. Data quality and veracity are important properties of data for data-driven applications.

2.1 Data markets as enabling mechanism for data economy

The establishment of the Open Data Market as a necessary stage in making a digital single market (DSM) in Europe a reality and working to facilitate the digital transformation of the European economy to embrace new trends related to Industry 4.0 [10] and data-driven technologies. Data markets are recognized as a priority topic in the Horizon 2020 Work Programme 2018-2020 [11]. Currently, data exchange is limited to open data (including social media data, city data, governmental data etc.) while customer activity data, personal data, business operational data and industrial data that has a strong potential of bringing global optimization to multiple human/social activities and resources use, still remain in proprietary/private custody by companies. This brings unfair benefits to big network(ed) and technology companies that all tend to create their own full set of data that they collect from their business and operational network. Data markets [11] and data exchange [12] have the potential to bring benefit for facilitating business activity, social creativity and technology advancement, but current infrastructure, data protection, legal conditions and regulations are not ready to offer the necessary environments. There are multiple initiatives and projects that have collected valuable experiences, however, most of them are of an ad-hoc nature and involved a limited number of stakeholders and they are not based on sufficient conceptual research, which in its own turn would require a significantly wider experience and observational basis.

2.2 Data-driven economy and technology integration

As described in recent studies and reports there are clear trends of all technologies' digitalization and an increased demand of advanced/smart data handling to enable the full potential of data-driven technologies. This is characterized by wide technology integration based on modern big data infrastructures and technologies which are predominantly cloud based.

Data processing and data analytics require continuous data storage and exchange between applications and processes, which need to be supported by a scalable distributed data storage infrastructure, such as using a Hadoop Distributed File System (HDFS) [13], Data Lake Storage [14], and facilitated by a special functional component such as data exchange, which can be a part of the data market [12] in a general business scenario.

Currently, there is no common approach and architectural model to address potential cross-industry, cross-application and cross-platform integration and exchange of data and data-driven applications. Most current developments and use of data in data analytics and data-driven applications require and remain inside vertical data processing stacks (typically belonging to one company). The best examples of effective data use for business purposes and offering data analytics platforms for business as services are provided by big data infrastructure and cloud providers such as Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform, IBM, Facebook. All these companies combine three key components of the modern hub/networked digital economy [15]: (1) (big data) computation and storage infrastructure together with business (data analytics) services, (2) direct business services and marketplace, (3) wide/extended customer base and/or access created around products or services. In recent research [16], these types of companies are defined as "superstars" that achieve their market dominance by using exclusiveness of information and data collection in a data-driven economy. Besides the regulation measures suggested in [16], necessary data market and data exchange mechanisms need to be created to allow data exchange at all stages of the data value chain from data production to data analysis and business application.

Few European projects address general issues of data protection, policies, and exchange. However, there is no common vision and widely accepted approach to create open data markets that would benefit big and small companies. European Union (EU) and Member States, anticipating the DSM paradigm, invested significant resources in building an advanced digital infrastructure and regulatory basis for a digital economy in Europe. A recent development of digital and data services is primarily based on individual companies' genius that discover and explore unique business opportunities. However, despite bright examples such as Facebook, Twitter, Instagram and other social network and social media platforms that actually demonstrated the potential of data-driven and digital technologies, the full potential of the data-driven economy has not been realized yet. Recent technology development is characterized by conceptual research lagging behind occasionally generated business opportunities.

Further development of data-driven technologies must combine conceptual research and development and implementation of economy-wide technological platforms for data collection, storage, exchange and integration with industry, business and social applications.

3. CHALLENGES IN DATA MARKETS ADOPTION

Based on the analysis provided above, we can define the general gap that creates a number of challenges for future data economy development and data market adoption: Data is becoming an economic good but there is no facility to unleash its full market potential.

We can list some of the identified challenges that motivated current research by the authors while actively developing the modern data infrastructure for data-driven research and industry:

- Data property as an economic good and commodity is not researched and not defined. Data is more than oil of the future economy.
- There is no common vision and model of how to trade data while retaining data ownership and sovereignty.
- The new data market model needs to be developed and adopted.
- GDPR provides common rules but there is no clear technology alignment to implement and enforce GDPR requirements. New ePrivacy legislation will make data management rules even stricter.
- There is no (or limited) coordination and interaction between industry and academia to develop new market mechanisms.

Technological aspects bring additional complexity to the topic of building data markets.

Use of modern cloud-computing and big data technologies and infrastructure is inevitable but there is no well-developed security and trust model for storing and processing sensitive/proprietary data on cloud.

4. DATA PROPERTIES AS ECONOMIC GOODS

This section provides suggestions about defining data properties as economic goods and explains their importance for enabling data markets and facilitating data exchange. The section refers to related works that created background in defining the proposed data properties.

4.1 From FAIR principles to STREAM properties

Initially proposed by the research community, the FAIR (Findable, Accessible, Interoperable, Reusable) data principles [6,19] found also strong support among industry and is currently included as a priority topic in the Horizon 2020 EOSC (European Open Science Cloud) Programme [20]. The FAIR principles are complied with and extend data governance and data management models at enterprises, as defined in industry standards: the DAMA Data Management Body of Knowledge (DMBOK) [21] and CMMI Data Management Maturity model [22].

Factors that facilitate the need for data trading as economic goods include but not limited to:

- IoT sensor networks and farms that continuously produce data that potentially may be used by different organizations and produce secondary data that may have added value;
- use of personal data for advanced market research and services development;
- earth exploration data collected over years (such as from mining or oil/gas companies) that can be also offered on the market;
- existing data archives whose value may increase if data is traded in a more flexible and measurable way;
- secondary data created from open data.

Emerging data-driven technologies and economies facilitate interest in making data a new economic value (data commoditization) and consequently the identification of new properties of data as economic goods.

The following properties are leveraging FAIR data principles and defined as STREAM properties for industrial and commoditized data:

[S] Sovereign
[T] Trusted
[R] Reusable
[E] Exchangeable
[A] Actionable
[M] Measurable

Other data properties that are important for enabling data commoditization and allowing data trading and exchange for goods include: quality, value, auditability/traceability, branding, authenticity, as well as the original FAI(R) properties findability, accessibility, interoperability. Special features that must be managed in all data transfer and transformation are data ownership and IPR. The data property originated from its digital form of existence defined as not-Rivalry, on one hand, makes data exchange (copying, distribution) easy, but on the other hand, creates problems when protecting proprietary, private or sensitive data or IPR.

Below we explain why these properties are important for effective data trading and exchange between data market participants along the whole data value creation flow/process.

a) *Data sovereignty*

Data sovereignty allows companies, data owners, to keep control over their data. It is important for business to enter the data market with their proprietary business data and be confident that their data is not compromised or used by third parties, without consent from the data owner (or data controller). Sovereignty is a key principle of the Industrial Data Space Architecture as defined by International Data Space Association [18].

Often the data sovereignty principle is opposed to general data storage and processing on clouds where data resides in the cloud provider's data centers and there is fear that companies may lose control over their data, or the cloud provider may have unauthorized access to data or their use for business purposes. However, modern cloud and infrastructure virtualization technologies, provider business models and compliance provide a sufficient number of controls to satisfy security and trust requirements by companies to operate their businesses and host data on clouds (see CSA Complete Cloud Security Governance, Risk, and Compliance (GRC) Stack, Cloud Security Alliance [23]).

b) *Trusted data*

Using data in decision making or in the processes control requires that data is trusted and verifiable. Trust in data is achieved by the whole process of data collection and by using verified models of the processes that data represent, which must be in general auditable. In most business cases, data trustworthiness is ensured by the reputation of the data provider but all aspects of data production and origin must be verifiable and auditable.

c) *Data reusability*

Data reusability should allow multiple uses of data, even if it is not for the original purposes that the data was created for. Normally, data represents events, entity or processes and are application agnostic. Data reusability can create multiple opportunities for data economy actors, including small and medium enterprises (SME) or individual researchers. Data reusability should be supported by well-defined metadata and well-documented data collection processes. Data reusability is part of the research data FAIR principles and is well supported by metadata management tools [19].

d) *Data exchangeability*

Data exchangeability ensures that data can be exchanged between a data producer and data consumer in general and be used for target applications or intended purposes. To much extent, this implies data format and platform and APIs compatibility, which is achieved by industry standardization. Ideally, data exchange should be possible between compatible processes or applications during the whole data processing flow or data lifecycle. In the context of economic data value, exchangeability of data can also mean the possibility of exchanging data for other economic goods or money. However, data pricing models are not addressed by the authors at this stage, although some references to related works are provided below.

e) *Data actionability*

When data is purchased by companies they should serve their business purposes and contain the necessary information to derive actionable decisions about operation or process optimization, in particular customer experience improvement or quality of services delivered. When data is used in industrial processes, the actionable data must be extracted and included in the industrial processes control. With increased use of artificial intelligence in industry, the spectrum and variety of data used in the industrial production value chain is increasing and may include process monitoring data, logistics data, market data and user feedback data.

f) *Data measurability*

Data measurability can be discussed at least in two aspects: as an important property for data valuation and exchange as economic goods, and as a part of data handling on the data infrastructure platforms. The former still requires additional research and effective data valuation models, yet to be created. The latter is concerned with the resources that are required or consumed by the data storage or data processing facility or applications. This area is well developed and supported by modern cloud based data infrastructure. Both public cloud platforms (such as AWS or Microsoft Azure) and Open Source platform (such as OpenStack or CloudStack) provide a rich opportunity to monitor cloud resources usage by all data handling processes with details up to processor cycles, storage transactions or network traffic volumes. Such information is essential for infrastructure resources planning and costs estimation.

4.2 Combining data and algorithms

With the complexity of modern data collected and generated by human activity and IoT, there is a need to store/archive data with the corresponding application programming interfaces (API) and containerized applications that can handle the stored data. Stored data must contain metadata and schema as well as the blueprint for use and deployment (ready for integration with the target applications). API management [22, 24] is an important part of data management in modern data-driven companies; this includes both accessing data from external providers/sources and providing own data to customers and application developers. "Algorithm economy" term was popular in 2016-2017 [25] as one of the development directions of the emerging data-driven economy but now API management is a necessary part of the modern enterprise data infrastructure.

4.3 Other related concepts and models

Defining data properties as economic goods is a new research area and requires a complex approach involving different research and technology domains. There are not many published researches on this topic. The authors' research is motivated by the need to consistently define requirements, architecture and services that must be provided by the big data infrastructure and such infrastructure components as data exchange and data markets.

Information and data value are admitted in many general publications and research related to the big data value chain [26], emerging data-driven hub economy [15] and macroeconomic model of big technology firms (referred to in the paper as "superstar") [16]. However, these publications do not discuss data as economic goods that can be exchanged or traded for money.

Few papers and blog articles explored information as an asset. The seminal paper "Measuring the value of information: An asset valuation approach" by Daniel Moody and Peter Walsh (1999) [27] formulated 7 Laws of Information, which we quote here to facilitate further discussion and research (refer for details to the original paper [27]):

- First Law: Information is (infinitely) shareable
- Second Law: The value of information increases with use
- Third Law: Information is perishable
- Fourth Law: The value of information increases with accuracy
- Fifth Law: The value of information increases when combined with other information
- Sixth Law: More is not necessarily better
- Seventh Law: Information is not depletable

Two blog articles of 2013 [28] and 2014 [29] attempted to re-define the Moody&Walsh laws to Data Science [28] or apply them to information emerging from IoT and sensors [29].

Although the presented paper does not discuss data pricing and cost models, we mention two other papers by Muschalle, et al (2012) [30] and Heckman, et al [31] that discuss pricing approaches for data markets. Their research can provide initial input for a developing a data pricing model and how it can be used in the data market operation.

5. DEFINING OPEN DATA MARKET

Data markets are an important component of the data economy that could unleash the full potential of data generated by the digital economy and human activity in general. Future data markets should incorporate an open data market (ODM) model that should allow multiple market participants to join by complying with established rules and using standard APIs for data and information exchange. This section briefly discusses the ODM model and its functional components as a part of the modern data architecture and future data (driven) economy.

The proposed ODM approach leverages the international data space (IDS) architecture and functional model [17, 18] that are strongly based on the concepts of data sovereignty and infrastructure federation to ensure architecture openness.

5.1 Modern data architecture

Modern data architecture is enabled by cloud and big data technologies and has the following characteristics:

- capable of handling big data V-properties: volume, velocity, variety, and addressing data variability, veracity, value [1, 3];
- cloud based, elastic;
- customer-centric;
- automated, smart;
- adaptable, powered by Agile and DevOps continuous improvement/deployment model [32];
- combining data and algorithms (as part of containerized applications and data);
- collaborative;
- governed;
- secure, trusted.

5.2 Characteristics of emerging data markets

Data markets should be able to support all major properties of big Data from/for all application domains, allow data exchange and integration at different stages while preserving data provenance and auditability. Below is non-exhaustive list of data market properties leveraging big data architecture and powered by cloud technologies:

- supporting heterogeneous data exchange at different processing stages;
- cloud powered/integrated;
- customer-centric;
- automated, smart;
- regional/sectoral specialized;
- collaborative;
- governed;
- secure, trusted;
- auditable;
- transparent;
- enabling data commoditization and operating with monetized data values.

The listed above properties should provide a necessary framework for definition of the open data markets architecture, at the same time they will require a better definition of data as economic tradeable goods.

5.3 Data market and data exchange functionality groups

Data markets and their associated data exchanges include technological, trust management, legal, commercial and operational aspects, as described below.

5.3.1 Technical infrastructure

Infrastructure is a necessary component of any service architecture. The following are essential components of the data market infrastructure:

- architecture and conceptual model of the data market space, including technological, organizational, legal and commercial aspects;
- shared/federated infrastructure to access and operate the data market;
- federated hybrid cloud based big data infrastructure to support data storage, processing and exchange in a secure and trusted way;
- DataHubs support for generic services for data suppliers such as caching, streaming, containerized delivery;
- support for on-demand connectivity and bandwidth provisioning between data handling services/hosts in the data lifecycle;
- gateway based and computational enforcement of market policies and rules.

5.3.2 Rules for trust

Rules for trust provide the basis for trusted relations between data market participants and reducing contractual and operational risk between participants. This includes the following rules, policies, services:

- policy framework and platform bound mechanisms to participate in and cooperate with parties in the data market;
- models for agreements between parties in the data market and end users, with engineering for scalable (software) contracts and supporting architecture;
- compliance assessment tools of the big data infrastructure to enable trusted interaction between market actors;

- infrastructure and transactions auditing for performance and disputes.

5.3.3 Matching of parties (brokerage)

Brokerage and third party services are typical components of traditional market exchange services, however it is not trivial to enable such services over data exchanges. The following are essential brokerage services to enable effective and trusted data exchange:

- directory and registry services to enable findability of data and services, including data providers, target data applications and typical data uses, conditions, etc.;
- data filtering and quality assessment services, data processing API, value added service/ product offerings;
- automated/assisted contracts negotiation and data “value/property” transfer/exchange, customer support.

5.3.4 Data catalogues to keep everything together

A data catalogue (or Directory) is an essential component of the intended data market and data exchange infrastructure that can be distributed and operated in centralized hierarchical and federated modes. Catalogue service should enable both hosting metadata information and historical/provenance information about data sets and data transformation. The following summarizes the data catalogue properties and services:

- cataloguing data sets
- cataloguing data operations
- metadata catalogue
- searching
- data curation
- data quality assessment and data categorization
- linking data properties and applications
- recommendations and relationships
- data sets evaluation
- data access policies and API
- usage metadata
- lineage/provenance
- integration and interoperability
- supporting data pricing models (current and future)
- data visualization (properties, usage, quality)
- security
- compliance.

5.3.5 Data exchange protocol

Data exchange protocols are important components of future data markets and data exchange infrastructure. Data exchange protocols can be provided as a part of the data exchange (as a data market functional component) and use a layered model benefiting from reliable and secure modern Internet protocols, while deploying specific data exchange protocols as upper layer protocols. Modern network and cloud virtualization technologies allow the building of program able virtualized networks that can be provided as part of virtual private clouds (VPC) and support data applications that produce and consume data and involve data exchange internally inside secure VPC and with external parties.

Recent developments in data descriptions and data management introduced a number of mechanisms that can be used for effective and consistent data exchange protocols and applications: persistent data identifiers (PID), data factories, metadata and data types registry, data annotation and data discovery mechanisms, all these are part of the Research Data Alliance (RDA) outputs and recommendations [33].

6. CONCLUSION

The presented paper is intended to provide a basis for further wider discussion on defining data properties as economic goods and building future open data market and data exchange infrastructure by involving concerned communities, in particular those that the authors are actively involved with. These are Research Data Alliance, Industrial Data Space, European Open Science Cloud, and National Institute of Standards and Technologies (NIST) Big Data Working Group. The proposed definition of the STREAM data properties is based on extensive study of related works and authors’ experience in building modern big data infrastructure and cloud based services.

ACKNOWLEDGEMENT

The research leading to these results has received funding from the Horizon2020 projects MATES (funded under grant 591889-EPP-1-2017-1-ES-EPPKA2-SSA-B) and EDISON (funded under grant n. 675419).

REFERENCES

[1] NIST SP 1500-1 NIST Big Data interoperability Framework (NBDIF): Volume 1: Definitions, Sept 2015 [online] http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1500-1.pdf

[2] NIST SP 500-292, Cloud Computing Reference Architecture, v1.0. [Online] http://collaborate.nist.gov/twiki-cloud-computing/pub/CloudComputing/ReferenceArchitectureTaxonomy/NIST_SP_500-292_-_090611.pdf

[3] Demchenko, Yuri, Peter Membrey, Cees de Laat, Defining Architecture Components of the Big Data Ecosystem. Proc. International Conference on Collaboration Technologies and Systems (CTS 2014), May 19-23, 2014, Minneapolis, USA.

[4] European Commission (2017): Building the European Data Economy, Factsheet. [online] https://ec.europa.eu/digital-single-market/en/policies/building-european-data-economy

[5] Commodification, Wikipedia [online} https://en.wikipedia.org/wiki/Commodification

[6] Barend Mons, et al, The FAIR Guiding Principles for scientific data management and stewardship [online] https://www.nature.com/articles/sdata201618

[7] European Data Market study measuring the size and trends of the EU data economy, EC-IDC, March 2017 [online] https://ec.europa.eu/digital-single-market/en/news/final-results-european-data-market-study-measuring-size-and-trends-eu-data-economy

[8] Custers, B.; Hof, S. van der; Schermer, B. (2014): Privacy Expectations of Social Media Users: The Role of Informed Consent in Privacy Policies. In: Policy & Internet, Vol. 6, No. 3, pp. 268–295

[9] Osimo, D.; Micheletti, G.; Cattaneo, G. (2017): Technical Barriers to Data Sharing in Europe, European Data Market (D3.10), SMART 2013/0063, https://view.publitas.com/open-evidence/d3-12-technicalbarriers_06-01-2017-1/page/2-3.

[10] The Fourth Industrial Revolution: what it means, how to respond. [online] https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond

[11] Supporting the emergence of data markets and the data economy. Horizon 2020 Work Programme 2018-2020. Call ICT-13-2018-2019 [online] http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/ict-13-2018-2019.html

[12] Data exchange as a first step towards data economy, PwC 2017 [online] https://www.pwc.de/en/digitale-transformation/data-exchange-as-a-first-step-towards-data-economy.pdf

[13] Hadoop Distributed File System (HDFS) [online] https://searchdatamanagement.techtarget.com/definition/Hadoop-Distributed-File-System-HDFS

[14] Chris Campbell, Data lakes in modern data architecture, June 2018 [online] https://www.blue-granite.com/data-lakes-in-a-modern-data-architecture-ebook

[15] Managing Our Hub Economy, by Marco Iansiti, Karim R. Lakhani, Harvard Business Review, September 2017 [online] https://hbr.org/2017/09/managing-our-hub-economy

[16] Anton Korinek, Ding Xuan Ng, The Macroeconomics of Superstars, November 2017 [online] https://www.imf.org/~/media/Files/Conferences/2017-stats-forum/session-3-korinek.ashx

[17] White Paper Industrial Data Space [online] https://www.fraunhofer.de/content/dam/zv/en/fields-of-research/industrial-data-space/whitepaper-industrial-data-space-eng.pdf

[18] IDS Reference Architecture Model: Industrial Data Space, Version 2.0 [online] https://www.internationaldataspaces.org/wp-content/uploads/2018/04/InternationalDataSpacesAssociation_ReferenzArchitecture2.0.pdf

[19] FAIR data [online] https://www.dtls.nl/fair-data/fair-data/

[20] European Open Science Cloud (EOSC), EC Open Science Portal [online] https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud

[21] DAMA-DMBOK: Data Management Body of Knowledge (2nd Edition), DAMA International, July 2017. ISBN: 9781634622479

[22] Data Management Maturity (DMM) Model, CMMI Institute, 2018 [online] https://cmminstitute.com/data-management-maturity

[23] A Complete Cloud Security Governance, Risk, and Compliance (GRC) Stack, Cloud Security Alliance [online] https://cloudsecurityalliance.org/research/grc-stack/

[24] Magic Quadrant for Full Life Cycle API Management, Published 30 April 2018, ID G00319327, Gartner [online] https://www.gartner.com/doc/reprints?id=1-4Y21K37&ct=180501&st=sb

[25] The Algorithm Economy Will Start a Huge Wave of Innovation, January 20, 2016 [online] https://www.gartner.com/smarterwithgartner/the-algorithm-economy-will-start-a-huge-wave-of-innovation/

[26] Curry E., The Big Data Value Chain: Definitions, Concepts, and Theoretical Approaches. In: Cavanillas J., Curry E., Wahlster W. (eds) New Horizons for a Data-Driven Economy. Springer, Cham, 2016

[27] Daniel Moody, Peter Walsh, Measuring the value of information: An asset valuation approach. Proc. European Conf on Information Systems, ECIS'99 [online] https://www.semanticscholar.org/paper/Measuring-the-Value-Of-Information-An-Asset-Moody-Walsh/677d018aa724aef71e2ba4a363f7ba1748ea5bfe

[28] Seven Laws of Information and its contextual reference to IoT, By Prashant, July 9, 2014 [online] https://startups.microsoft.com/en-us/blog/seven-laws-of-information-and-its-contextual-reference-to-iot/

[29] Data valuation: Seven laws of data science, By Dr. J, June 8, 2013 [online] https://datascientistinsights.com/2013/06/08/data-valuation-seven-laws-of-data-science/

[30] A. Muschalle, F.Stahl, A.loser, G.Vassen, Pricing Approaches for Data Markets, Proc. International workshop on business intelligence for the real-time enterprise, 2012

[31] J.Heckman, E.Peters, N.G.Kurup, E.Boehmer, M.Davaloo, A Pricing Model for Data Markets, iConference 2015 Proceedings

[32] Pat Patterson, DataOps: Applying DevOps to Data - Continuous Dataflows Built With a DataOps Platform, 24 May 2018 [online] https://dzone.com/articles/dataops-applying-devops-to-data-continuous-dataflo

[33] Research Data Alliance Recommendations and outputs [online] https://www.rd-alliance.org/recommendations-and-outputs/all-recommendations-and-outputs