Data Properties as Economic Goods

Yuri Demchenko,
System and Network, University of Amsterdam

Session “Driving Digital Transformation - The need for Governing Data Value”
ICT2018, 4-6 December 2018, Vienna

Outline

- RDA12 BoF on Data properties as economic goods
  - Part of IDW2018, 5-8 November 2018, Botswana
  - https://www.rd-alliance.org/bof-data-properties-economic-goods-rda-12th-plenary-meeting
- STREAM Data properties and related research
- Data Markets and Data Exchange

Agenda RDA12 BoF:
Data Properties as Economic Goods

1. BoF goals, Agenda bashing and round of introduction
2. Overview of existing activities, initiatives and available resources on defining economic value of data and future data markets (moderator - Yuri Demchenko, UvA)
    - 2.1. FAIR data principles adoption in industry and business (Luiz Bonino, GO FAIR)
    - 2.2. Data Factories and other concepts to enable Data Market (Jianhui Li, CNIC, CH)
    - 2.3. Data Conceptualizations: Resource, Object, Role, Abstraction (Gary Berg-Cross, Ontolog, US)
    - 2.4. Contribution and statements from BoF audience (open list)
3. Discussion of the future market demand and identified STREAM properties
4. Discussion of possible activities and next steps
    - WG/IG vs information document
    - Contributors and editors/chairs

Decision

- *To create Interest Group on Data Properties as Economic Goods*
- *Hold next next meeting during RDA13 Plenary in Philadelphia, US on 2-4 April 2019*

Gaps: Data is becoming an economic goods but no facility to unleash their full market potential

- Data use as Economical model
  - Data use and re-use
  - Data localization
  - Vendor lock-in (apps and data)
  - Legal uncertainty (non-personal data, cross-border, GDPR, provenance)
- Data property as economic goods is not researched and not defined
  - ***Data is more than oil of the future economy***
  - ***Data is reusable to large extend but more than water***
- There is no common vision and model how to trade data while retain data ownership (and sovereignty)
  - **GDPR** provided common rules but there is not clear technology alignment
  - New **ePrivacy legislation** will make data management rules even stricter
  - The new Data Market model needs to be developed and adopted
- Use of modern Cloud Computing and Big Data technologies and infrastructure is inevitable
  - There is not well developed security and trust model for storing and processing sensitive/proprietary data on cloud
- There is no (or limited) coordination between industry and academia/research to develop new market model and mechanisms

Research and Technology Overview

- On the way to define data properties as economic goods
  - FAIR principles and STREAM data properties
  - Information value research, economics of “superstars”
- Modern Data Market architecture and components
  - Cloud and Big Data technologies enabled
- Industrial Data Space (IDS) Architecture and activities
- Data related standards as basis for Data Exchange design
  - RDA outputs: PID, Data Factories, Data Type Registries, Repository certification, etc.
  - NIST, OASIS, BDVA, IIC in US
  - Industry best practices on Data Management and Governance

Data Properties as Economic Goods

STREAM data principles for industrial and commoditised data

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

- Other data properties: Important ***to commoditise*** data
  - **Quality, Value, Auditability/Trackability, Branding, Authenticity**, as well as original FAI(R) properties Findability, Accessibility, Interoperability.
  - **Data ownership and IPR**: Special features that must be managed in all data transfer and tracked along all data transformation.
  - **Not-Rivalry:** data is not depleted because of sharing and exchange

- Leverages FAIR principles for research data
  - Findable – Accessible – Interoperable - Reusable

Data Markets and IoT

- IoT is considered as a key use case and a facilitator for Data Markets
  - Potentially many consumers for centrally or locally operated IoT infrastructure
- IoT (sensors) infrastructure often created by community/federal project and will produce open or community value added services
  - IoT data can be exchanged and traded
  - **Sensor networks are core of many Research Infrastructures**
- IoT is supported and powered by Edge/Fog computing infrastructure
  - Special cloud services by major cloud providers
- Numerous EU and industry studies and roadmap for IoT Data

- Data Market initiatives around Europe
  - Amsterdam Data Exchange – calling for pilot projects, Joint meeting with EC and JRC
  - Data Market Austria @ ICT2018 4-6 Dec 2018, Vienna
  - Data Exchange pilot by Copenhagen city, Barcelona city, etc.

Need for Usable Data Pricing Model

- Data production and maintenance cost money
  - In particular, sensor and IoT network
- Data Exchange infrastructure cost money
- Buying data from professional data producers (or markets/brokers) will save money and allow focusing on application aspects
- Need for smart data contracts and properties embedding
  - Ensuring policy enforcement

Data pricing model - Origin and related papers

- Daniel Moody, Peter Walsh, Measuring The Value Of Information: An Asset Valuation Approach, 1999
  - 7 Laws of information
  - Few followers to map to data properties
- J.Heckman, E.Peters, N.G.Kurup, E.Boehmer, M.Davaloo, A Pricing Model for Data Markets, iConference 2015 Proceedings
- A. Muschalle, F.Stahl, A.Loser, G.Vassen, PricingvApproaches for Data Markets, Proc. International workshop on business intelligence for the real-time enterprise, 2012

7 Laws of Information by Daniel Moody and Peter Walsh (1999)

"Measuring the value of information: An asset valuation approach" by Daniel Moody and Peter Walsh (1999)

- 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

Data Pricing Factors (J.Heckman, 2015) - Attributes selection

- Value based parameters (value of data to the consumer)
  - The value of data in terms of saving time, efforts or money
  - ROI for customer
  - Risk exposure
  - Data exclusivity
  - Level of ownership (ownership transfer)
- Qualitative parameters (attributes or meta-attributes of the datasets)
  - Age of data
  - Credibility
  - Accuracy of data elements
  - Quality
  - Format and structure
- Fixed and marginal costs parameters (directly measurable costs)
  - Data collection, storage, maintenance
  - Delivery cadence

Philosophy research on Information and Data

- Pyramid: Data – Information – Knowledge
  - Many research and publications but time to revisit
- The philosophy of information by Luciano Floridi. Oxford University Press (2011)
  - 18 Open Problems
- Floridi's "Open Problems in Philosophy of Information", Ten Years Later, by Gordana Dodig Crnkovic (Sweden), Wolfgang Hofkirchner (Austria)
- An application of the dynamic knowledge creation model in big data, By Jestine Philip (2018)
  - SECI process including social/community phase/cycle
  - Leverage 'Ba' concept of time-space-nexus" from Japanese philosophy to include the shared context for knowledge creation

Data Markets and Economics of “superstars”

- Amazon, Microsoft, Google, Facebook, Apple, IBM, Baidu/Alibaba?
- Big technology companies use exclusiveness of information and data to get (sometimes unfair) market leadership and dominance
  - Discussed at the last IMF 2017 conference (paper by A.Korinek)
  - Possibly regulation is needed
- Data value increases with more (different) data collected

- 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
- 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

Data Market Architecture components

Data Exchange – Data Connectors – Catalog - Brokers - Trust

- Architecture and conceptual model of the Data Market space, including **technological, organisational, legal and commercial aspects**
- **Data Exchange(s)** as the main component for Data Market actors interaction and data exchange
- **Data Connectors** to enable sovereign end-to-end data provider and consumer connection
- **DataHubs** to support for generic services for data suppliers such as caching, streaming, containerised delivery
- **Federated Access Control and Trust Management** 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
- Support for on-demand connectivity and bandwidth provisioning between data handling services/hosts in the data lifecycle
- Gateway based and computationally enforcement of market policies and rules

Data Exchange protocol - Components

- PID, OID, Data Registries, Data Factories – *All RDA output*
- Metadata and data types registry, data annotation and data discovery mechanisms
- Directory and registry (to enable publish - subscribe)
- API and schema
  - Combining data and algorithm
- Provenance, Auditing
- Blockchain for data exchange tracking and validation
- Trust Management and Trust establishment Protocol
- Policy and rules construction (+ enforcement)

- Build on top of reliable and proven by practice Internet protocols
  - ***Data exchange protocols defined as upper layer protocols***
- Leveraging Virtual Networks and Virtual Privat Clouds
- Leveraging IDS architecture

Data Tao

- From records to logs and monitoring to raw data as economic goods
  - Meaning of data as representation of the process
  - Value of data as guide for action – Actionable data
- Data – Information – Knowledge – Action – Production – Development (social) – (Life?)
  - How deep link to data should be maintained?
  - Data well – The deeper data we can maintain the greater transformation we can do
- Data value transformation
  - Data transformed are the same or new data?
- Data commoditization
- Data monetization
  - Requires quality measure and price model

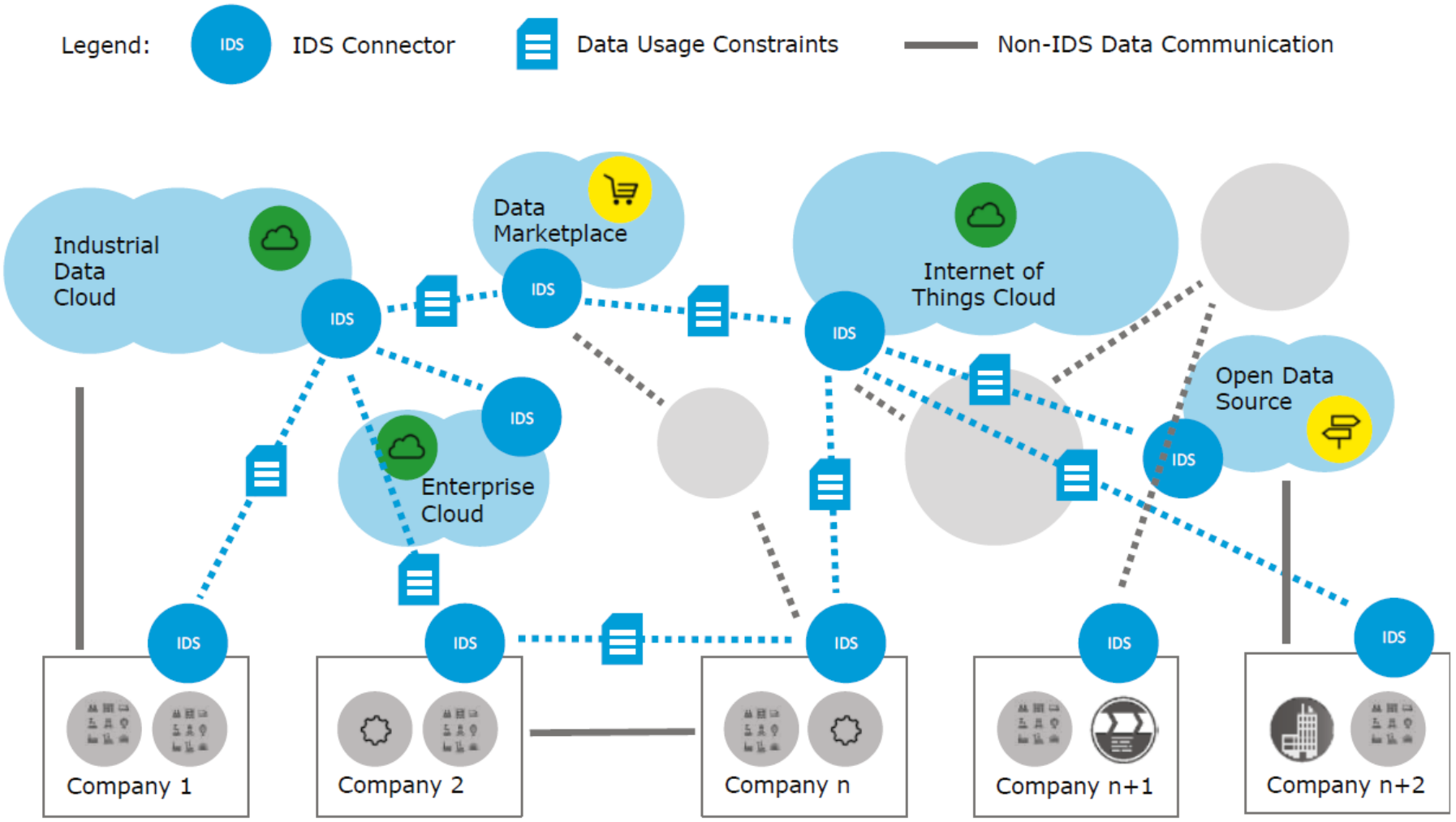
References

- Data as Economic Goods: Definitions, Properties, Challenges, Enabling Technologies for Future Data Markets, by Yuri Demchenko, Wouter Los, Cees de Laat. Position paper. To appear in ITU Journal: ICT Discoveries, Special Issue “Data for Goods”, 23 November 2018 https://www.itu.int/en/journal/002/Documents/ITU2018-12.pdf https://www.itu.int/en/journal/002/Pages/default.aspx
- Research Data Alliance (RDA) BoF on Data as Economic Goods https://www.rd-alliance.org/bof-data-properties-economic-goods-rda-12th-plenary-meeting
- RDA12 Poster "Bringing Data to Market: Data Properties as economic goods" http://www.uazone.org/demch/posters/rda12poster29-Data-economic-goods-markets-v01.pdf
- FAIR Data Initiative - https://www.dtls.nl/fair-data/
- Industrial Data Space Reference Architecture, Version 2.0 https://www.internationaldataspaces.org/wp-content/uploads/2018/04/InternationalDataSpacesAssociation_ReferenzArchitecture2.0.pdf
- DAMA-DMBOK: Data Management Body of Knowledge (2nd Edition), DAMA International, July 2017.

Technical Addendum

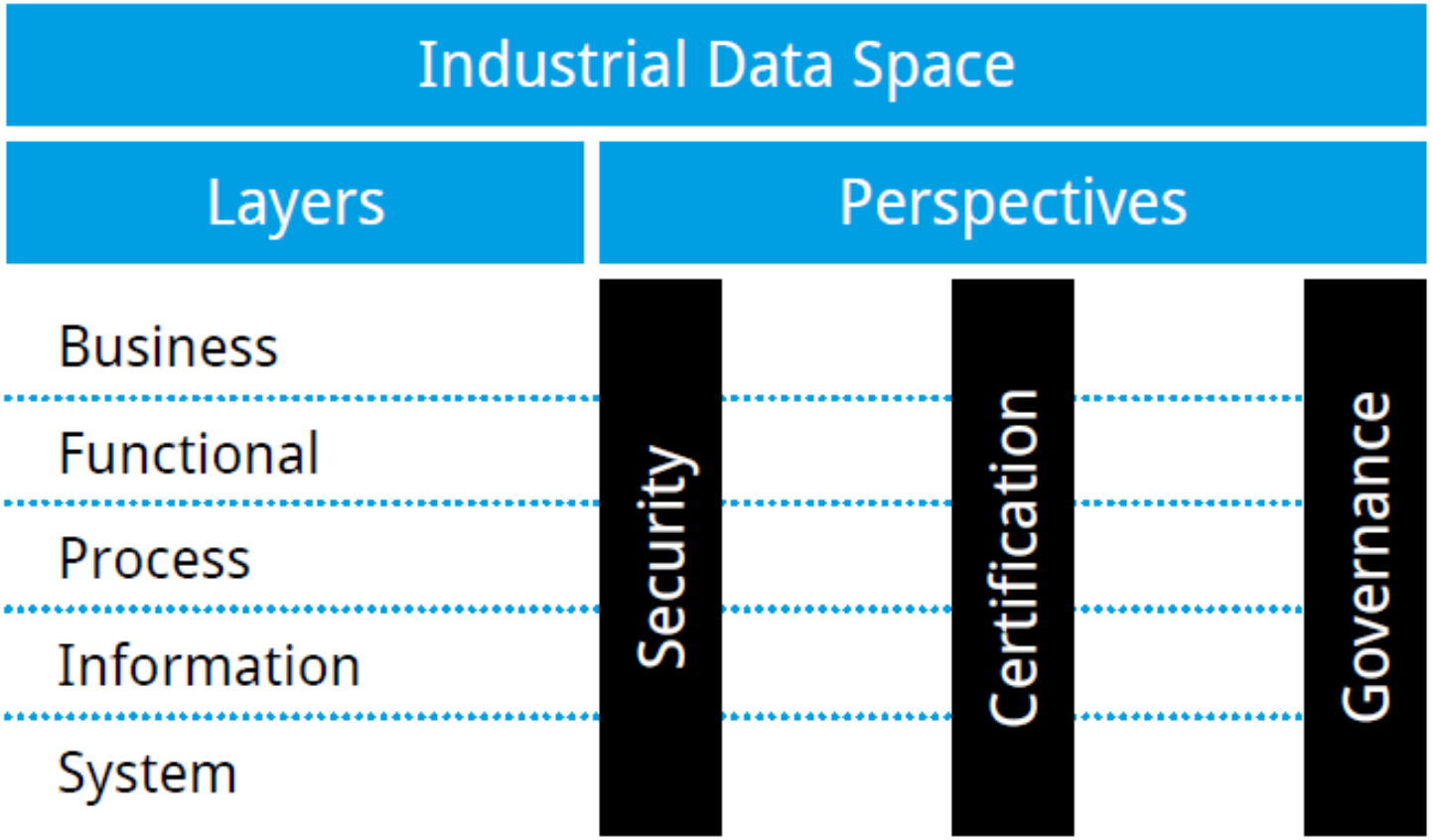
- International Data Space Architecture overview
- Existing standards
- Data types revisited
- Big Data Infrastructure and Data Markets Tchnology components

IDS and Cloud



- IDS Connector is the main functional component
- No specifically defined infrastructure

General Structure of IDS Architecture



- Specification defines functionalities by layers
- Details are sufficient to define processes, functional components and API