Interest Group on Data Economics (DE) Charter

IG Chairs
- *Yuri Demchenko*, Faculty of Science Informatics Institute, University of Amsterdam. Expertise: Computer science, metadata, research in general data management, and commerce
  Expertise: Big Data, Data Science, Data Analytics, Data Management and Governance, Computer Science, Educational technologies

- *Jane Greenberg,* Ph.D., Alice B. Kroeger Professor and Director of the Metadata Research Center, College of Computing and Informatics, Drexel University.
  Expertise: Metadata (automatic methods), Big Metadata, Linked data/semantics, Knowledge Organization, Information Economics, and Research Data Management.

Name of Proposed Interest Group: Data Economics

Introduction (A brief articulation of what issues the IG will address, how this IG is aligned with the RDA mission, and how this IG would be a value-added contribution to the RDA community):

Data is commonly recognized as a resource containing value associated with informetric qualities and measuring scholarly impact [1, 2]. Increased production of data from sensor networks for environmental monitoring, Digital Twins, and IoT devices have introduced another value driven data stream. Data from these developments are increasingly being integrated into data markets and data exchange services that support research and industrial processes [3, 4]. Although these services are generally commercial and require subscription fees or one-time payments for data access and use, they are part of the research data management environment. Furthermore, their design, regardless of a fee, supports **FAIR (Findable – Accessible – Interoperable – Reusable)** data principles and falling within the scope of RDA's mission. This topic warrants further discussion and community understanding on how data value is measured, including the economic value (commoditization) of data. Key issues that align with and add value to the RDA and motivate this proposal, include:

- **Informing RDA community about how data value is measured**, including exploration of data as an economic good.
- **Developing and endorsing a set of shared metrics** for measuring the value of data across different environments.
- **Linking with other RDA IG and WG** where data commoditization has important connections and impacts. For example, the Legal Interoperability IG, Big Data IG, and FAIR and metadata related IGs.

For these reasons, and to further enrich the RDA community, we propose the Data Economics IG. The IG will build on existing RDA and wider community experience in defining models and related properties (and attributes) to enable data exchange/trading as economic goods (including paid data services) and adoption of the different data economics models. It is important to recognize that paid data services are part of the research ecosystem, however data economics is not limited to financial matters, as value and impact can be measured in different ways. The IG will leverage and extend the **STREAM (S**overeign, **T**rusted, **R**eusable, **E**xchangeable, **A**ctionable, and **M**easurable**) properties [5]**, defined for industrial data as an extension to the FAIR principles. All of these aspects are important factors in forming this Data Economics IG Case Statement.

User scenario(s) or use case(s) **the IG wishes to address** *(what triggered the desire for this IG in the first place):*

Real world scenario triggers: Across the global research sphere, the data commoditization is gaining attention. For example, the European Union's initiative, "Building a European Data Economy" [8] recognises Data Markets as one of the priority topics in the European Horizon2020 Framework programme for 2019-2020 [9] to enable data exchange between industrial parties and increase data effectiveness. Looking specifically at cost and value, there has been research specifically focusing on Metadata Capital [e.g. 10], and, recently the U.S. Academy of Sciences charged a task force with, "Forecasting Costs for Preserving, Archiving, and Promoting Access to Biomedical Data" [11]. Additionally, a number of examples from industry demonstrate practical interest to facilitate and obtain value from data exchange, interoperability, and adopting common architectures for data markets. These includes but not limited to the proposed International Data Space Association (IDSA) Architecture [12]that defined the Data Market and related data exchange mechanisms for industrial data, or recent Open Data Initiative (ODI) by Microsoft and associates SAP and Abode, and others [13]. The examples are all concerned with research data, data valuation, data commoditization, and aspects of data markets, and highlight the need for an international forum and discussion.

RDA community interest: The proposed Data Economics IG is a direct and well supported outcome of the BoF, "Data as Economic Goods" that took place at the RDA 12 Plenary on 5-8 Nov 2018 in Gaborone. The BoF drew 29 participants from 13 countries; there were four presentations followed by an engaged discussion and support for forming an IG. The BoF agenda and notes are available are made available via the RDA portal [6]. Following the initial BoF gathering, a Case Statement was prepared, and a second meeting took place at the RDA Plenary 13 (P13) in Philadelphia, U.S.A., April 2-4, 2019. This second meeting drew nearly 60 attendees (two times the number of participants compared to the first meeting) from 18 countries. The result was clear interest in establishing the Data Economics IG, and meeting notes from this second meeting are also accessible via the RDA portal [7].

User Scenarios/RDA fit for the Data Economics IG: RDA, as a global organization, is an ideal community to advance discussion on data economics. The Data Economics IG will address the following *user scenarios/needs*: 1) the need to understand data costs and associated factors, 2) the need create consistent and workable models for measuring data valuation, and 3) the need understand the costs and commoditization related to data markets.

Consistent data pricing/monetisation and data markets models are equally important for government funded and sponsored research, open data and governmental data. The proposed IG will provide a needed, open forum for addressing these scenarios and build important connections with other RDA IGs and WGs.

Objectives (A specific set of focus areas for discussion, including use cases that pointed to the need for the IG in the first place. Articulate how this group is different from other current activities inside or outside of RDA.):

The *overriding objective* of this IG is the exchange of information about existing developments and initiatives and promotion of best practices in data valuation and data markets and operations. A set of specific objectives are to:

1. Provide and open forum and venue for gathering individuals interested in discussing and defining set of properties that are related to data costs as data as economic goods.
2. Compile initial agreed set of data properties and validate currently denoted STREAM properties.
3. Identify key actions to further research topics and coordination actions.
4. Identify set of documents to be produced by IG, in particular STREAM-like properties definition and overview of business models for research data and commoditized data trading and exchange.
5. Provide input to other RDA WGs and IGs, as well as to other associations and standardization bodies such as BDVA, IDSA, NIST, ISO and IEEE, using existing RDA channels or establishing new channels.

Participation (Address which communities will be involved, what skills or knowledge should they have, and how will you engage these communities. Also address how this group proposes to coordinate its activity with relevant related groups.):

IG members, as a community, will discuss and prepare a set of documents meeting the above objectives, and furthering exploration of the topic of data economics. Communication will be supported by RDA infrastructure, such as the community portal, wiki, and mailing list.

Interaction with other IG and WG is expected, external experts and contributors will be invited. Interaction with external organizations is expected: NIST, IEEE, International Data Space Association.

Outcomes *(Discuss what the IG intends to accomplish. Include examples of WG topics or supporting IG-level outputs that might lead to WGs later on.):*

At the initial stage of 1-1.5 year, the IG plans to produce a set of documents that prove an overview of existing approaches and concept related to data economic values, data pricing, economical models related to several use cases identified by RDA community on using and managing research data.

Other general outcomes may include:

- A framework to increase use and sharing data in secure and economical way; in particular between research community and industry
- A metadata scheme defining measurable data properties and data exchange models, which is useful for both the research community and industry.
- A shared understanding of economic value as applied to data, that supports broad research collaboration with industry, specifically data.

The IG activity will be driven by community and other expected outcomes will address community needs and requests. In particular, the IG will make the case for creating a taxonomy of the data properties as economic good that be a spin off Working Group.

Mechanism *(Describe how often your group will meet and how will you maintain momentum between Plenaries.):*

The IG will meet at the RDA Plenaries, will manage discussion on the list, and establish regular call (expected monthly) between plenaries. Activity between plenaries will be facilitated by working of the IG outcomes and documents.

Timeline *(Describe draft milestones and goals for the first 12 months):*

The suggested timeline is presented in the table below

Time span or milestone	Activity	Goals and outcomes
2019-2020	• Meetings at RDA Plenaries (first expected at RDA13 Spring 2019) – twice a year	Build community

Interest Group on Data Economics (DE) Charter

	• Regular tel/skype/video call – as required by work items (expected 3-4 between plenaries)	
December 2019	Initial draft information document on data properties as economic goods	Initial Draft, Discussion at RDA14
RDA15 Spring 2020	Interacting with other RDA IG/WG and external standardisation bodies and organisations	IG network is built
Summer 2020	Final draft on Data properties as economic goods. Call for IG and community comments.	Final draft published
Fall 2020	IG informational document finalisation	Document published by RDA
2021	Review IG results outcome and identify needs for renewed Charter and mandate	Renewed IG Charter

REFERENCES

 [1] Ding, Y., Rousseau, R., & Wolfram, D. (2016). Measuring scholarly impact. Springer International Pu.

[2] Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. *Journal of the American Society for Information Science and Technology*, 58(13), 2105-2125.

[3] Data marketplace powered by Blockchain and the global network of sensors [online] https://www.datapace.io/

[4] Data Intelligence Hub Data treasure instead of digital graveyards [online] https://www.t-systems.com/blob/811254/1accc153c557294e846ad0416a8430a6/DL_Best-Practice_BP05_Setlog.pdf

[5] Demchenko, Y., Los, W., de Laat, C. (2018). Data as Economic Goods: Definitions, Properties, Challenges, Enabling Technologies for Future Data Markets, ITU Journal: ICT Discoveries, Special Issue 2 "Data for Goods": https://www.itu.int/en/journal/002/Documents/ITU2018-12.pdf

[6] RDA 12 Plenary on 5-8 Nov 2018 in Gaborone. The BoF, Data as Economic Goods: https://www.rd-alliance.org/bof-data-properties-economic

[7] Philadelphia, U.S.A., April 2-4, 2019. Data Economics IG: https://docs.google.com/document/d/1HCnqfGKWu6NCbCtkIIRzjGyj4u59bCOdAkfFSPY-ZjE/edit#heading=h.euj6tazbbttq

[8] "Building a European Data Economy" https://ec.europa.eu/digital-single-market/en/policies/building-european-data-economy

[9] Towards a Common European Data Space', European Commission, April 25th 2018, https://ec.europa.eu/digital-single-market/en/news/communication-towards-common-european-data-space

[10] Greenberg, J. (2017). Big Metadata, Smart Metadata, and Metadata Capital: Toward Greater Synergy Between Data Science and Metadata, Journal of Data and Information Science, 2(3), 19-36. doi: https://doi.org/10.1515/jdis-2017-0012

[11] U.S. Academy of Sciences charged a task force with, "Forecasting Costs for Preserving, Archiving, and Promoting Access to Biomedical Data: https://www8.nationalacademies.org/pa/projectview.aspx?key=51436

[12] International Data Space Association (IDSA): https://www.internationaldataspaces.org/wp-content/uploads/2018/04/InternationalDataSpacesAssociation_ReferenzArchitecture2.0.pdf

[13] Fueling the Next Generation of Customer Obsession [online] https://theblog.adobe.com/open-data-initiative-summit2019/