This panel aims to share structures and issues of the Eco-System.
Panel Membership

Panelists
Anna Walek, Gdańsk University of Technology, Poland
anna.walek@pg.edu.pl

Petre Dini, IARIA, USA/EU

Yoshihisa Udagawa, Tokyo University of Information Sciences (TUIS), Japan
yu207233@rsch.tuis.ac.jp

Chair
Outlines

1. Back Ground
2. Definition of Data Ecosystem
3. Structure of Data Ecosystem
4. Involved Information Technology
5. Open Issues and Challenges
1. Back Ground

• Now, **data systems are taking a significant role** in companies and governments.

• **Needs for continuous improvement** of a data system regarding size and variety of data, and functions of a system over the years **are showing up**.

• The term “**data ecosystem**” likens growth of data system to a natural ecosystem.

• Issues and challenges can be rephrased as “Can the concept of **data ecosystem** be brought up to the level of giving social impact, and developing a new technical framework?”
2. Definition of Data Ecosystem

A data ecosystem is a collection of infrastructure, analytics, and applications used to capture and analyze data, for making better pricing, operations, and marketing decisions.

https://mixpanel.com/topics/what-is-a-data-ecosystem/
https://www.cognizant.com/glossary/data-ecosystem

A data ecosystem refers to a combination of enterprise infrastructure and applications that is utilized to aggregate and analyze information for better understanding of customers and for superior marketing, pricing.

https://www.cognizant.com/glossary/data-ecosystem

Eco-Software System can be defined analogously.

- The definitions are not yet finalized. They can be revised as needed.
3. Structure of Data Ecosystem (an example)

(1) Data acquisition and access function

① Files (CSV, Txt, etc.)
② APIs for structured/unstructured data on Web pages
③ APIs for Clouds
④ SQL for RDBMS
⑤ Data cleaning
Etc.

Data management
(reliability, security, integrated access, etc.)

(2) Data analytics function

① Statistical methods
② Data/pattern mining
③ Correlating findings
④ Generate insights/forecasts
Etc.

(3) Presentation function

① Result visualization
② Actionable plan/report generation
Etc.
4. Involved Information Technology (examples)

- Infrastructure (e.g., hardware, internet, clouds)
- File-based data access
- API-based data access (e.g., Web, clouds)
- Data-management tools (e.g., RDBMS, non-SQL)
- Statistical analyses tools/software
- Domain specific analyses
- Data mining algorithms (e.g., Apriori, Clustering)
- Machine leaning and AI
- Visualization tools/software
- Etc.
5. Open Issues and Challenges (examples)

- How to make a win-win condition for all the participants in a data ecosystem
  - Small instances of implementation experiences
  - Lack of consensus of data ecosystem structure, concept of operation and system growth

- How to get a data ecosystem to drive business innovations
  - Lack of methodology to collect high quality, related data in distributed environment
  - Immature technology for handling unstructured data
  - Information security and copyright
5. Open Issues and Challenges (continued)

- **Impact on development process**
  - Is the traditional waterfall model applicable? If not, what comes next?
  - For example, how to meet the request to upgrade 6 times every 3 months for the next 2 years?

- **Education of development personnel**
  - Technologies are not fixed.
  - Interoperability standards are not developed.

- **Responsibility concerning distributed data and open data**
  - Accounting for data provision
  - Responsibility for dissemination of fake information
  - Partial service outage due to system trouble

- **Cost estimation of administration and evolution of a data ecosystem for stakeholders**

Etc.
These are a brief introduction to the panel.

Thank you for taking time to watch.