Data policy of research data for commercial uses

#5 Open Science Data Promotion Workshop

2nd March, 2018

Japan Agency for Marine-Earth Science and Technology (JAMSTEC)
Center for Earth Information Science and Technology (CEIST)
Hajime Nishimura
https://orcid.org/0000-0001-8639-1927

Past presentations

My presentations are based on Environment issues should be solved inter-disciplinary with commercial activities.

#1: Action Items for Open Science from the view point of Inter/Trans-disciplinary Collaboration on Environmental Issues

#2: Cross-disciplinary collaboration platform - ubiDIAS

#3: Cross-disciplinary collaboration platform using MMORPG technology – Art and modern history.

#4: Proposal of overall strategies for Promoting Open Science.

#5: Today!: Data Policy for commercial uses.

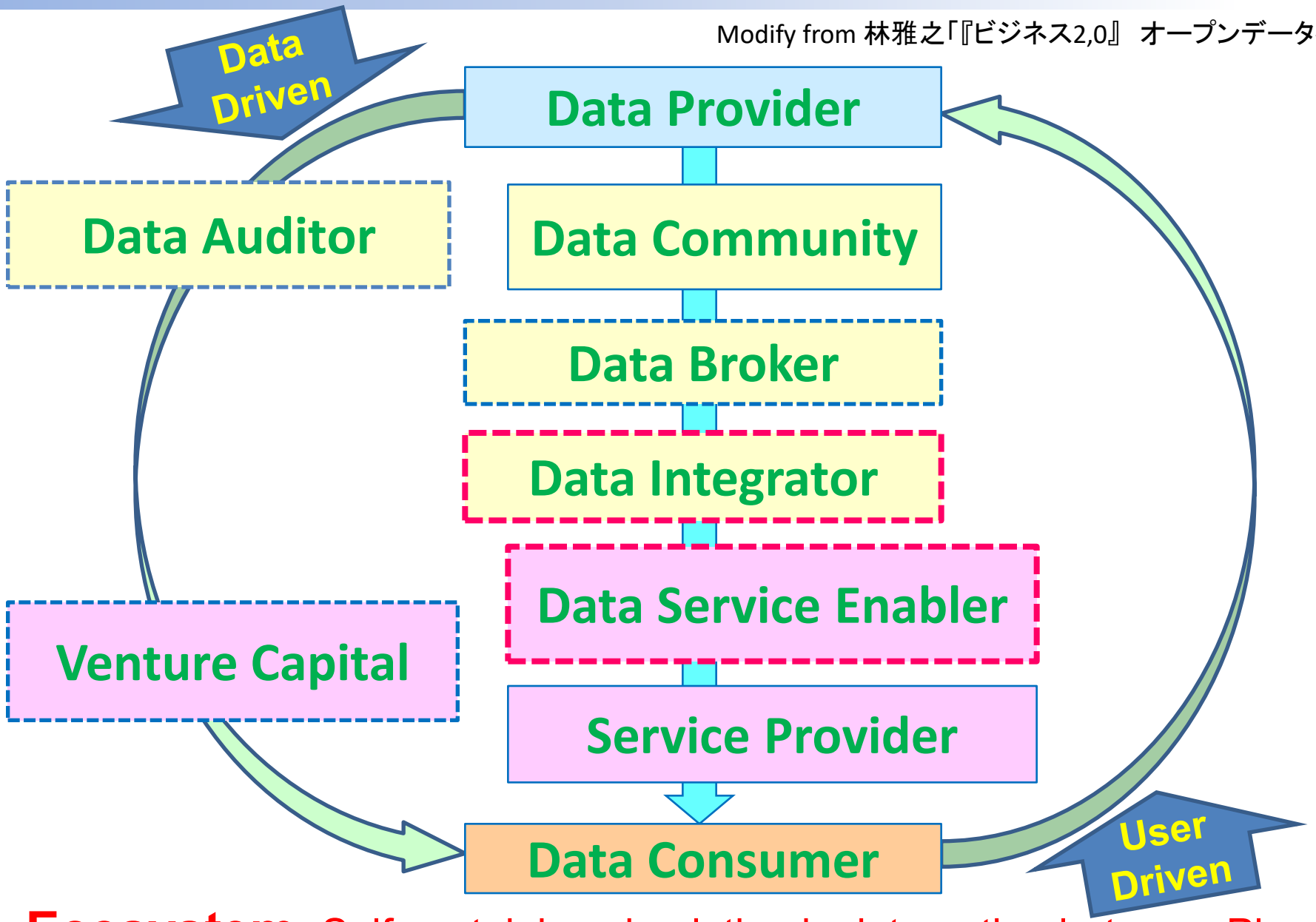
(#6: Cross-disciplinary data marketplace and analysis platform)

Today's Outline

1. Ecosystem, players, platforms
2. Why are almost use research data limited for research purposes?
3. Limitation of **Creative Commons Licenses**
4. Proposal of **Data Policy Guideline** for Commercial Uses.
5. Collaboration with **Data Trading Alliance**

1. Data Ecosystem and Players (Future)

Modify from 林雅之「『ビジネス2,0』 オープンデータ社会」



Ecosystem: Self-sustaining circulation by interaction between Players

Players and Platforms in Data Ecosystem

- Various players gradually grown up in accordance with growth of open data utilization.
- Data Driven: Two big data providers of meteorology data and geospatial data developed big data market. But from business view point, almost public research data are not provided for business use.
- User Driven: Users know not solutions but needs. So there are so many small Service Providers who close to each user.
- Above two driving forces are not enough for self-sustaining circulation of data utilization. Common platforms are need to improve efficiency of data utilization.
- How to initiate new inter/ trans-disciplinary data market?

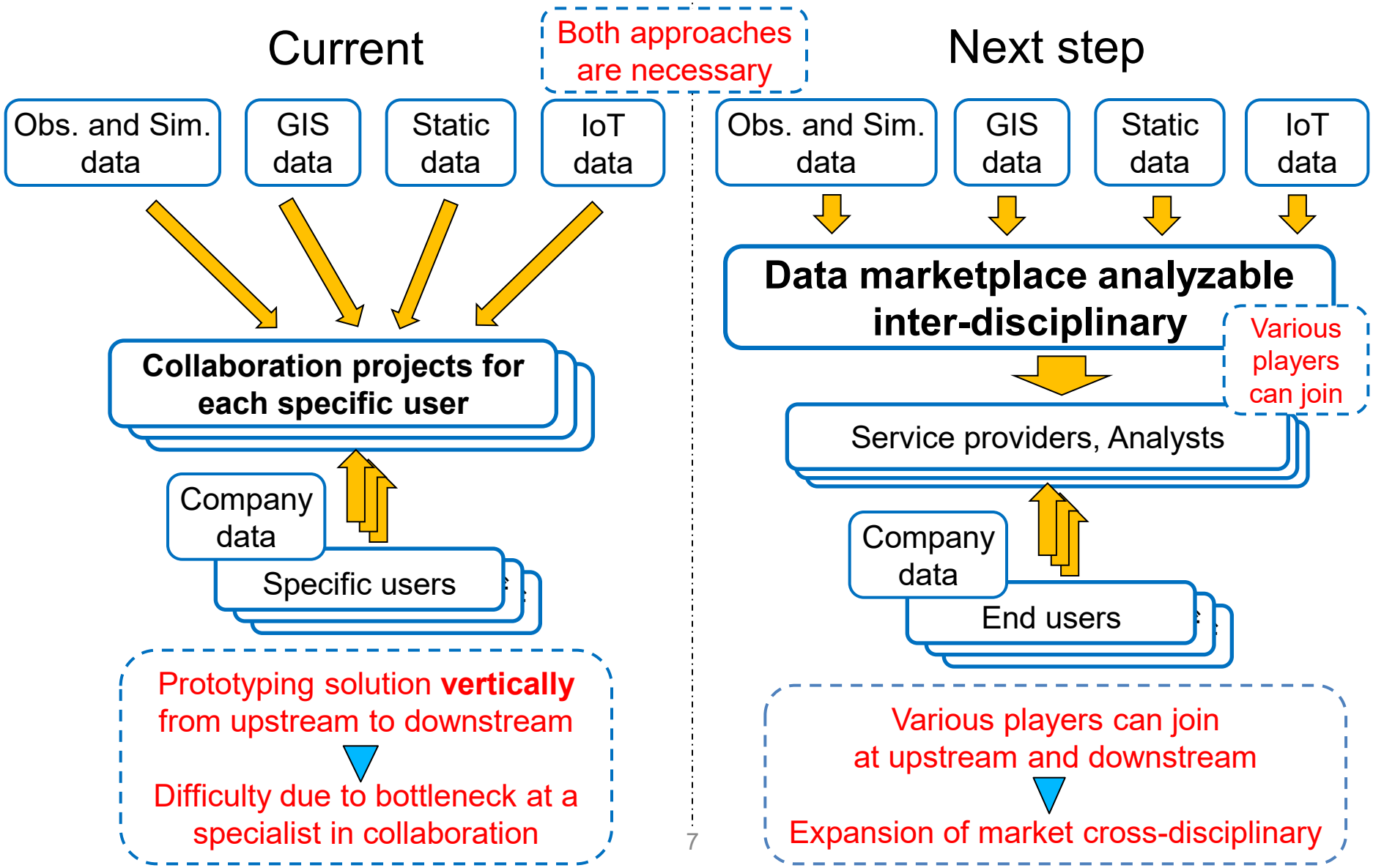
Difficulties for self-sustaining circulation of data ecosystem

- Incentive of data providers and ownership*
- Cost for storage and disclosure of data with enough security
- Finding and matching of Data
- Registration for each data provider
- Data format and the viewer depends on each discipline (see next page)
- Data Policy including Billing for commercial uses* (today's main topic)

Various efforts on data platform in each discipline or inter-discipline

(*: need to be authorized politically)

Inter/cross-disciplinary collaboration for Prototyping of Solutions



2. Why are almost research data limited for research purposes?

In spite of the government policy of open data, almost Research Data are research purposes only.

- Data provider wants to know who uses own data for what usage.
- Data provider is concerned about own data being misused.

(Provider (researcher) has no reserve capacity to help users.)

- Data provider does not always have ownership of the data. (collaborative projects, data compiled/ aggregating various prime data)
- There is no expert on data policy assuming commercial uses.
- Difficulty to respond to billing, donation etc.

Data policy + Goverment strategy

3. Creative Commons Licenses

(combination of BY, NC, SA and ND)

Icon	Description	Remark
	Freeing content globally without restrictions	Public Domain Metadata
	Attribution alone (give a link for 4.0)	Commercial OK, Derivatives OK
	Attribution + NonCommercial	Derivatives OK for Noncommercial use
	Attribution + ShareAlike	Derivatives OK in same licenses
	Attribution + NoDerivatives	Derivatives OK in same licenses for noncommercial
	Attribution + NonCommercial + ShareAlike	Shared OK without modify for Commercial
	Attribution + NoCcommercial + NoDerivatives	Shared OK without modify for Noncommercial
	NoCcommercial + Sampling	

Components of Data Policy (1/2)

- DIAS Open Science Study Group
  - Chaired by Asanobu Kitamoto（NII)
  - NII, Tokyo-U/EDITORIA, Kyoto-U, JAMSTEC, NIES, NiPR
- DIAS Data Policy Taskforce → MEXT Metadata WG (2016.3)

- Analysis from;
  - Copyright（著作権）
  - Moral rights of Authors（著作者人格権）
  - Creative Commons Licenses

- Compare and check with many current data policies in DIAS

23 data policies: AGURAM, AMY, CEOP, GRENE, JAMSTEC, JAMSTEC-TIT-NIED, JAMSTEC-NIED, RECCA, Ministry of Environment, JMA, JMA-Sousei, JMA-Himawari, Kyosei, NIES, MLIT, AIST, SIO, Government of Japan Standard Terms of Use, Sousei, TIT, NARO, Micro Geodata Forum, Nagoya-U

Components of Data Policy (2/2)

Components of Data Policy	Copyright/ Moral rights of Author	Creative Commons Licenses
1. Submitting user information		
2. Availability for commercial use		NC: NonCommercial
3. Availability of secondarily distributing non-modified data	Partially defined	Basically, CCL is for open sharing
4. Availability of secondarily distributing modified data	Integrity (Moral Right), Partially defined (Copyright)	ND: NoDerivatives SA: ShareAlike
5. Quotation display on the deliverables	Attribution (Moral Right)	BY: Attribution
6. Submission of deliverables using data		
7. Disclaimer / Usage Notes		
8. Billing conditions		
(9. Anonymity processing)		

4. Outline of the data policy guideline (1/2)

1. Submitting user information

   Almost researchers need user information but CC BY using Data DOI (or ORCID) is more convenient for both providers and users.

   (Registration for each data provider is bothersome. Inter-disciplinary data platform or mutual authentication Is desirable.)

2. Availability for commercial use

   At least, possibility study for commercial uses should be included in "Research purpose".

3. Availability of secondarily distributing non-modified data

   Almost data providers want to distribute own data from own data site for integrity.

Outline of the data policy guideline (2/2)

4. Availability of secondarily distributing modified data

   Recommend products of user's models using a data from which original data can't be reconstructed easily can be distributed by the user. (Major meteorological agency admits secondary distribution of products in which the data is used for initial/ boundary condition of re-analysis, prediction, downscaling using own model).

5. Quotation display on the deliverables

6. Submission of deliverables using the data

   CC BY using Data DOI (or ORCID) is more convenient for both providers and users. The data providers can search the deliverables on the internet..

7. Disclaimer / Usage Notes

8. Billing conditions

5. Data Trading Alliance

- Aims to establish a technical and institutional environment in which data providers can provide data safely and smoothly and users can easily find and collect data.
- Established based on the discussions of the Working Group of the “Cabinet Secretariat IT Comprehensive Strategy Office”, “Ministry of Internal Affairs and Communications” (METI) and “Ministry of Economy, Trade and Industry” (MIC).

  - Technical standard committee,
  - Operation standard committee,
  - Data utilization committee, etc.

Thank you!

If you are interested in “Proposal of Data Policy Guideline” (Japanese only),

please send E-mail to hajimen@jamstec.go.jp