

# WDSなどの国際データ活動の展望

村山泰啓,

NICT

渡邊堯

WDS-IPO、名大STE研、NICT

# Open (Research) Data

## **G8 Science Ministers Statement London UK, 12 June 2013**

### **Introduction**

We, the G8 Science Ministers met in London on Wednesday 12 June with Presidents of our respective national science academies, as part of the UK's G8 Presidency. At this unique meeting we discussed how our nations could lead efforts to improve the transparency, coherence and coordination of the global scientific research enterprise in order to address global challenges and maximise the social and economic benefits of research.

### **1. Global Challenges**

### **2. Global Research Infrastructure**

### **3. Open Scientific Research Data**

Open enquiry is at the heart of scientific endeavour, and rapid technological change has profound implications for the way that science is both conducted and its results communicated. It can provide society with the necessary information to solve global challenges. We are committed to openness in scientific research data to speed up the progress of scientific discovery, create innovation, ensure that the results of

### **4. Expanding Access to Scientific Research Results**

# 研究データのオープン化

---

- 科学研究におけるデータ公開：  
2つの文脈
  - 研究成果の再検証担保
    - ⇒先行して制度やツールの基本実証段階（出版社など）
    - ・やることそのものは、国際的に否定できない
  - 自由なデータアクセス・情報処理ツールによる新たな研究創出
    - ⇒目指すはデータ中心科学？等。
    - ブレークスルーへの努力中。
- 研究者が作成・管理の労力をかけたデータ
  - 欧米先行（EISCAT3D、[www.openradar.org](http://www.openradar.org)（MIT Haystack）...）。
  - 国・機関としての情報資産：
    - 保有・占有？ オープンサイエンスによる成果の総量？
  - 国内：
    - 途上。いまなら太陽地球系科学は先行例になれるのでは？

# Why Open Data, Open Access?

<http://www.getchemistryhelp.com/chemistry-lesson-scientific-method/>

## Traditional scientific method

Ask a question

Construct a hypothesis

Test with experiments

Analyze the results

Formulate a conclusion

- Sharing research information and evidence
- For open discussion and re-examination by third party.
- Reuse of information resources
- **Important is to ensure the mutual trust between Science and Society**

Various research information

Software code

Data

Research papers

Open discussion, Re-examination

Scientists, Community, Society

Toward next sciences



# 科学の体制(これまでは主に論文文献の世界)

## 研究事業

### 研究主体・科学研究者



## 研究情報の出版・保存・利用事業

### 出版社・出版組織



### 「ライブラリ」、 学術研究情報組織



### 研究情報マネジメント、 インフラ構築



研究結果、データ、メタデータ、のやりとり



# 「研究」側の事業例：EUDAT



## EUDAT

- Start date: 1st October 2011
- Duration: 36 Months
- Budget: 16.3 M€ (9.3M€ EC)
- EC Call: INFRA-2011-1.2.2
- Consortium: 25 partners from 12 countries
  - National data cent
- Objectives:
  - Cost-efficient and
  - Meetings users' ne
  - Across geographic

<http://www.eudat.eu>

## European Data



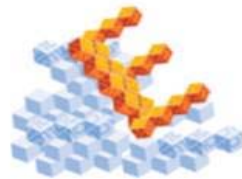
Researchers, citizens,  
Industry and society...

EC-Funded

Promote  
“Open Data”

EISCAT-3D joins

## User Forums + 25 communities



1st User Forum  
7-8 March 2012,  
Barcelona



# 「研究情報マネジメント」「出版・保存・公開」側を中心にした事業・機関間リンクの一側面

ICSU-World Data System



**WDS-IPO**  
Hosted by **NICT**  
Ionosphere, Space Weather

**Earth Science**  
Space Physics  
Geography  
Atmosphere  
Agriculture  
Informatics

**IRIS** Seismology

**Earth Sciences**  
Atmosphere & Ocean

**Space Physics**  
Computer Science, Physics, etc.

**CLARIN**  
Common Language Resources and Technology Infrastructure  
Linguistics, History, Psychology

**ICPSR** INTER-UNIVERSITY CONSORTIUM FOR POLITICAL AND SOCIAL RESEARCH  
Social Sciences

**WILEY** Publishing  
**ELSEVIER**

...76 members in total.



Future Earth  
(ICSU, UNESCO, UNEP, UNU, Belmont Forum,...)



GEOSS/DIAS  
(U of Tokyo, JAXA, JAMSTEC etc)



Science Council of Japan



Council for Science and Technology Innovation, Japan

RDA  
(Research Data Alliance)



Open Data Charter, Open Research Data  
(Science Minister Meeting, June 2013)



Joint WG/  
MOU planned



# データ・マネジメント組織の例

- ICSU-WDS
  - 学術会議のような立場。50年スパンでの推進。委員会活動が主。具体的作業は下部組織で。
- RDA
  - 政府関係の推進力が背後にあり(G8+06)
  - 具体的・短期的(1.5年)目標について自由参加。
- DataCite
  - データサイテーションのコンソーシアム、自己資金拠出。出版事業者、研究機関など。
- 社会科学系
  - Data Seal of Approval, DataVerse, 多数先行例。



# こんなこと(データマネジメントの仕事)に関わって、メリットがあるのか？

- 科学と社会(国、公共)の関係を考える上で重要。
- 国際的な議論・情報が先行して入手しやすい。
- 「標準、スタイル」を決める現場に立ち会う  
→ 先行的に仕事を進められる。今後を見通しやすくなる。
- 分野限定でない議論が多い  
→ 研究、サイエンス全体のためになる。その目的を掲げて、リソース獲得を狙える(のでは)。
- これらを通じて、自分たちの国やコミュニティの活動を強化できる(のでは)。

# 我々の周りでのオープンデータ関連動向



- ICSU-WDS: 国際オフィスの国内設置(2012年3月～)
  - アカデミーとしての組織化・調整機能、下部構造が具体化作業。



- RDA設立 ←→ G8における研究データ施策の議論
  - 科学大臣会合: 研究データのオープン化声明(2013年6月12日)
  - G8+O6 Data Infrastructure WGが2011年頃から活動



- 国立国会図書館(NDL)の電子化、電子アーカイブ事業
  - 「国立国会図書館の資料デジタル化に係る基本方針」(2013.5)等
  - 電子図書館事業の推進、NDLデジタルアーカイブシステム

- 大学付属図書館、機関リポジトリ

- 文献を対象に普及、運用安定化・DOI付与などが進行中。

- 科学技術振興機構(JST): 科学技術情報委員会

- 内閣府・文科省へのインプットを目指す(産業界、学界...)

- 今後の方向性: 情報提供事業から情報循環事業へ変革したい。

- ジャパン・リンクセンター(JaLC)におけるDOI付与

- DOI付与権限をもつ国内機関。日本ではここだけ。(JST内)

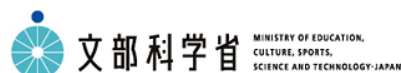
- JaLC運営員会コンテンツ拡大検討分科会(小山、村山)

- 文科省科学技術・学術政策研究所(NISTEP)

- 第5期科学技術基本計画と研究データオープン化の議論

- JpGU/PEPS、EPS

学術機関リポジトリ構築  
NII Institutional Repositories Program



# Steps by Major scientific publishers encouraging data deposition

---

- **Wiley/AGU** publication policy:  
”...in AGU’s journals, **all data** necessary to understand, evaluate, replicate, and build upon the reported research **must be made available and accessible whenever possible...**”
- **SpringerOpen/**”Earth, Planets and Space”, “Geoscience Letters” ...  
“...Electronic archiving of data enables readers to replicate, verify and build upon the conclusions published in papers in the journal. **It is recommended that all data** which are not directly attached to a publication as electronic supplementary files **be deposited...**”
- **Elsevier/JASTP**:  
“...Elsevier encourages **authors to deposit raw experimental data sets** underpinning their research publication in data repositories, and to enable interlinking of articles and data...”

# WDS's Practical Steps to Go 今後の重点課題

Working Groups has been Established.

WGを設置して、下記課題を議論

WDS-WG



RDA/WDS-  
joint WGを  
併用

## 1. Data Publication, Data Citation

⇒RDAジョイントWG(共通メンバー多数)

Standards, Practices: データ出版・参照(サイテーション)、  
管理運用の枠組を議論。

- 学術出版社、図書館もふくめた連携:

Elsevier, Thomson-Reuters, Wiley ,Springer,...

WDS-WG



RDA/WDS  
Joint化も  
検討

## 2. Open Metadata Catalogue

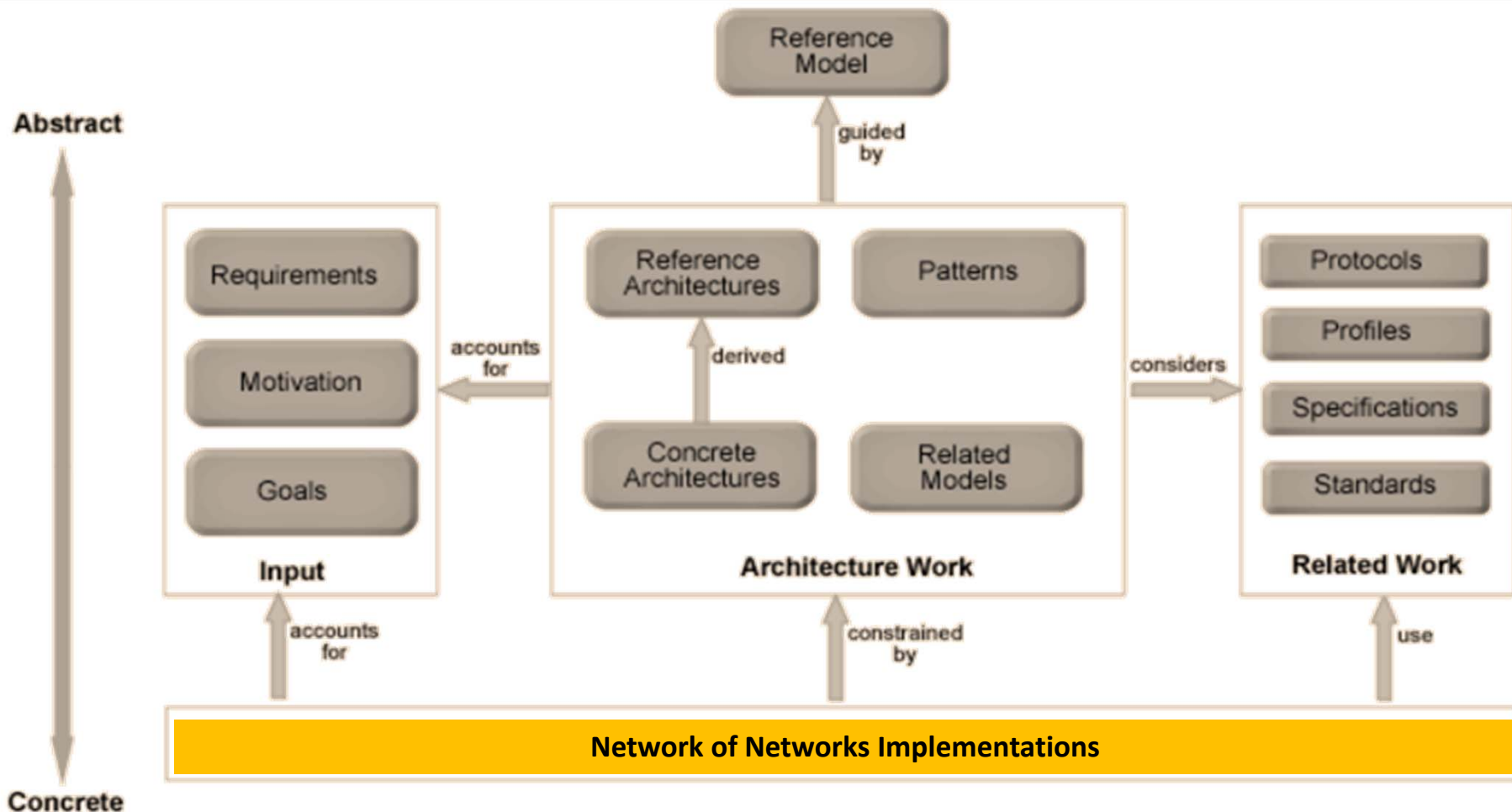
WDS加盟機関のグローバル検索を可能に。

## 3. Scalable Knowledge Network

将来の国際データネットワークのための最適な  
メタデータ利用や参照モデルを議論。

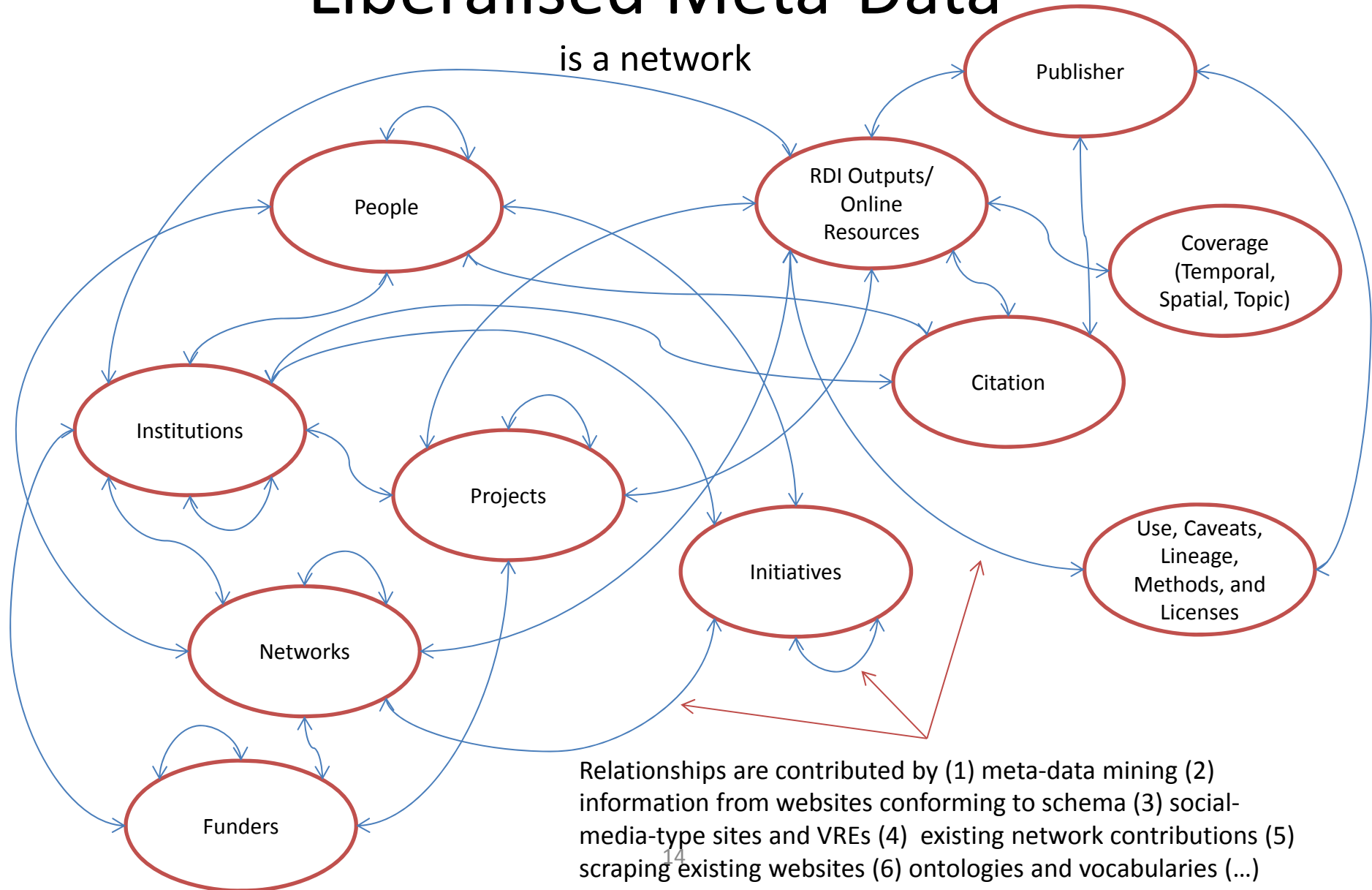
# Reference Models

Adapted from 'Reference Model for Service Oriented Architecture', OASIS, Committee Draft 1.0, 7 February 2006. <http://www.oasis-open.org/committees/download.php/16587/wd-soa-rm-cd1ED.pdf>



# Liberalised Meta-Data

is a network

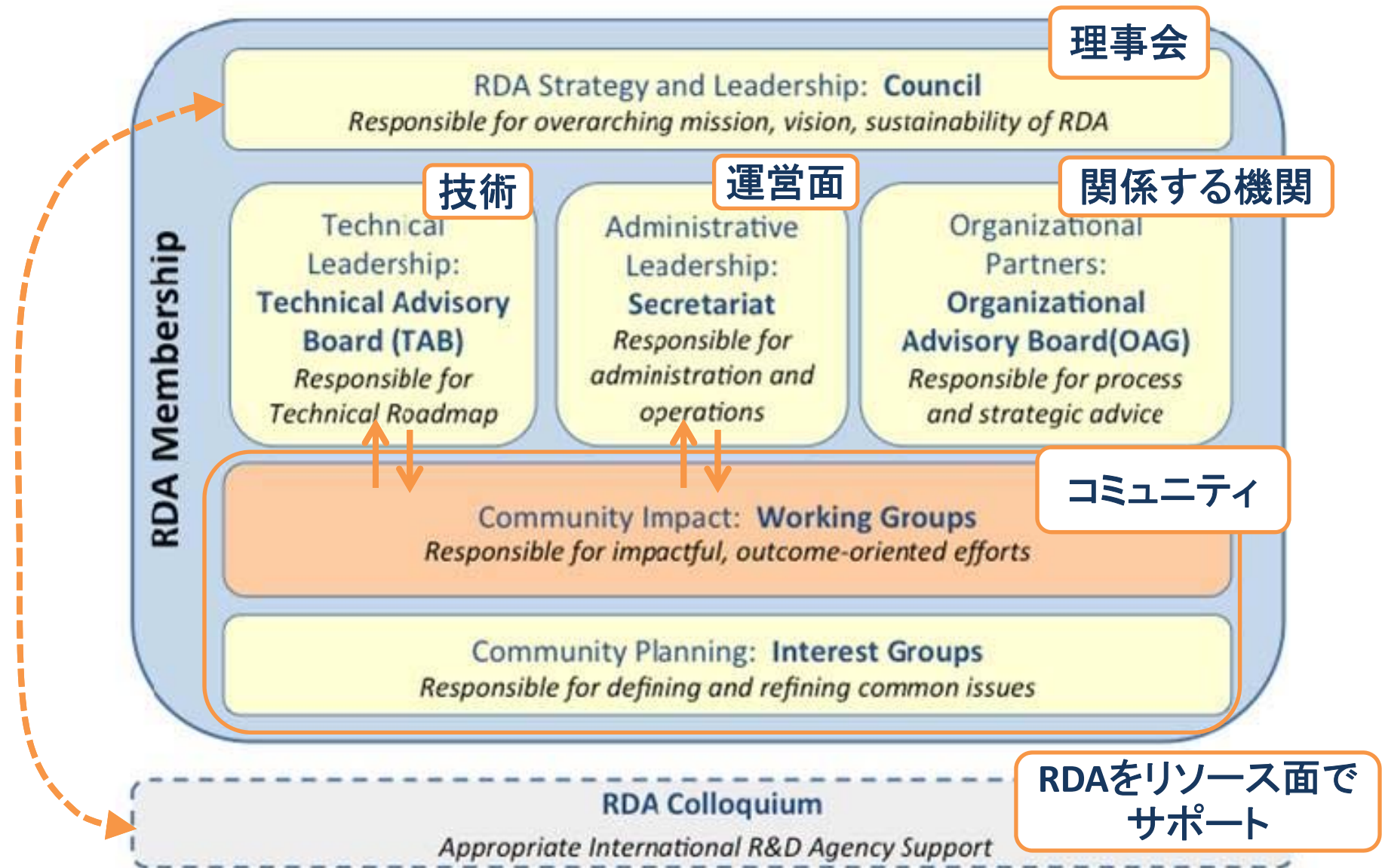


# RDA (Research Data Alliance) について

- 研究データ共有を加速し、実験・プラクティスを実現していくコンソーシアム。
  - 2013年3月正式発足。
  - 米、欧、豪がそれぞれにローカルオフィスを設置
  - Mark Parson(米)がSecretary General(2014年3月～)
  - G8・GSO (Group of Senior Officials) 下のデータWG議論が契機。
- 組織モデル: IETF (Internet Engineering Task Force) の手法を、科学データに適用。
  - 研究者・技術者による自由な参加。ボラティア・合意ベース。
  - 実質的な国際標準・国際相互結合理体制の形成を目指す。
  - ICSU、CODATA、WDS、社会科学分野などとも協力・連携  
(国際的に活躍する専門家はだいたい限られているが...)

# RDAの組織構成、ガバナンス

(オレンジ色枠は村山加筆)



[Mark Parsons, 2013]



# Groups that Met at the RDA Plenary

テーマ提案・検討中

BOLD = new since last Plenary

## Birds-of-a-Feather

- Linked Data
- Chemical Safety Data
- Education and Skills Development in Data Intensive Science
- Libraries and Research Data
- Cloud Computing and Data Analysis Training for the Developing World

## Working Groups

- Data Type Registries
- Metadata Standards
- Practical Policy
- Persistent Identifier Types
- Data Foundations and Terminology
- Data Categories and Codes

WG検討中

## Interest Groups

- Agricultural Data
- Big Data Analytics

ミッション  
遂行  
(18か月)

- Data Brokering
- Certification of Trusted Repositories (joint with ICSU-WDS)
- Long tail of Research Data
- Marine Data Harmonization
- Community Capability Model
- Data Publishing (joint with WDS)
- Toxicogenomics Interoperability
- Research Data Provenance
- Data Citation
- Metadata
- Economic Models and Infrastructure for Federated Materials Data Management
- Engagement
- Preservation e-Infrastructure
- Legal Interoperability (joint with CODATA)

- Global Registry of Trusted Data Repositories and Services
- Digital Practices in History and Ethnography
- **Data Citation Harmonization Summit**
  - DataCite, FORCE11, CODATA/ICSTI, ESIP, DCC, etc.

[Mark Parsons/RDA-US, 2013]

**Info:**  
[enquiries@rd-alliance.org](mailto:enquiries@rd-alliance.org)  
[@resdatall](https://twitter.com/resdatall)

# 今後の課題

- 国内体制の構築
  - リポジトリ、解析、サービス
  - 分野ごと
  - 複数分野・全体ポリシー
- 国際的な共用システムへむけて