

資料5

第3回科学技術情報整備審議会  
平成25年7月25日

# *NDL & NII*

喜連川優  
国立情報学研究所

# NII ミッション

*情報学基礎研究*

+

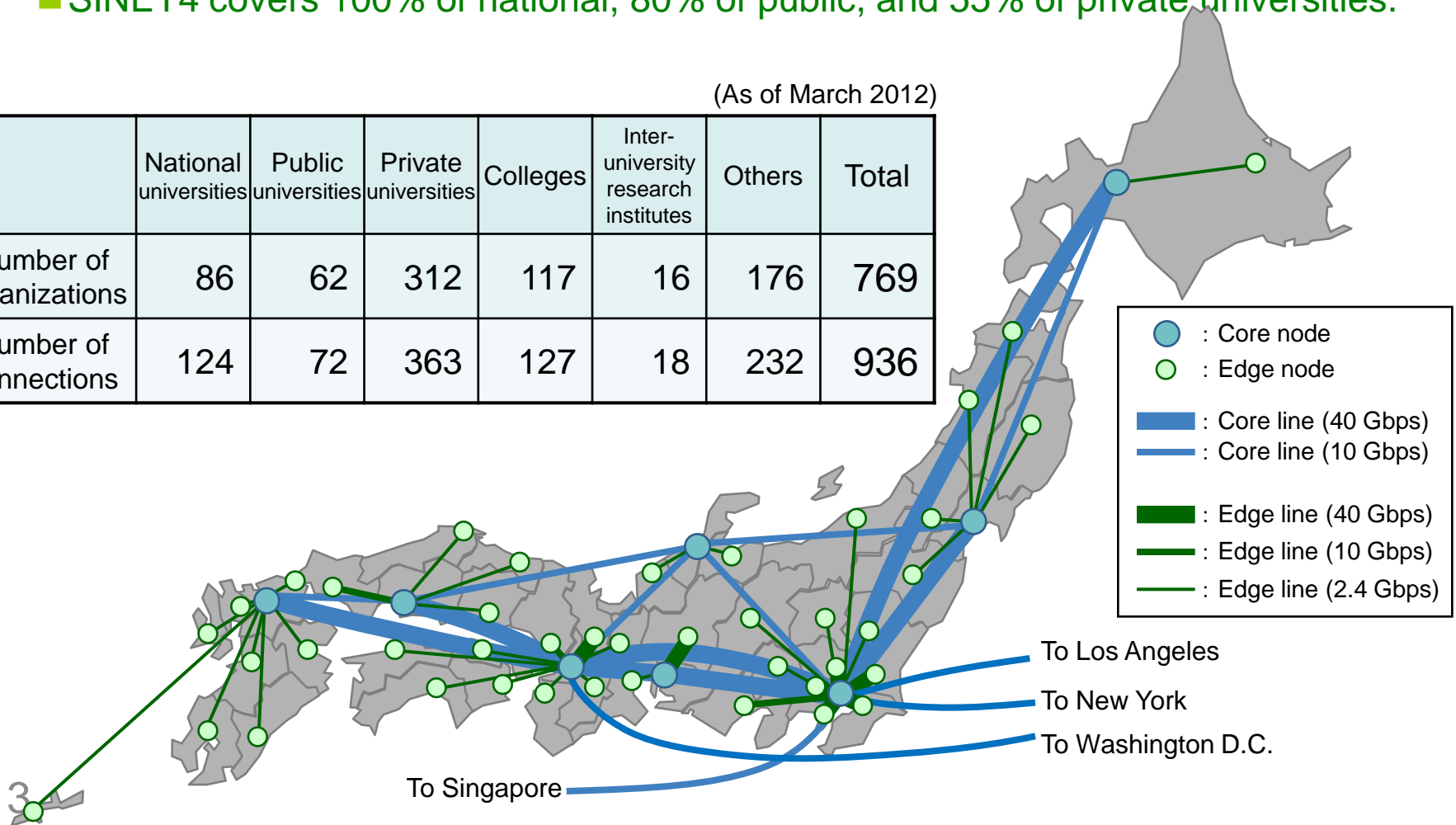
*情報基盤(SINET)の運用*

# Science Information Network (SINET)

- ❑ Academic backbone network for more than 700 universities and research institutions and more than 2 million users.
  - SINET4 covers all 47 prefectures. (SINET3 did 34 prefectures.)
  - SINET4 covers 100% of national, 80% of public, and 55% of private universities.

(As of March 2012)

	National universities	Public universities	Private universities	Colleges	Inter-university research institutes	Others	Total
Number of organizations	86	62	312	117	16	176	769
Number of connections	124	72	363	127	18	232	936



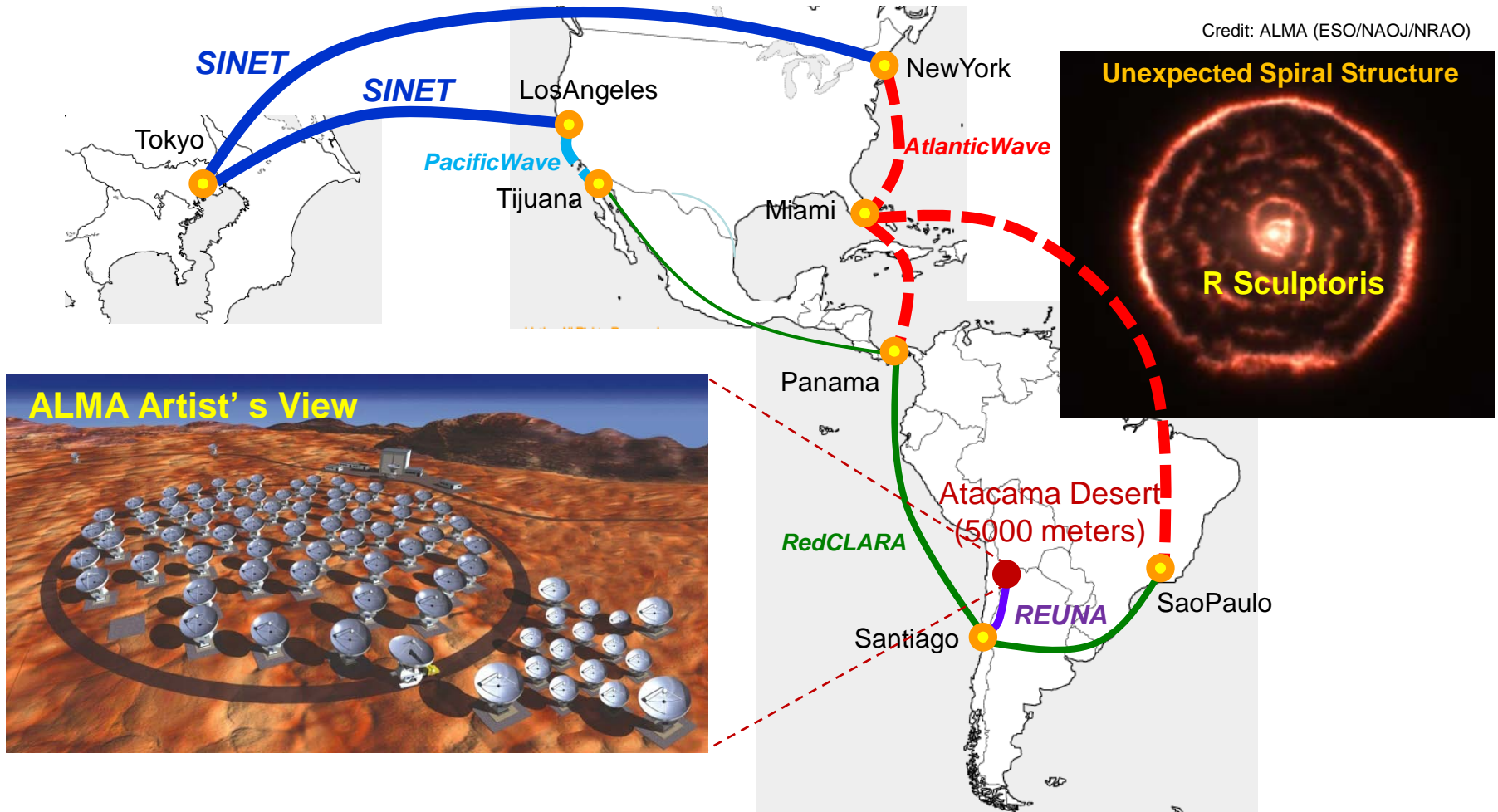
# Usage Example in ALMA Telescope

- The ALMA project is a global partnership of Europe, North America and East Asia (led by Japan) in cooperation with Chile to operate a ultra-high performance radio telescope consisting of 66 high-precision antennas. The Initial operation started by using 16 antennas in September 2011.

ALMA: Atacama Large Millimeter/submillimeter Array

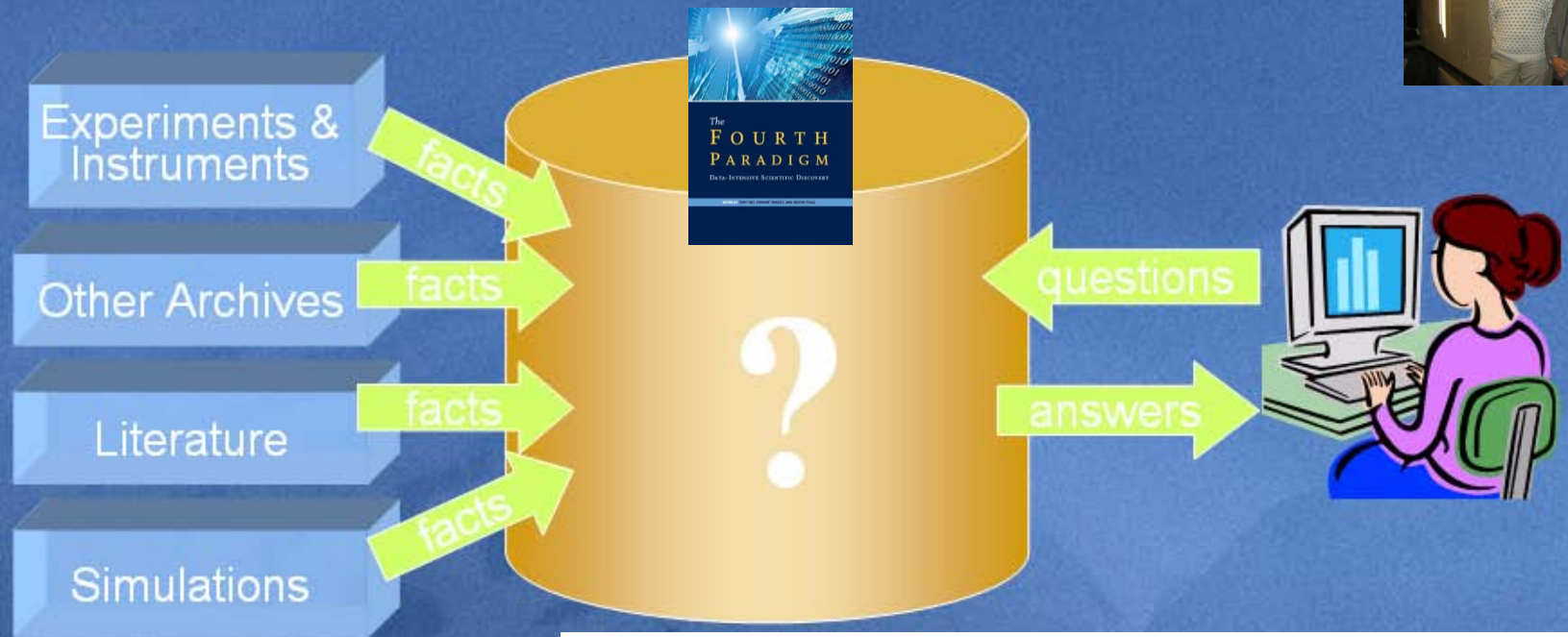
Source: National Astronomical Observatory of Japan (NAOJ)

Credit: ALMA (ESO/NAOJ/NRAO)



# Big Data

# The Problem for the e-Scientist



## 大規模データ基盤の有無が生命線

- ◆ Data ingest
- ◆ Managing a petabyte
- ◆ Common schema
- ◆ How to organize it?
- ◆ How to reorganize it?
- ◆ How to coexist & cooperate with others?
- ◆ Data Query and Visualization tools
- ◆ Support/training
- ◆ Performance
  - Execute queries in a minute
  - Batch (big) query scheduling

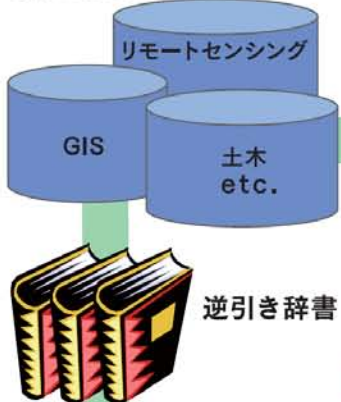
DIAS :  
地球環境情報統融合データベース

Only ONE

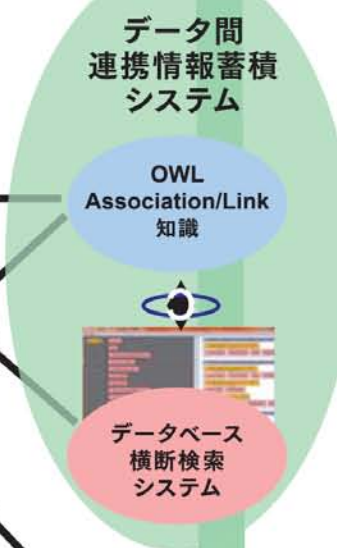
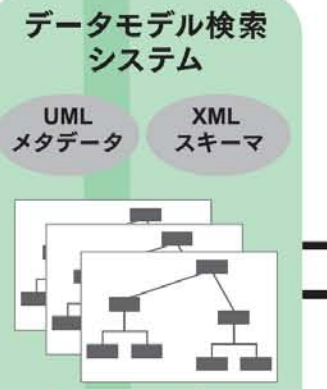
Non-Portal!

# データ統合・解析システム ⇨ 高度化・拡張

専門用語辞書



データ・情報の超多様性・複雑な関連性

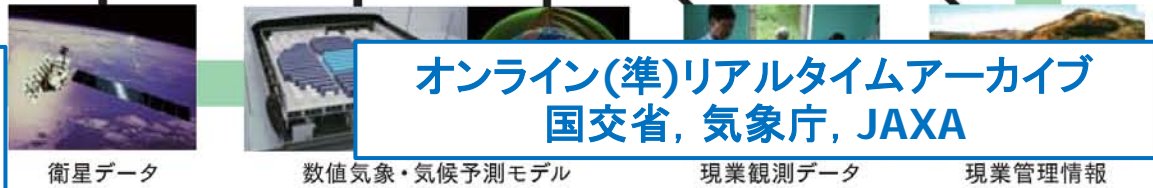


データセンタ間  
大規模データ  
統融合機能:20

様々な情報源からの超大容量データ

データアーカイブシステム改良

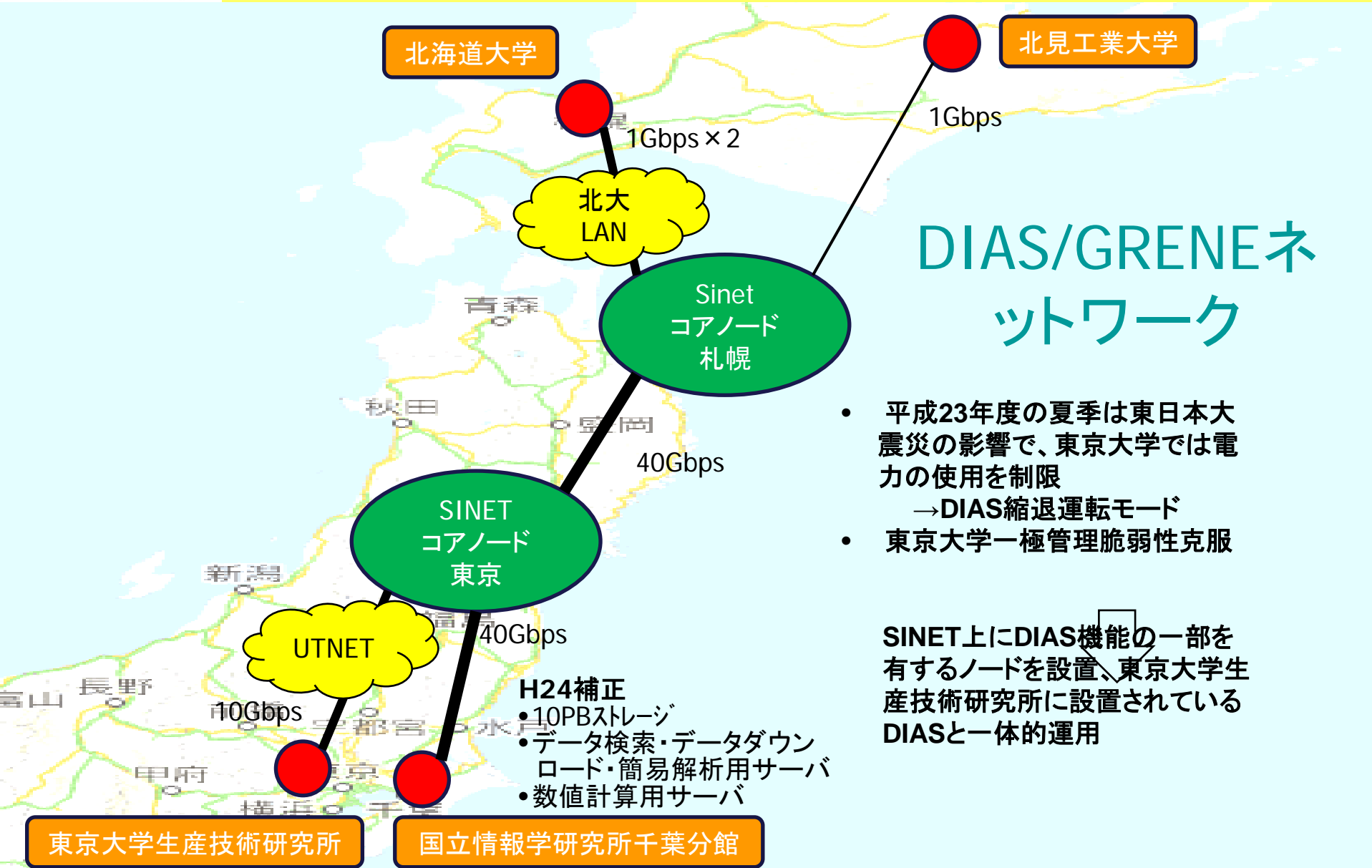
- GEOSSアジア・アフリカ:248
- アジアモンスーン年:317





# 地球環境情報統融合プログラム(DIAS-P)

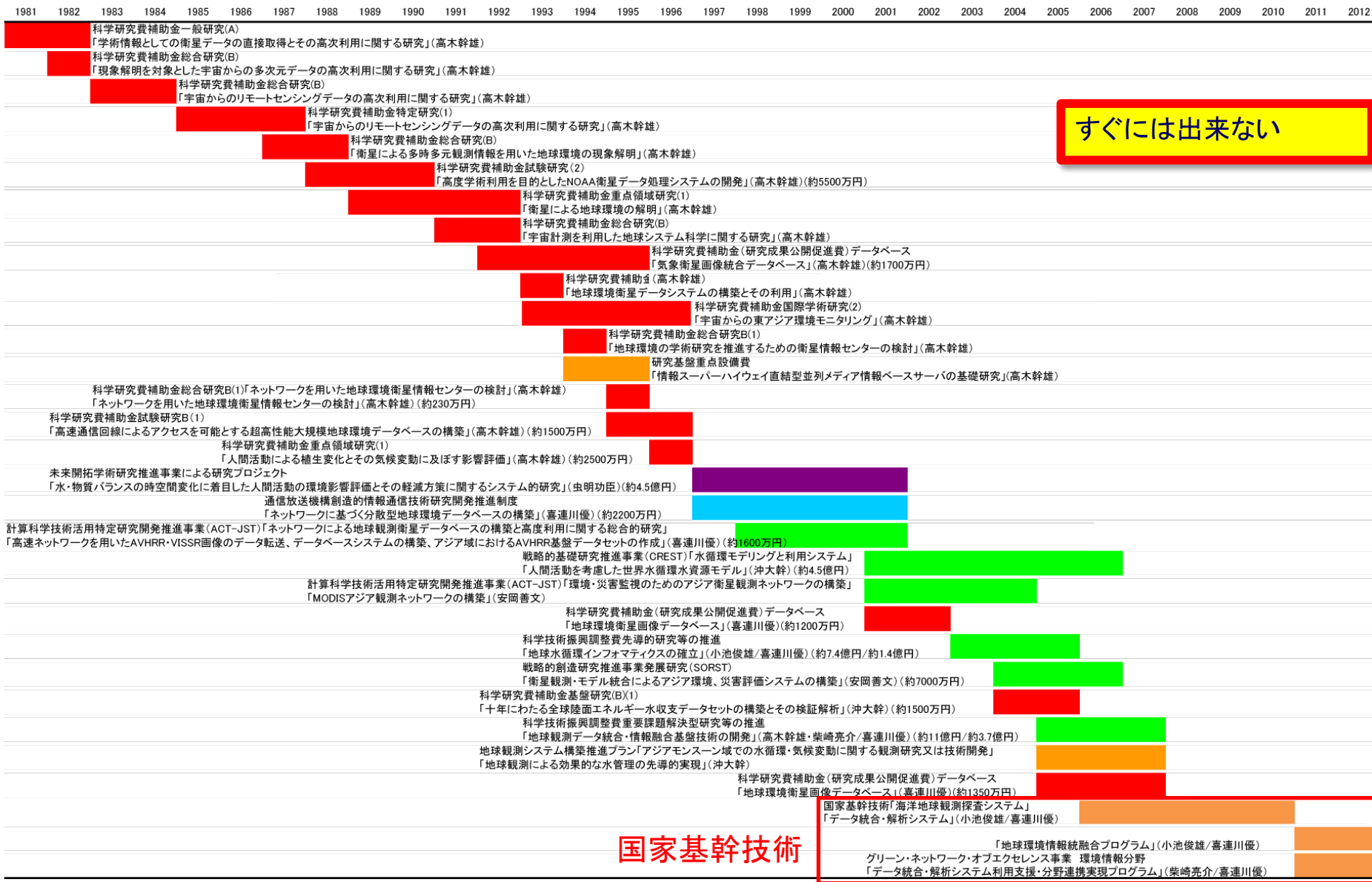
## 中間評価時報告: 計画外事象への対応



- 平成23年度の夏季は東日本大震災の影響で、東京大学では電力の使用を制限  
→DIAS縮退運転モード
- 東京大学一極管理脆弱性克服

SINET上にDIAS機能の一部を有するノードを設置、東京大学生産技術研究所に設置されているDIASと一体的運用

# History of funding



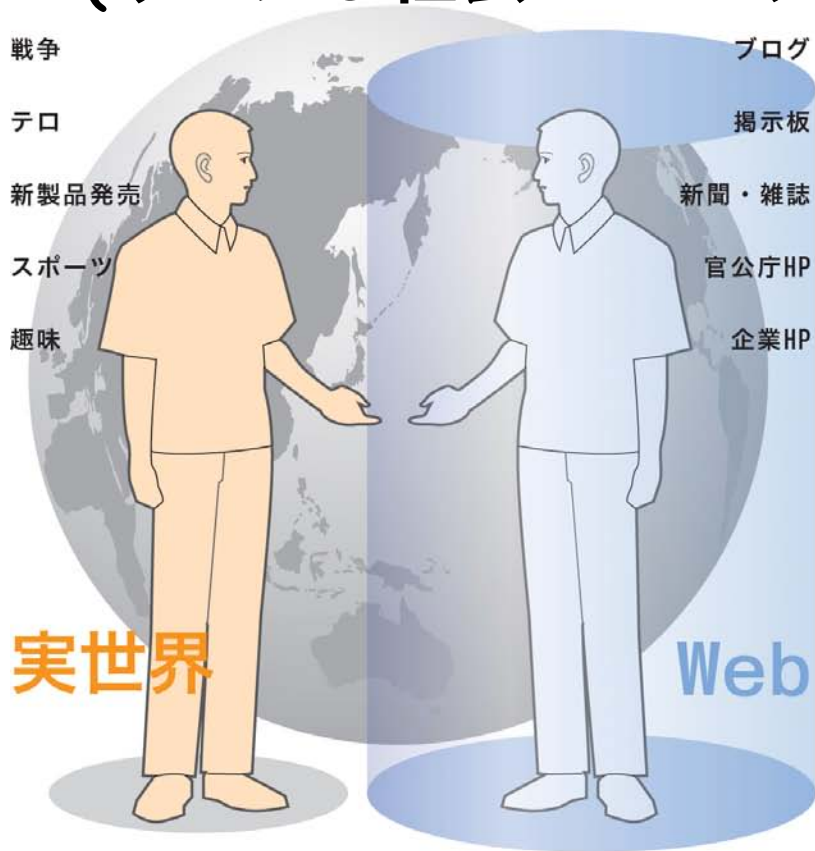
すぐには出来ない

国家基幹技術

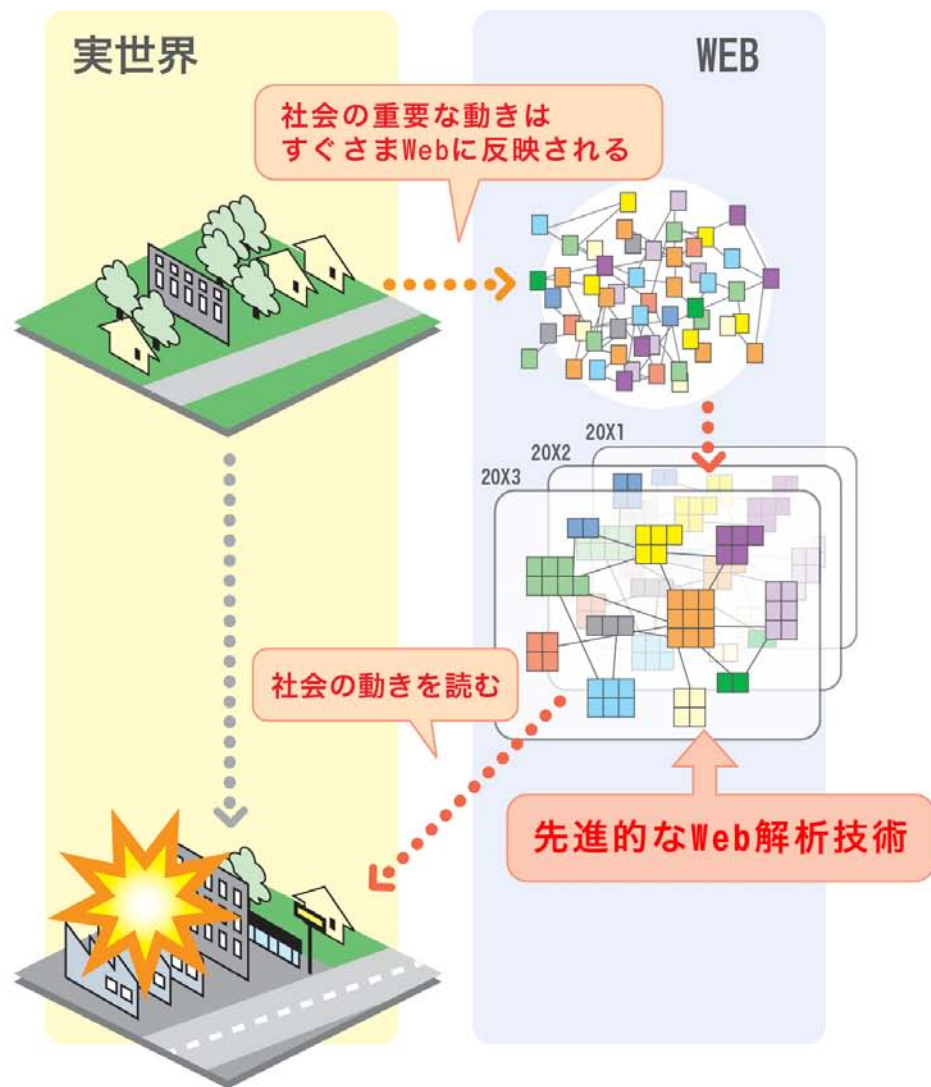
# Socio Sense:

## 実社会の射影としてのITメディア

(ウェブは社会のセンサ)



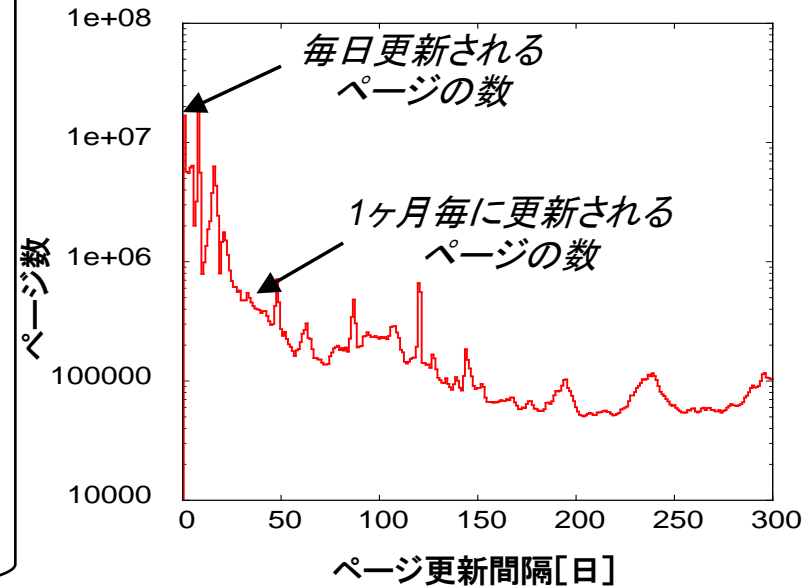
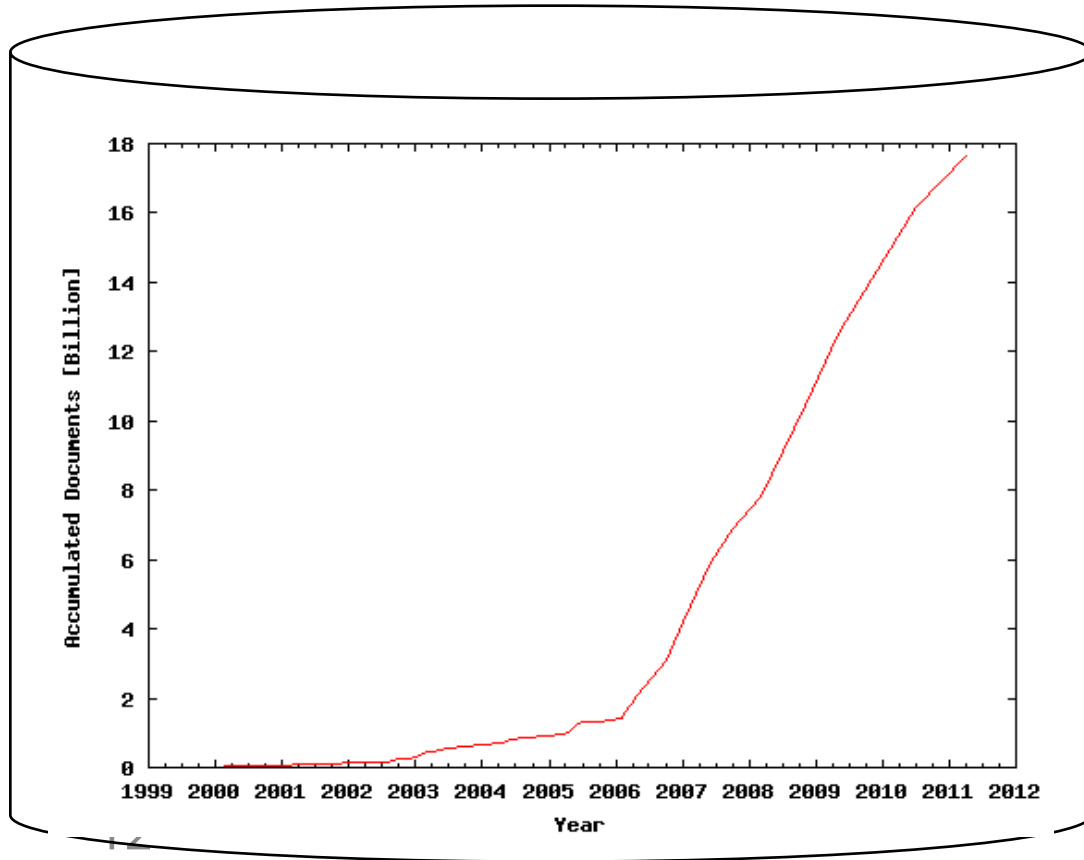
Webは実世界の転写構造を形成



# 日本語ウェブアーカイブの構築

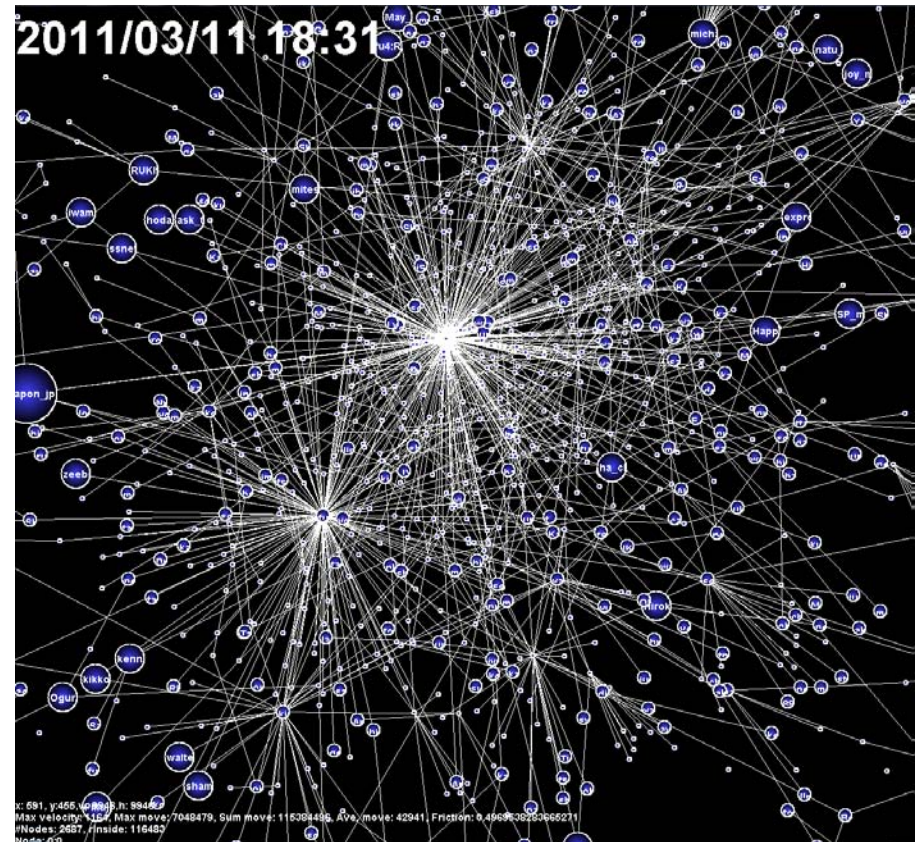
- 12年間にわたり220億ページ(2012年3月)規模の日本語ウェブページを集積し、継続期間および規模において**アジア圏最大級**のウェブアーカイブを構築
- 各URLの更新頻度に応じた可変周期収集

すぐには出来ない



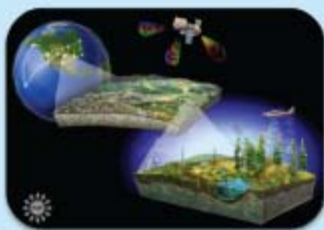
# 震災直前からのTwitter アーカイブ

- 約50万ユーザのtweetsを  
3/11午前中まで遡って収集、現在も継続中
- 累計 約100億tweets
- キーワード検索及び、  
リツイート・@関連の  
追跡・可視化



# NSF Keith Marzulloより

## Big Data and national priorities



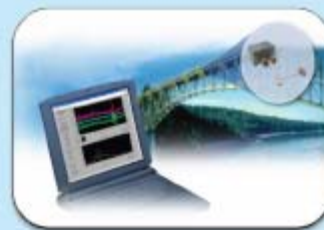
**Environment & Sustainability**



**Health & Wellbeing**



**Emergency Response & Disaster Resiliency**



**Manufacturing, Robotics, & Smart Systems**



**Secure Cyberspace**



**Transportation & Energy**



**Education & Workforce Development**



Big Dataは継続的な維持が大切  
more active maintenance is a must than books

先進的ITを駆使出来るかどうか  
が  
生命線

NDL Services



```
graph TD; NII[Powered by NII] --> NDL_Services[NDL Services (production)]; NII --> NDL_Lab[NDL Lab. (ambitious services)]; NDL_Lab --> NDL_Services;
```

NDL Services  
(production)

NDL Lab.  
(ambitious  
services)

Powered by NII

Big  
Data

NDL Services  
(production)

NDL  
Lab.  
(ambitious  
services)

Univ.  
and  
Research  
Institute  
s  
and  
Agencies

Powered by NII & Coordinated by  
NII

# CSIとNIIのコンテンツ事業

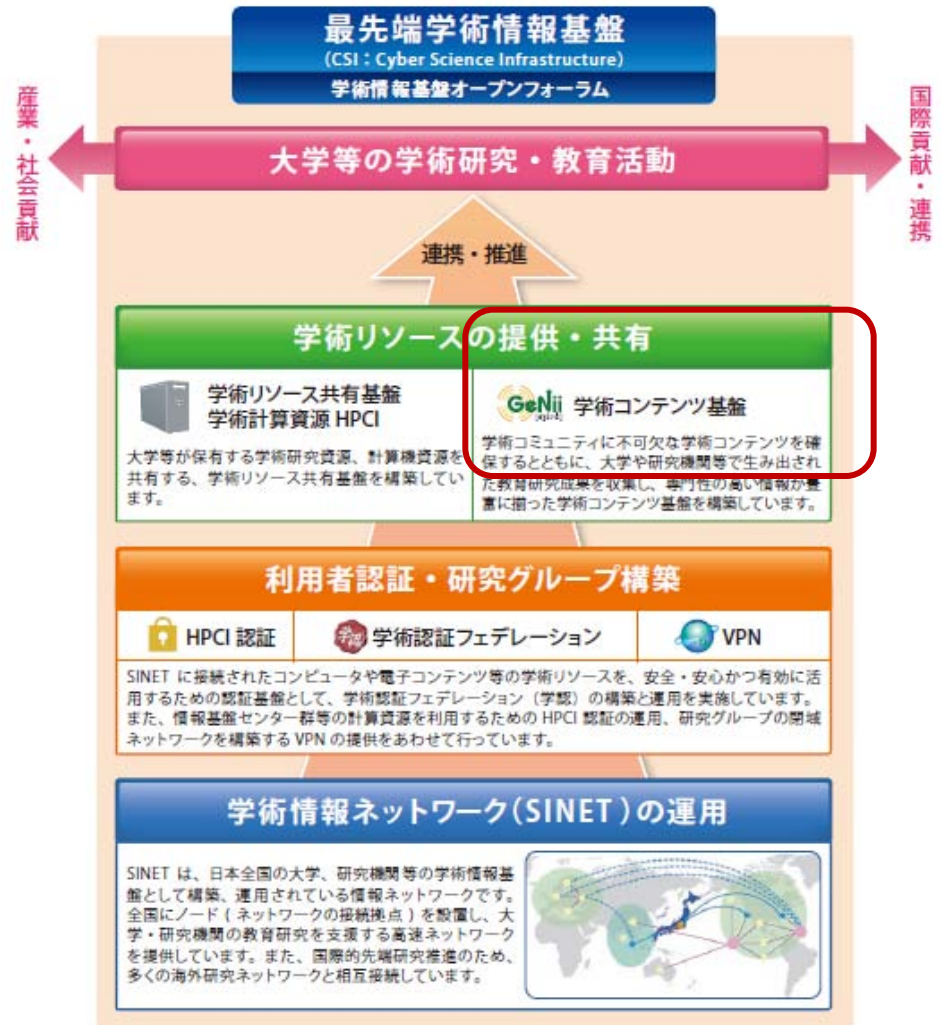
# CSIとは？

## ■ CSI (Cyber Science Infrastructure)

- 我が国の大学等の研究機関が有しているコンピュータ等の設備, 基盤的ソフトウェア,  
コンテンツ及びデータベース,  
人材, 研究グループそのものを超高速ネットワークの上で共有する「最先端学術情報基盤」

## ■ 学術コンテンツ基盤

- 大学等の教育・研究機関の構成員(研究者・学生)が必要とする学術コンテンツを, 学術コミュニティと連携して確保・形成し, 付加価値を付けて提供することで, 教育・研究に貢献する



# 学術コンテンツ基盤：学術コンテンツの確保と提供

提供

## NIIの提供する学術コンテンツ

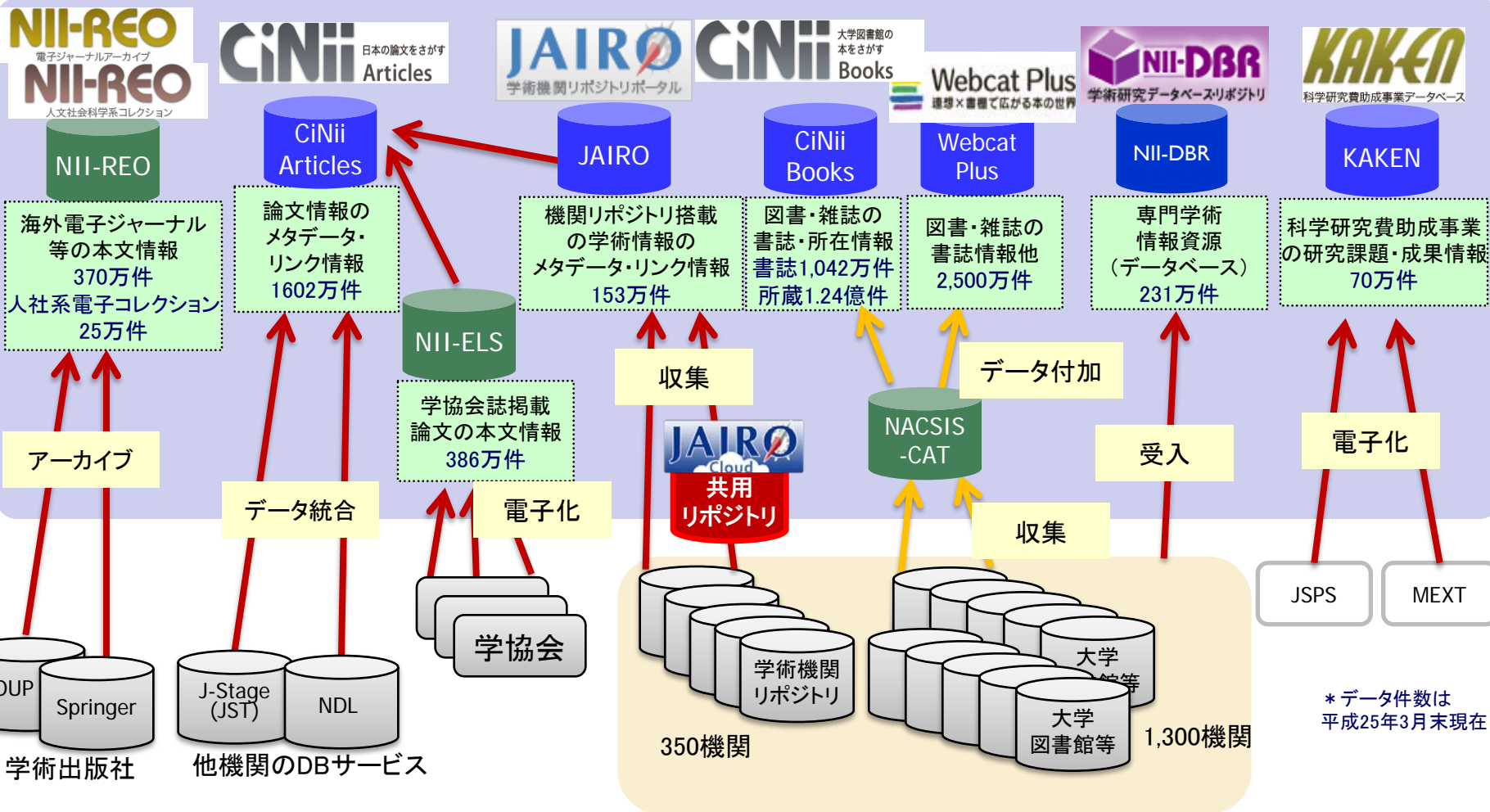
論文情報

機関発信情報

図書・雑誌情報

専門学術情報

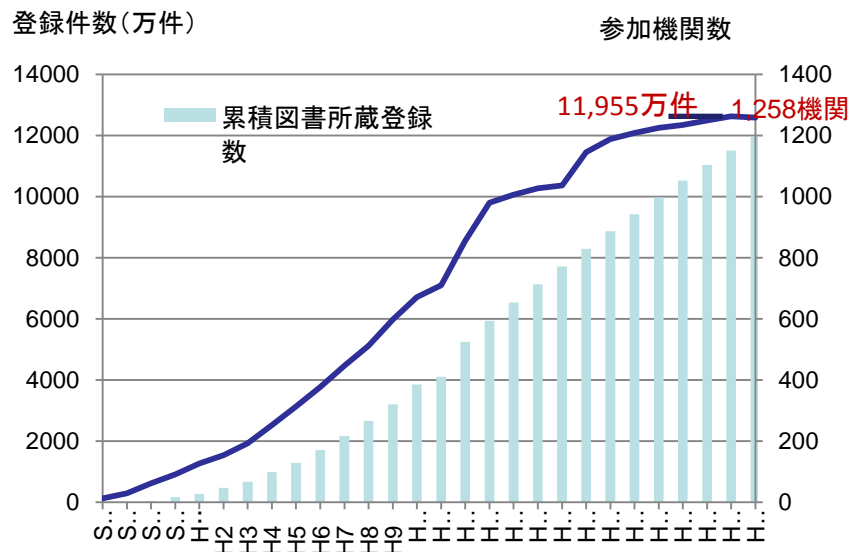
研究情報



# 目録所在情報サービス—CAT/ILL

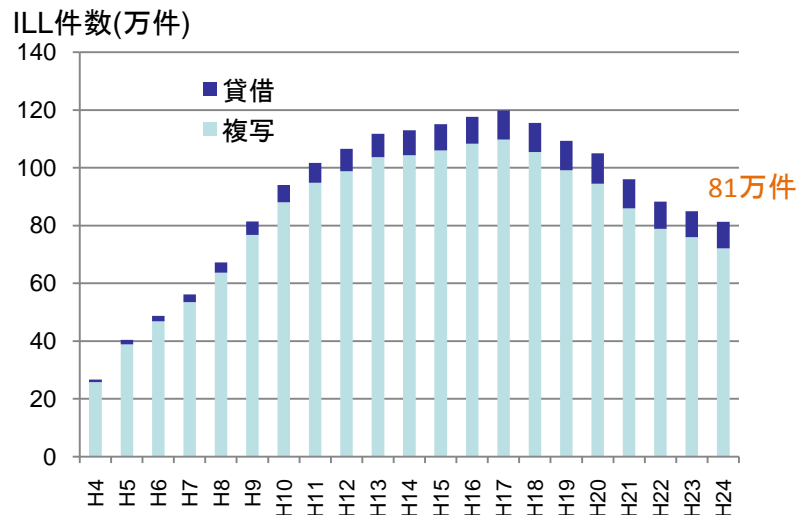
## □ NACSIS-CAT (総合目録データベース)

- 書誌レコード 約1,042万件
- 所蔵レコード
- 約1億1,955万件
- 参加機関 1,258機関



## □ NACSIS-ILL (図書館間相互貸借システム)

- 複写年間 約72万件
- 貸借年間 約9万件
- 海外ILL (OCLC, KERIS)



# 学術リソースナビゲーターCiNii

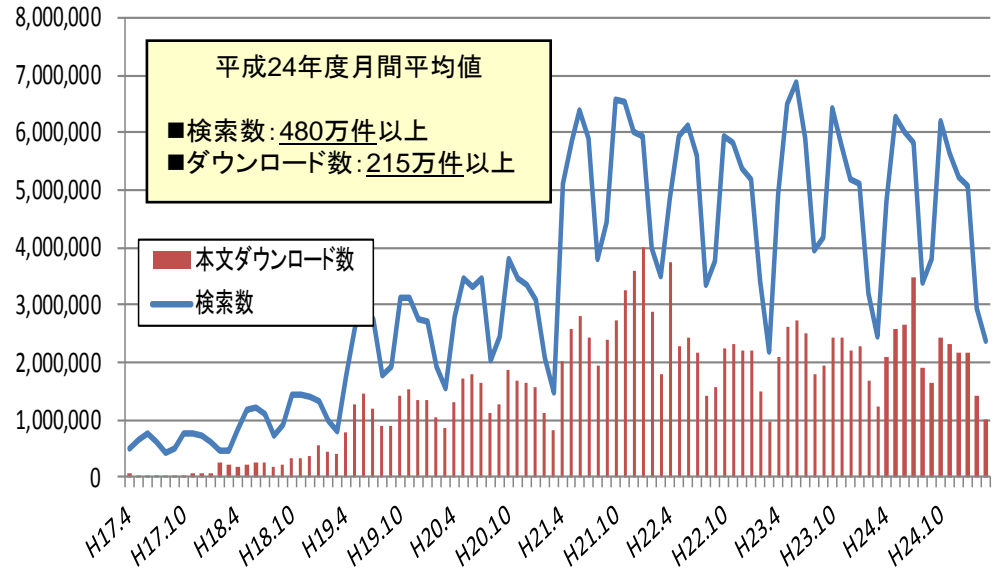
## □ CiNii Articles

- 日本の学協会誌・紀要の論文情報を網羅的・統括的に提供
- 収録誌数 約18,500誌  
収録件数 約1,602万論文(本文あり386万件)
- 学協会誌の約半分がオープンアクセス、さらに定額制により9割が個人の課金なしで利用可能
- 国内4年制大学の約7割以上が定額制を利用
- **NDLからは「雑誌記事索引」の提供をいただき、活用させていただいております。**

## □ CiNii Books

- 平成23年11月にNACSIS Webcat(平成25年3月8日終了)の後継サービスとして公開
- 目録所在情報サービス(NACIS-CAT)に蓄積された全国の大学図書館等が所蔵する図書・雑誌の情報を検索可能。  
書影、内容説明・目次も表示。
- 書誌情報 約1,010万件  
所蔵情報 約1億1,955万件  
図書館 1,258館
- 著者検索、図書・雑誌検索の拡充により、論文情報ナビゲータから学術リソースナビゲータへ

CiNii Articlesの利用状況(平成24年度末)



CiNii Booksの利用状況(平成24年度末)



# 科学研究費助成事業データベース—KAKEN

## □ KAKEN

- 科学研究費助成事業の採択課題・研究成果を蓄積
- 科学研究費助成事業は全ての学問領域にわたって幅広く交付されており、本データベースによって、我が国における全分野の最新の研究情報について検索することが可能

### □ 収録データ

採択課題	約70万件(1965年～)
実績報告	約84万件(1985年～)
成果概要	約15万件(1985年～2007年)
研究成果報告書・自己評価報告書	約5万件(2008年～)

## 学術コミュニティ・社会

科研費研究成果の発信と還元





# 機関リポジトリ (Institutional Repository) と オープンアクセス (OA)

# 機関リポジトリの役割

## 機関リポジトリの二つの役割

オープンアクセスによる情報の公開・提供

### ■ セルフアーカイブ

#### 研究者コミュニティへの貢献

- ◆ E-journalとして収録されている論文のアーカイブ
- ◆ 日本の学術論文の執筆数を考慮すると、日本は世界の約8.1%を生産しこの割合で貢献しなければならない
- ◆ STM (Science, Technology and Medicine) 分野では、業績データベースや科研の申請書に記載するような成果論文を公開していくことに該当する。

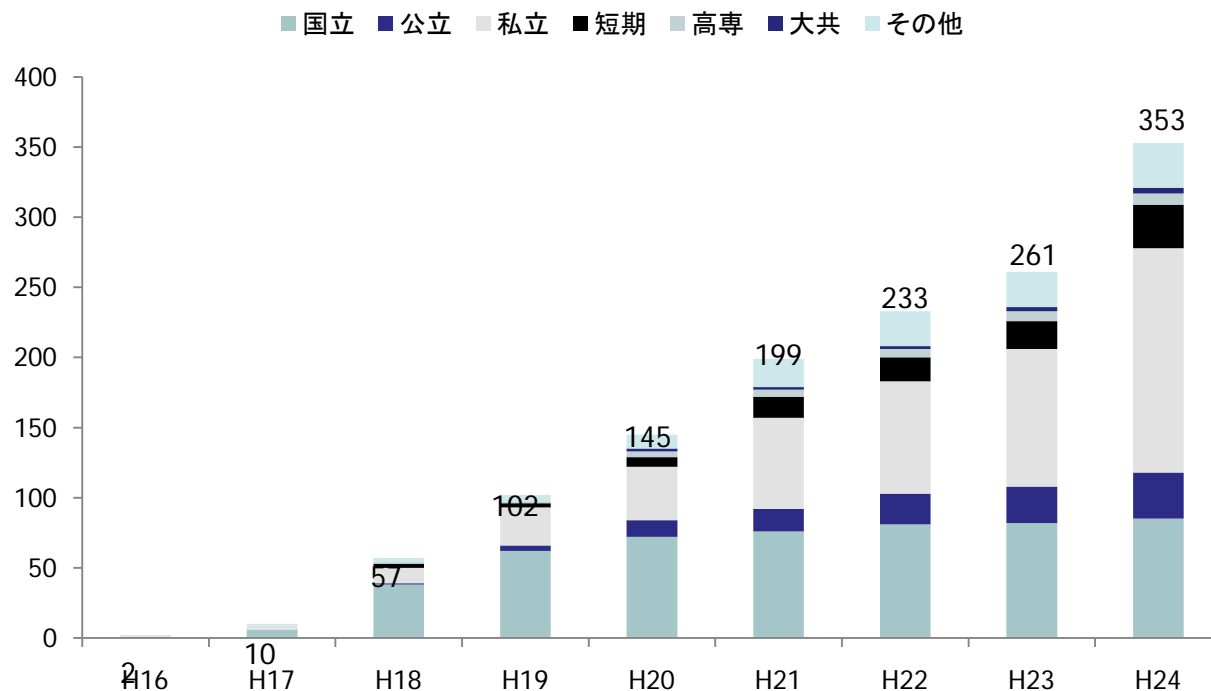
### ■ 大学の生産する情報の公開

#### 社会へのアカウンタビリティ

- ◆ 学位論文
- ◆ 紀要
- ◆ 研究に付随する電子データ

# NIIによる機関リポジトリ委託事業

- NIIは、3期8年（H17-H24）に渡り、大学図書館とともに機関リポジトリを育成するプロジェクトを実施してきた。この成果は多岐にわたるが、最も端的な成果としてはリポジトリの設置数の急激な増加を達成したことが挙げられる。
- また、平成25年度から博士論文のインターネット公開が義務づけられたことにより、機関リポジトリの役割がさらに重要になっている。



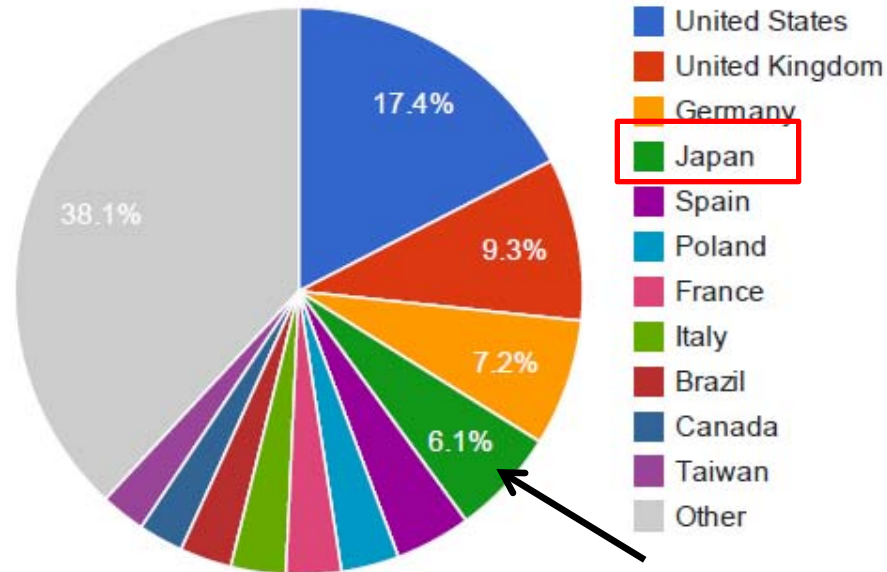
## 機関リポジトリ構築機関数

国立大学	85
公立大学	33
私立大学	160
その他	75
合計	353

（平成25年3月末時点）

# 機関リポジトリ構築数：世界の状況

Proportion of Repositories by Country - Worldwide



138機関

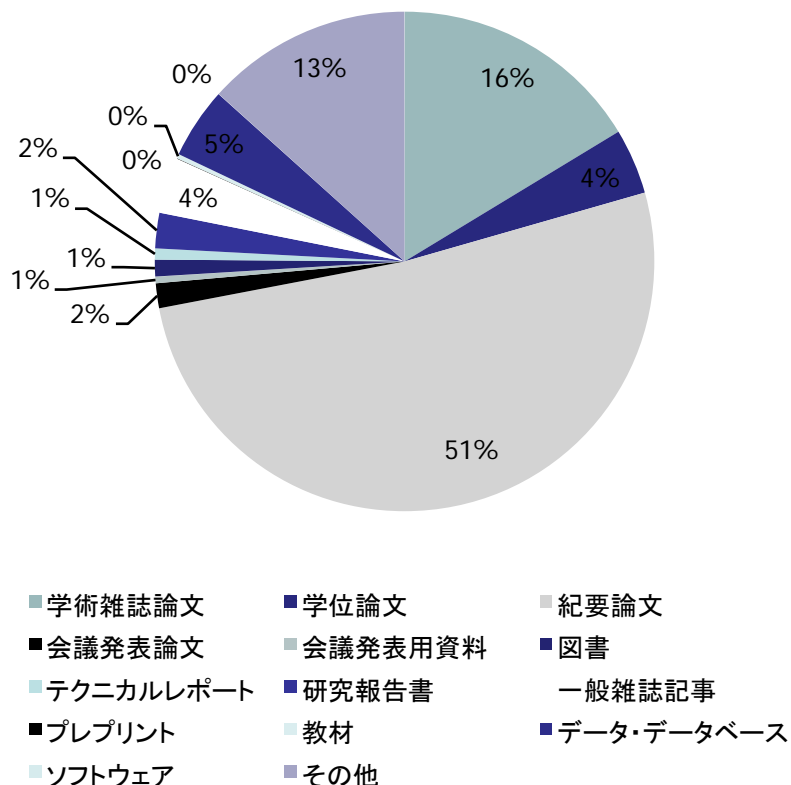
Total = 2269 repositories

OpenDOAR - 03-Apr-2013

実数では日本は世界第三位である。当サイトに未登録のため、順位が下がっている。

出典: Open DOAR: Directory of Open Access Repositories  
<http://www.opendoar.org/> (参照: 2013/4/3)

# 本文を含むコンテンツの登録状況

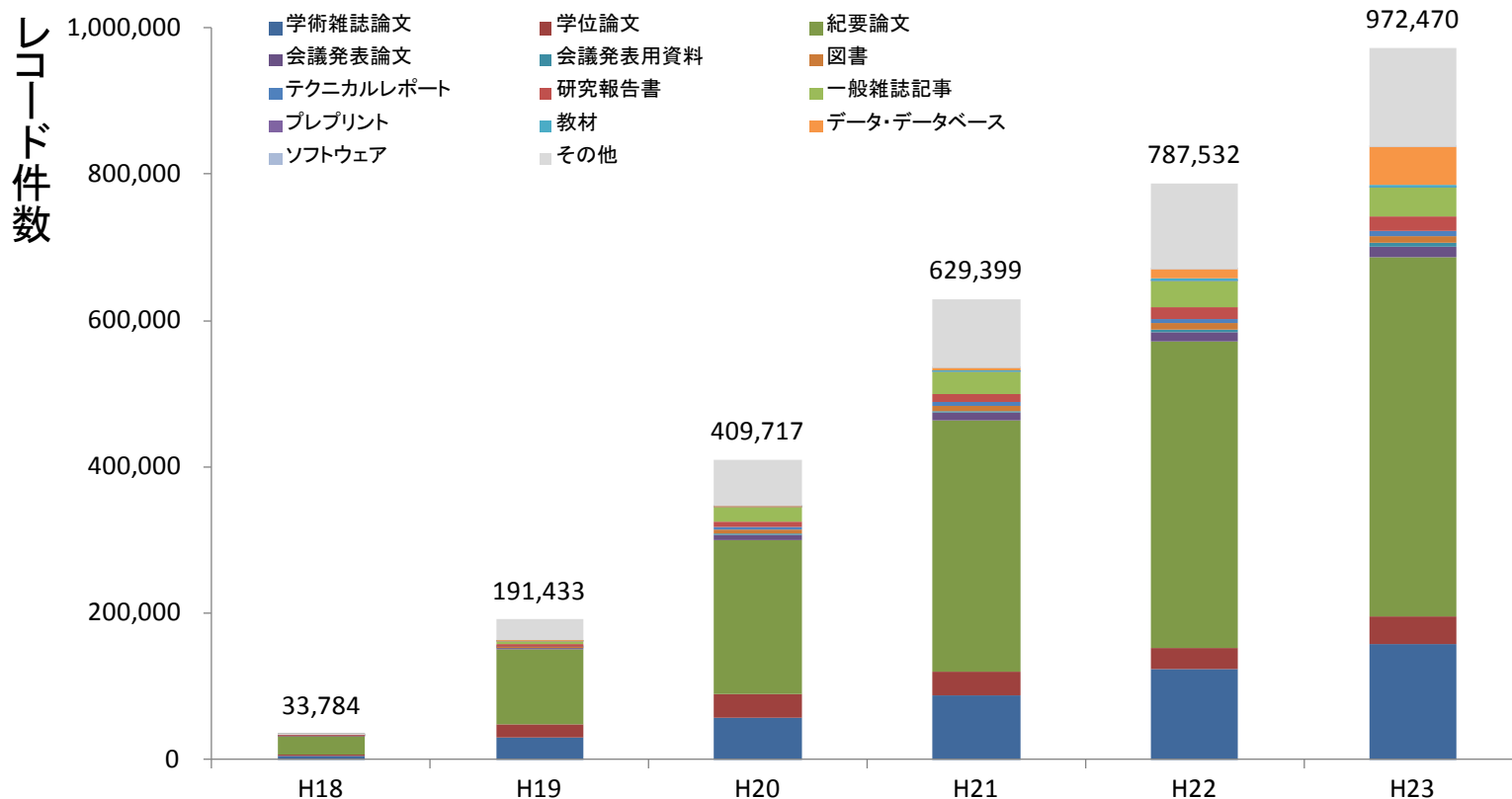


資料の種類別	コンテンツの数	割合
学術雑誌論文	185,397	16.3%
学位論文	48,371	4.3%
紀要論文	584,352	51.4%
会議発表論文	18,384	1.6%
会議発表用資料	4,774	0.4%
図書	12,394	1.1%
テクニカルレポート	8,220	0.7%
研究報告書	25,979	2.3%
一般雑誌記事	41,203	3.6%
プレプリント	336	0.0%
教材	3,160	0.3%
データ・データベース	52,412	4.6%
ソフトウェア	26	0.0%
その他	151,500	13.3%
合計	1,136,50	

IRDBコンテンツ分析システムによる <http://irdb.nii.ac.jp/analysis/index.php>

(2013年3月末時点) (参照:2012/06/03)

# 本文を含むコンテンツの登録数の年次変化



年度	学術雑誌論文	%	学位論文	%	紀要論文	%	他	%	計	%
H18	4,916	15%	667	2%	25,834	76%	2,367	7%	33,784	100%
H19	30,597	16%	16,628	9%	103,289	54%	40,919	21%	191,433	100%
H20	56,251	14%	32,307	8%	211,457	52%	109,702	27%	409,717	100%
H21	87,748	14%	32,588	5%	343,543	55%	165,520	26%	629,399	100%
H22	123,434	16%	28,303	4%	419,569	53%	216,226	27%	787,532	100%
H23	157,705	16%	38,017	4%	491,247	51%	285,501	29%	972,470	100%

※ 各年度3月末時点の数値

# 共用リポジトリ-JAIRO Cloud

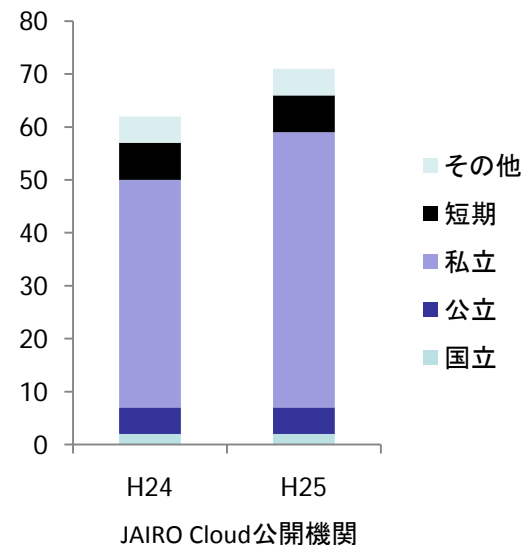
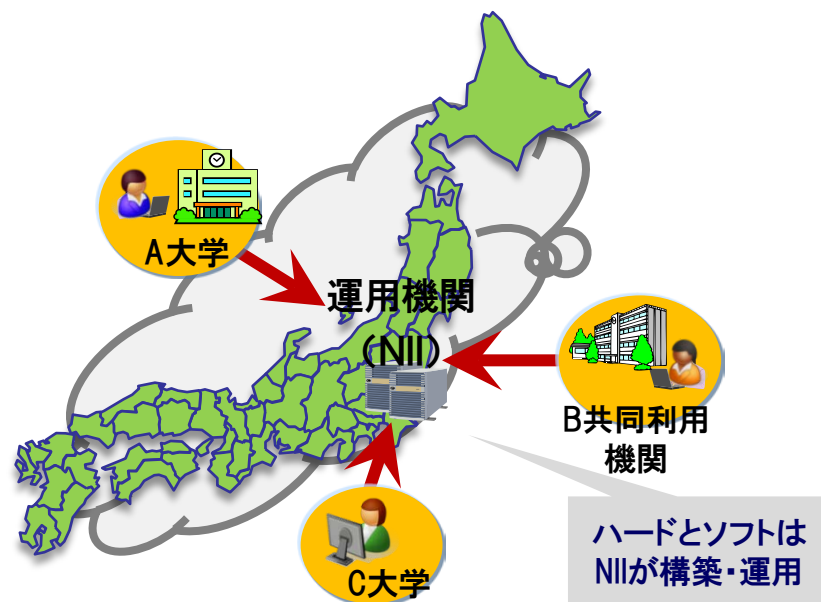
## □ 本事業のめざすもの

- 機関リポジトリの自力構築が困難な機関向けに、リポジトリのシステム環境を提供し、運用を支援する。学位授与機関数はおよそ400、平成27年度までにこの規模にリポジトリを拡大する。

NIIが開発した機関リポジトリのオープンソフトウェア「WEKO(ウエコ)」をベースに、クラウドとしてNIIがシステム環境を提供

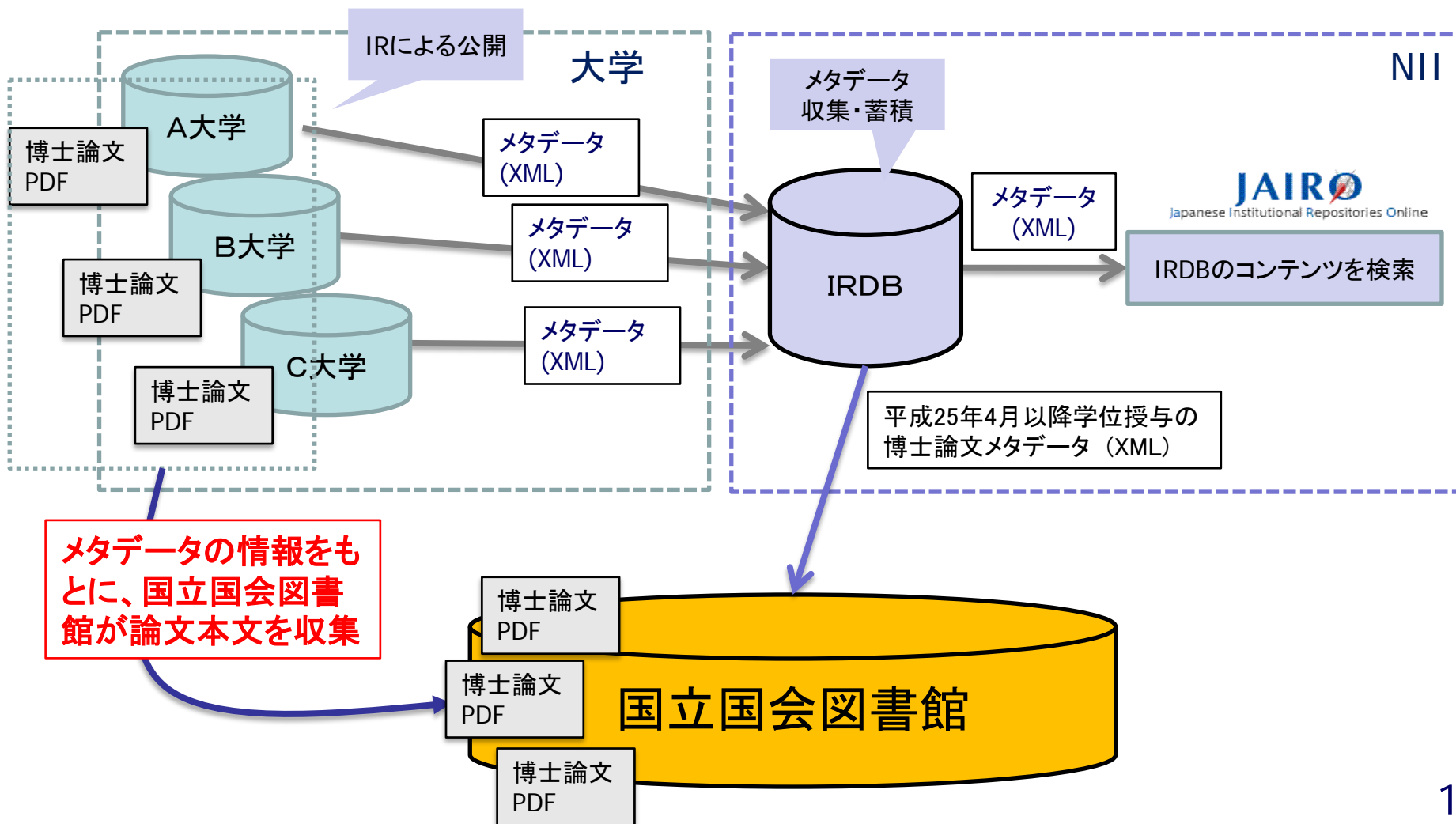
## □ サービス対象

- 日本国内の大学、短期大学、高等専門学校、大学共同利用機関のうち、当面は次の機関を対象とする。
  - 新たに機関リポジトリを構築する機関
  - 地域共同リポジトリを構築する機関
  - 既設リポジトリからの移行実験も実施中



# 博士論文のインターネットによる公表

- 学位規則改正により、各大学等は、平成25年4月以降に学位を授与した博士論文を「インターネットの利用により公表」することとなった。
- 大学、NDL、NIIは、連携・協力して博士論文の公表と蓄積を行う。





# 国際環境下でのオープンアクセスの推進

## □ 国際学術情報流通基盤整備事業(SPARC Japan)

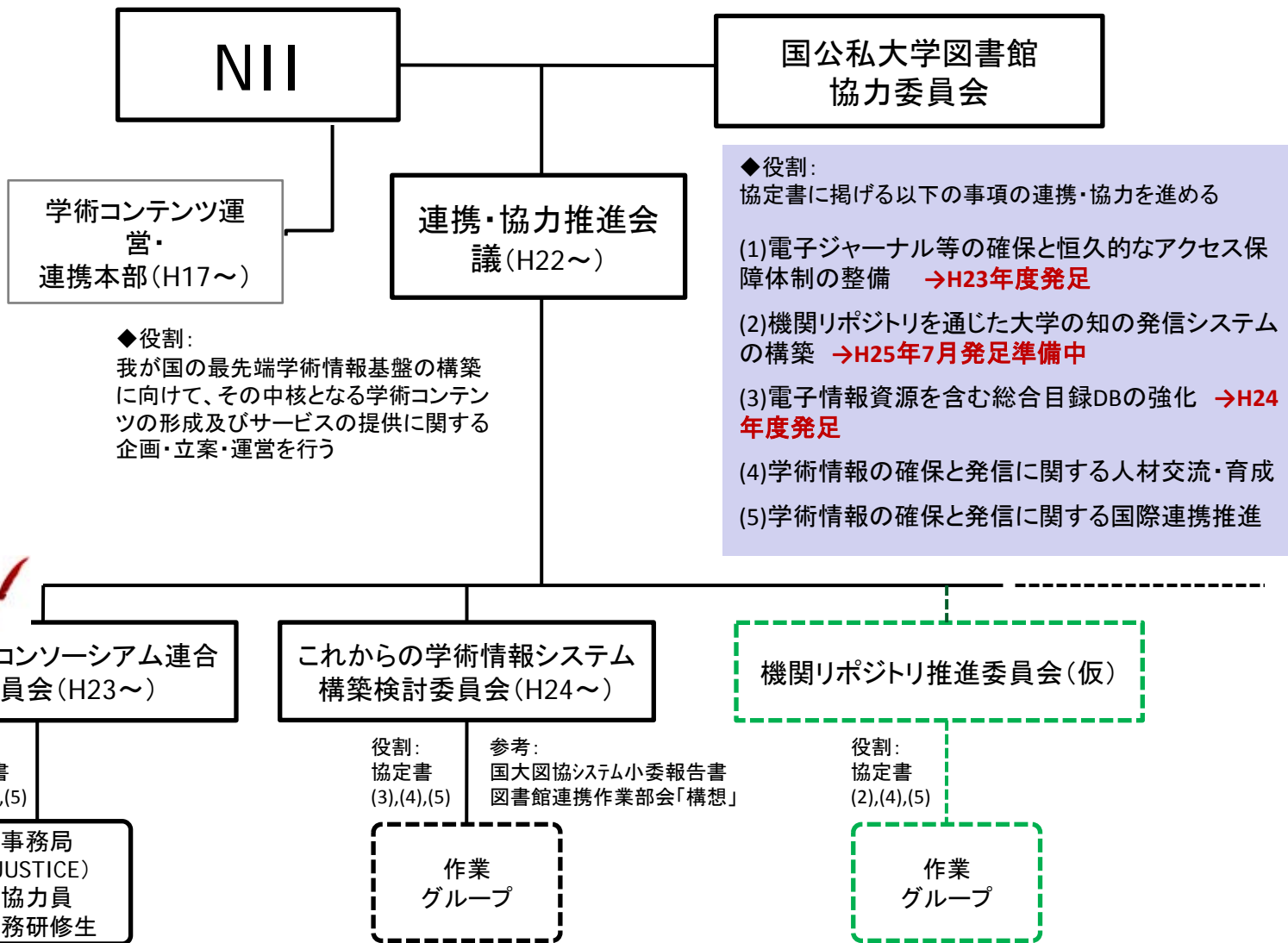
- 我が国の環境に合ったオープンアクセスの在り方を検討
  - 機関リポジトリ事業との連携
  - セミナー, ニュースレターに代わるウェブフライヤー, 年報によるAdvocacy活動
- 国際連携の強化(国際的なオープンアクセスプロジェクトへの参画)
  - SCOAP<sup>3</sup> 支援
    - 大学図書館とのタスクフォースを組み, 国内参加館とりまとめ
    - Reconciliation, MoU締結予定 → 2014年1月 OAサービス開始予定
  - arXiv.org 支援
    - 大学図書館, 関連機関の協力により, arXiv.org支援機関の窓口
    - 2013年より, arXiv会員制プログラムへ移行
- OA支援のためのパイロットプロジェクト実施に係る検討
  - APC機関負担モデル調査

## □ ジャパンリンクセンター(JaLC)運営への参画

- JST、NDL、NIMSとの共同運営
- 機関リポジトリ内のコンテンツへのDOI付与促進→OA流通推進

# 大学図書館との連携活動体制

# 大学との連携体制



# 東日本大震災アーカイブ事業 に関する連携状況

## ◆ 「東日本大震災アーカイブ」基盤構築プロジェクト ラウンドテーブルに構成員として参加

### 主な検討事項

- ・東日本大震災アーカイブのシステム構築に係る技術的要件
- ・東日本大震災アーカイブの利活用法
- ・デジタルアーカイブの構築・連携・運用に関する技術・ルール等

### 参加者

- |             |               |
|-------------|---------------|
| ラウンドテーブル    | : 高野 明彦教授(座長) |
| 技術ワーキンググループ | : 高野 明彦教授(座長) |