

Cyber Science Infrastructure in Japan

December 18, 2006

Masao SAKAUCHI

National Institute of Informatics, Japan

Today's Story

- ❑ e-Science and Data-centric Science
- ❑ Japan's Cyber-Infrastructure, aka "CSI"
- ❑ CSI Architecture
- ❑ Networking, Grid, ...
- ❑ Information Explosion Era
- ❑ Japan's Strategy for Scholarly Information
- ❑ NII's Scholarly Information Services
- ❑ Institutional Repositories

Approaches in Sciences

- ❑ Theoretical Science
- ❑ Experimental Science
- ❑ Computational Science
- ❑ Data-centric Science
 - Large amount of experimental data produced by sensors, instruments, ...
 - Large amount of simulation data produced by supercomputers

The Evolution of Science

Courtesy of Jim Gray

□ Observational Science

- Scientist gathers data by direct observation
- Scientist analyzes data

□ Analytical Science

- Scientist builds analytical model
- Makes predictions.

□ Computational Science

- Simulate analytical model
- Validate model and makes predictions

□ Data Exploration Science

Data captured by instruments
Or data generated by simulator

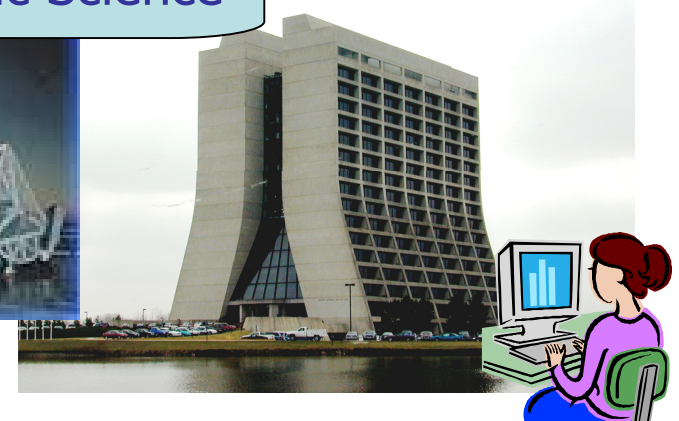
- Processed by software
- Placed in a database / files
- Scientist analyzes database / files

Theoretical Science
Experimental Science

Computational Science

Data-centric Science

e-Science



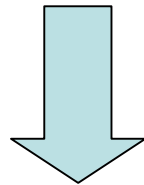
Cyber Science Infrastructure: background (1)

- A new information infrastructure is needed for boosting today's advanced scientific research.
 - **Integrated information resources and system**
 - super computer and high-performance computing
 - software,
 - databases and digital contents such as e-journals
 - “Human” and research processes themselves
 - **U.S.A: Shared Cyber-Infrastructure (SCI)**
 - **Europe: e-Infrastructure (EGEE)**

Cyber Science Infrastructure: background (2)

□ Break-through required in information research

- the key to industry/academia cooperation:
from 'Science' to 'Intellectual production'

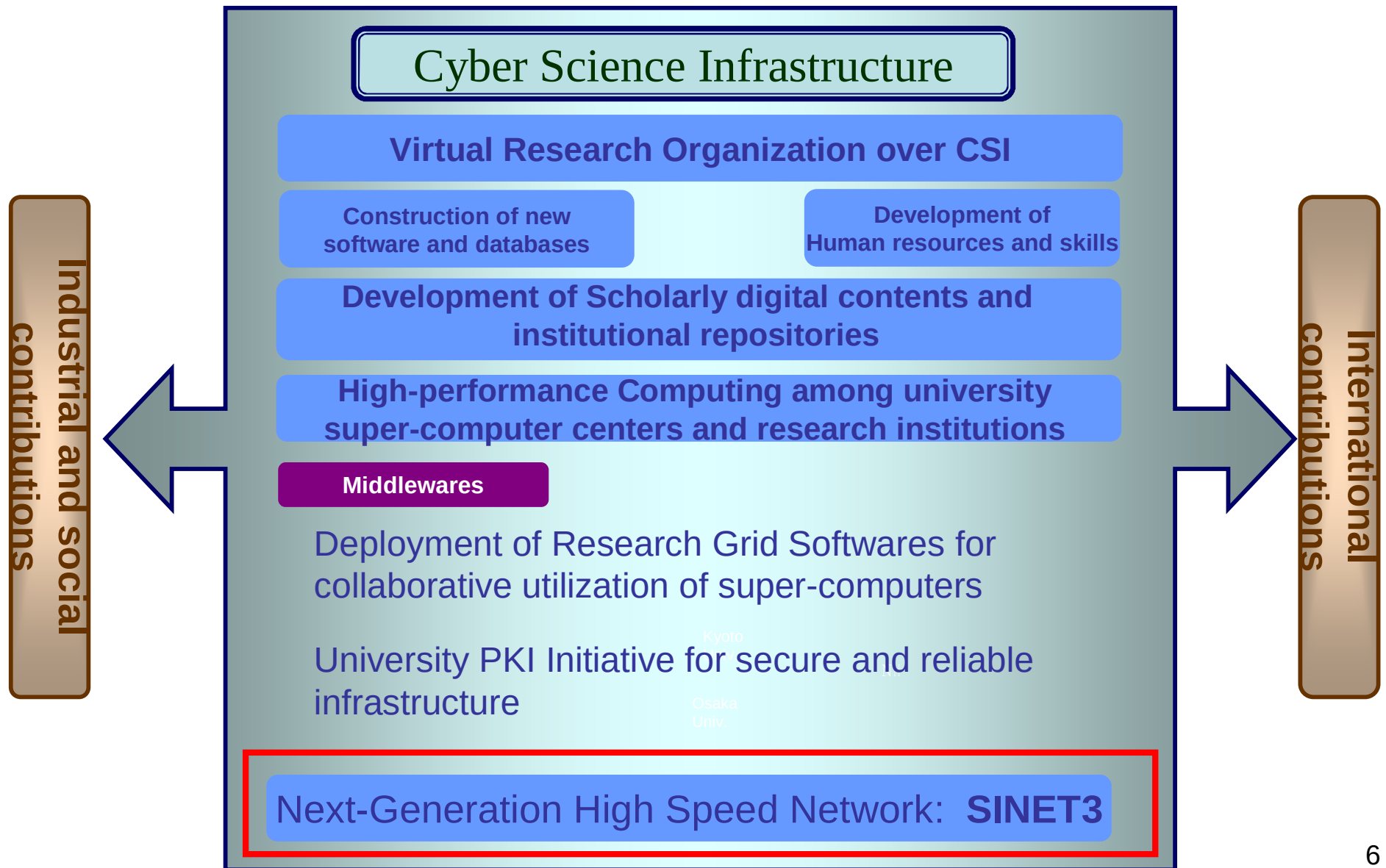


Advanced information infrastructure for research will be the key in international cooperation and competition in future science and industrial fields

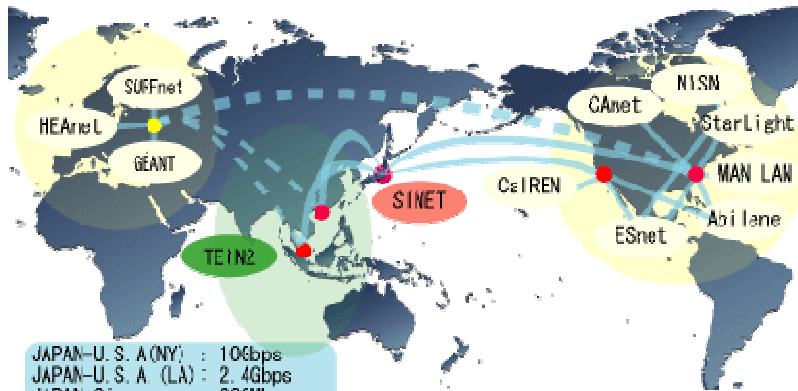
A new comprehensive framework of information infrastructure in Japan

Cyber Science Infrastructure CSI

CSI Architecture



Japanese Academic Network: SINET/SuperSINET→SINET3



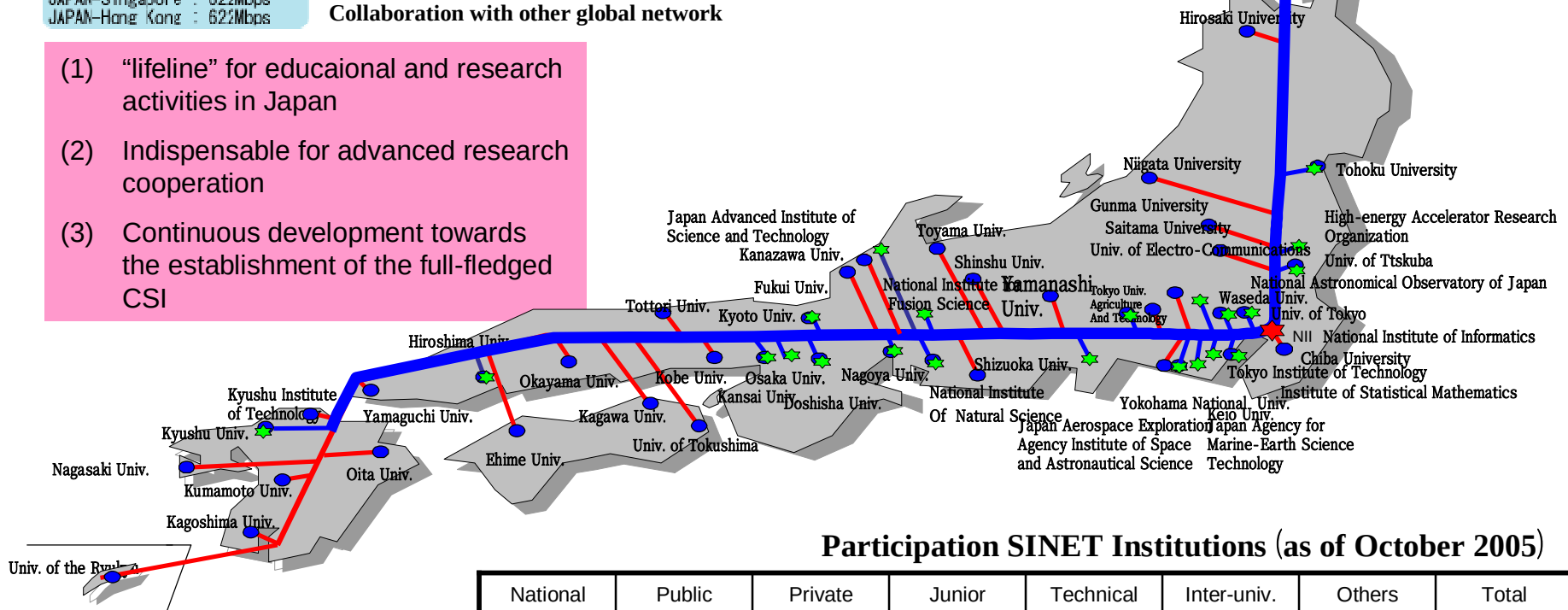
Collaboration with other global network

Circuit Speed

SINET(44 nodes)	1 Gbps
Super SINET (35 nodes)	10 Gbps
International Line	U.S.A: 10 Gbps×1 2.4 Gbps×1 Singapore: 622 Mbps×1 Hong Kong: 622 Mbps×1

40Gbps
From April 2007

- (1) “lifeline” for educational and research activities in Japan
- (2) Indispensable for advanced research cooperation
- (3) Continuous development towards the establishment of the full-fledged CSI



Participation SINET Institutions (as of October 2005)

National Univ.	Public Univ.	Private Univ.	Junior College	Technical College	Inter-univ. institute	Others	Total
82	48	266	78	42	14	184	713

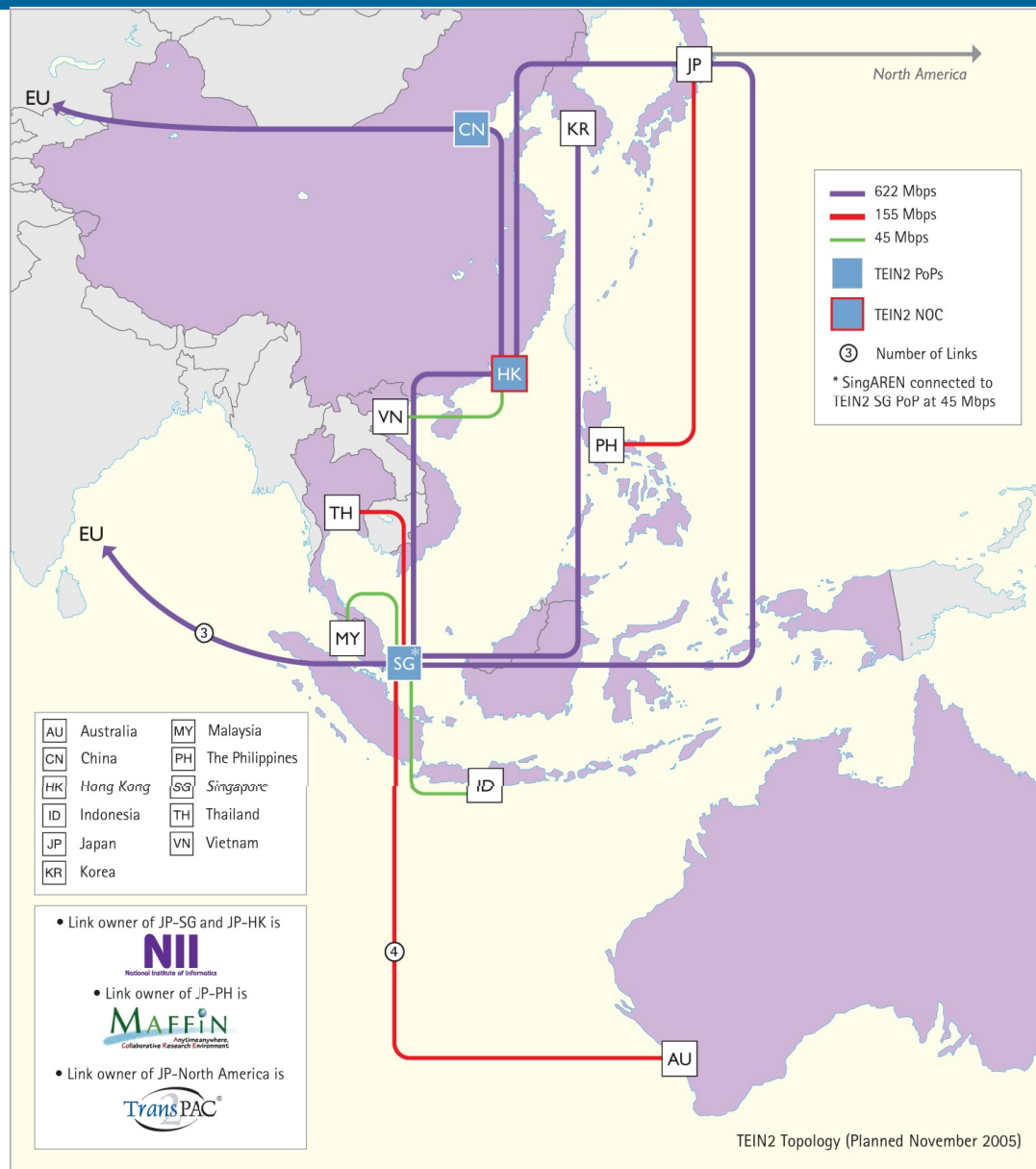
Features of Next-generation Network

- ❑ Maximum 10Gbps Ethernet connection for user interface
- ❑ Providing multiple-layered services: IP, Ethernet, Layer-1 (dedicated line) (L3/L2/L1).
- ❑ Virtual network connecting multiple points
- ❑ High quality multicast/multi-point connection services
- ❑ IPv6 services
- ❑ Various information and statistics data of the network

Deployment will start in April, 2007

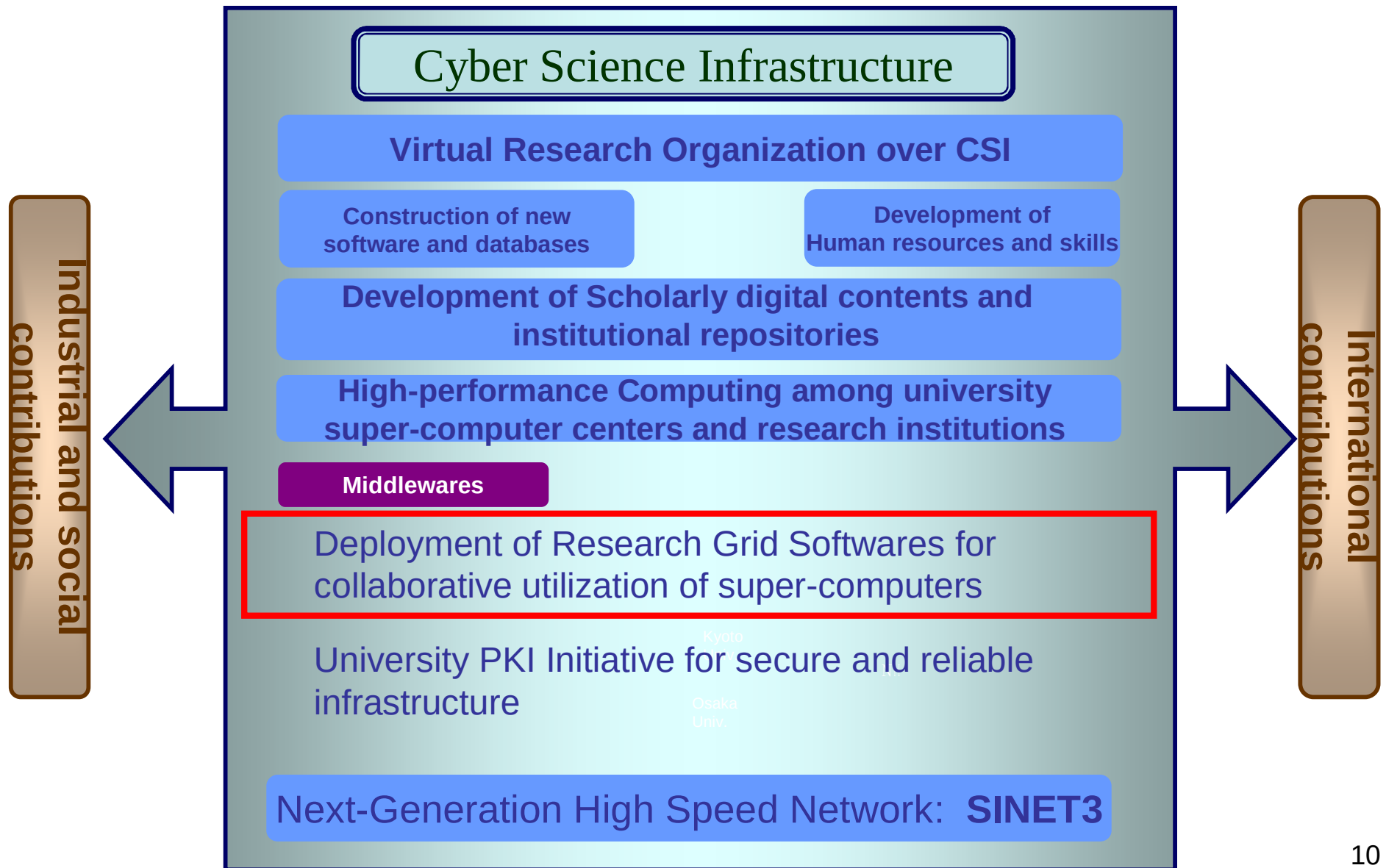
Academic Networking of SINET (Asia) and Partners

TEIN2

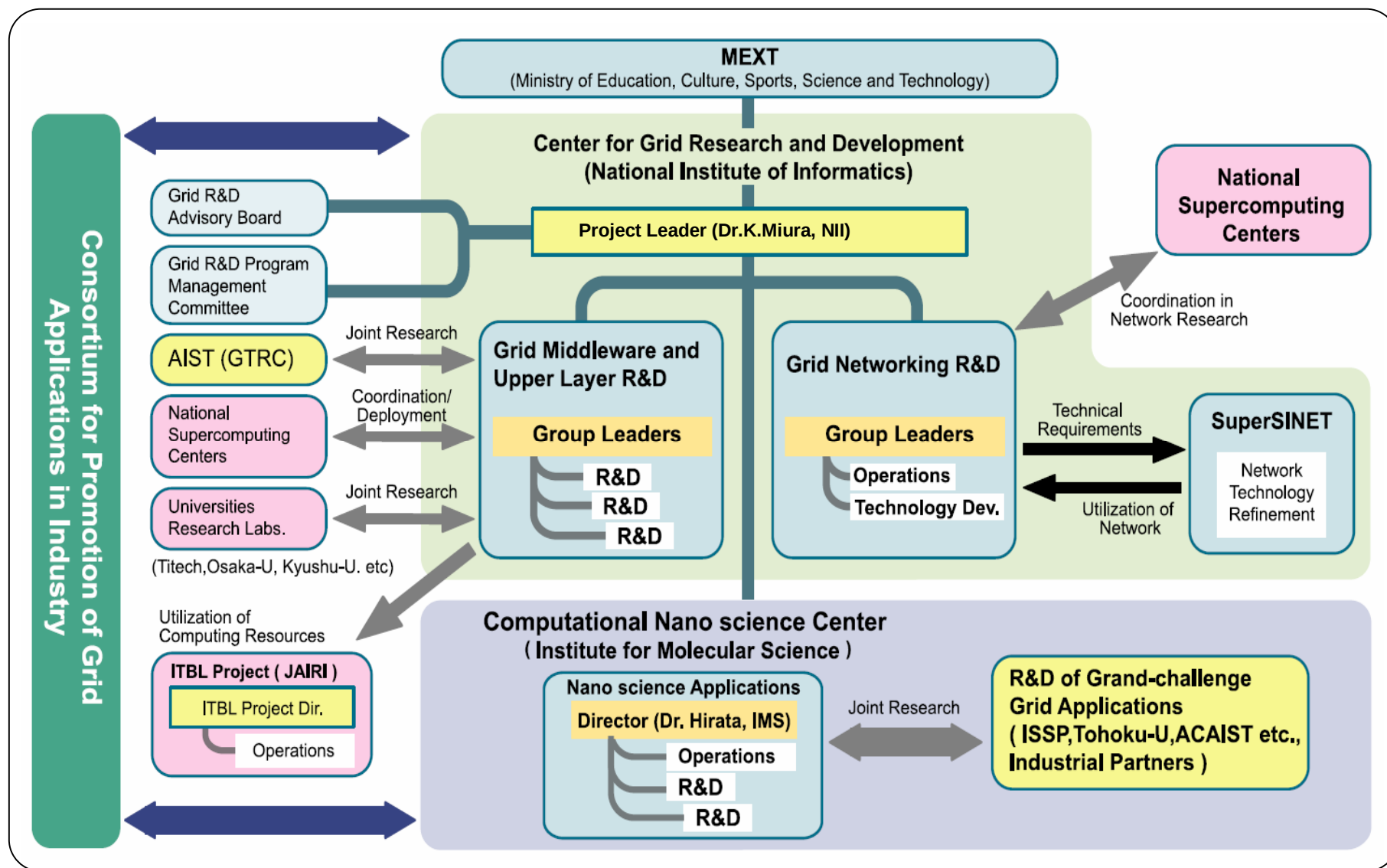


NII
National Institute of Informatics

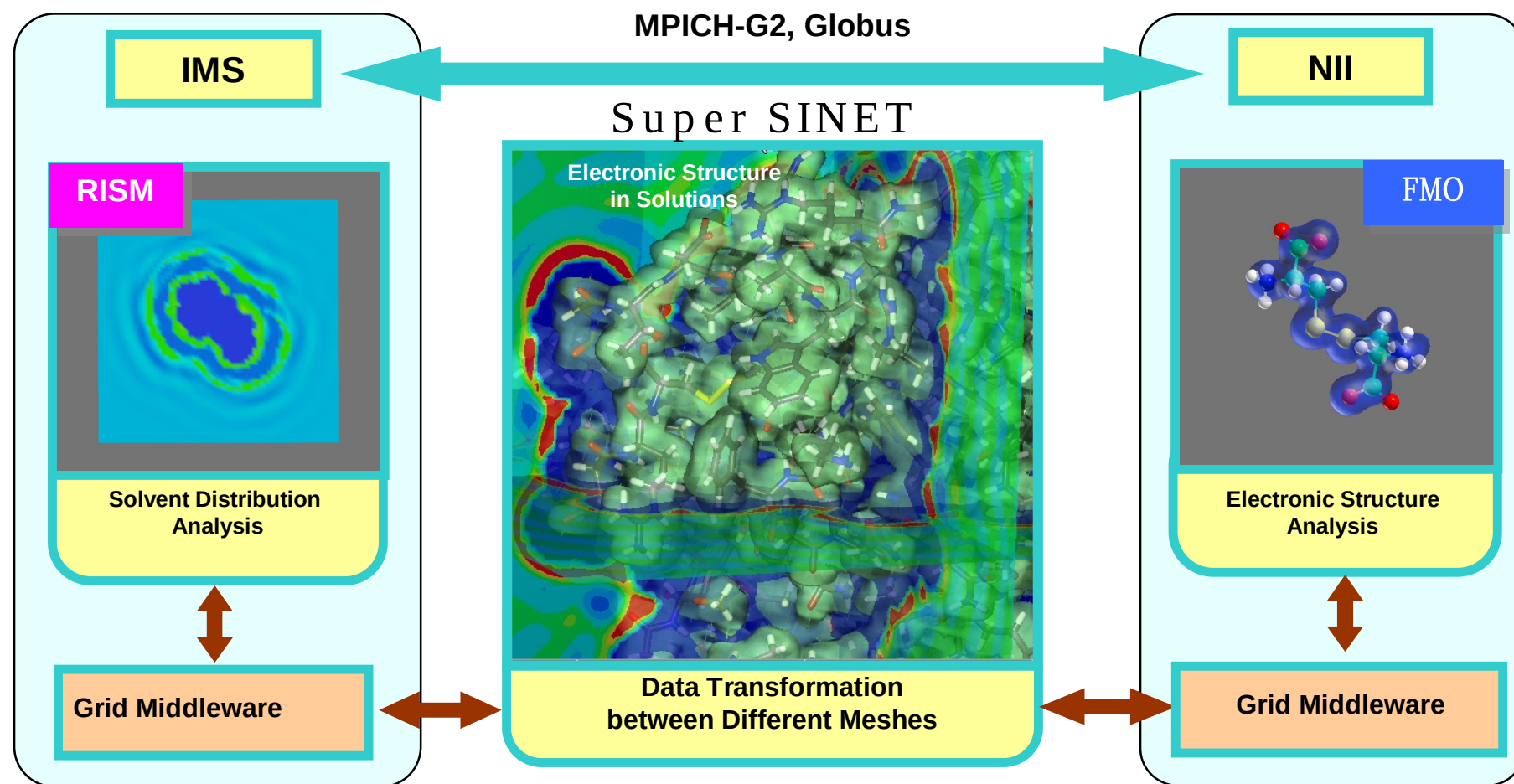
CSI Architecture



NAREGI: National Research Grid Initiative



Adaptation of Nano-science Applications to Grid Environment



RISM

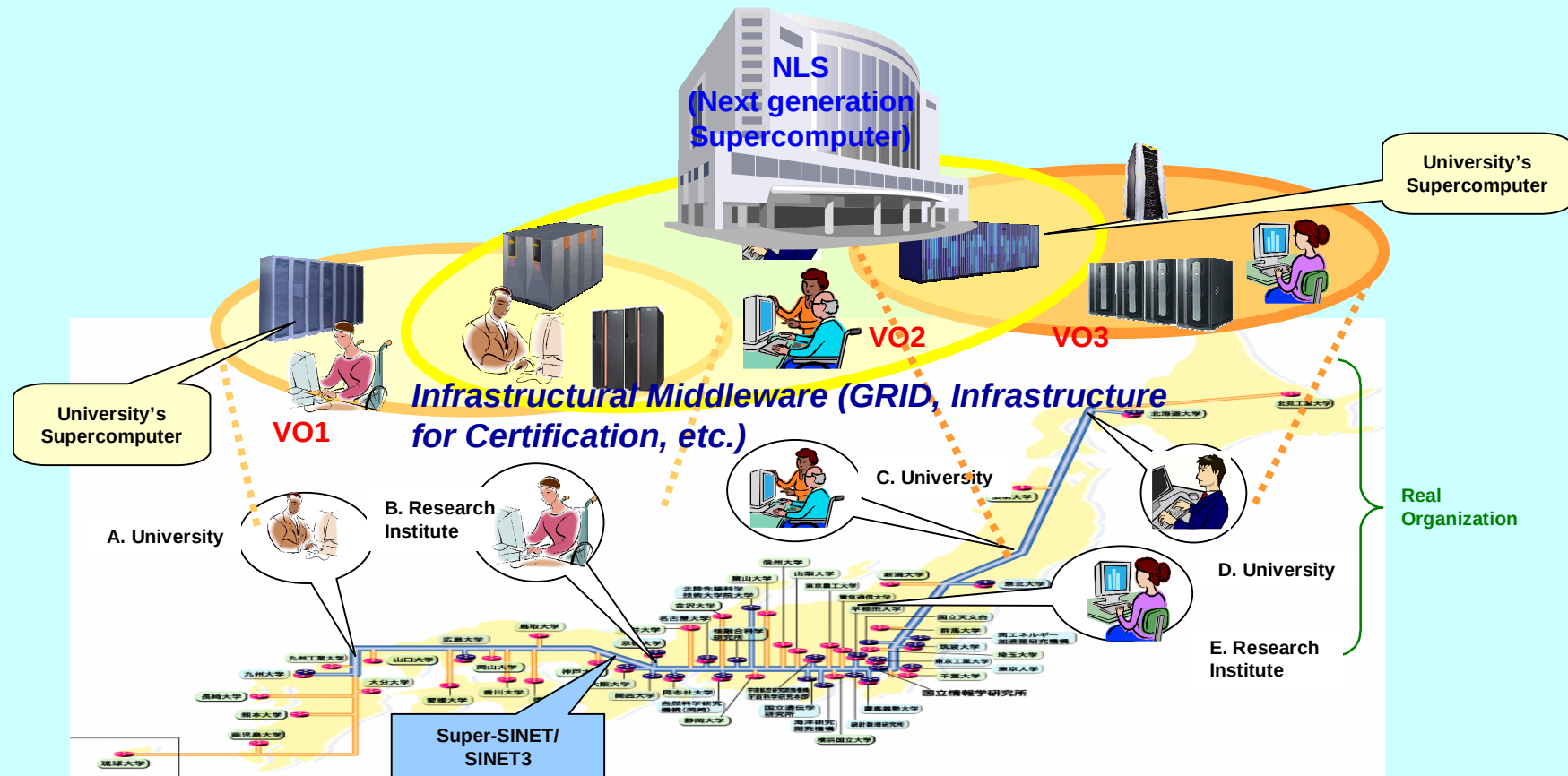
Reference Interaction Site Model

FMO

Fragment Molecular Orbital method

VOs (Virtual Organizations) on NAREGI

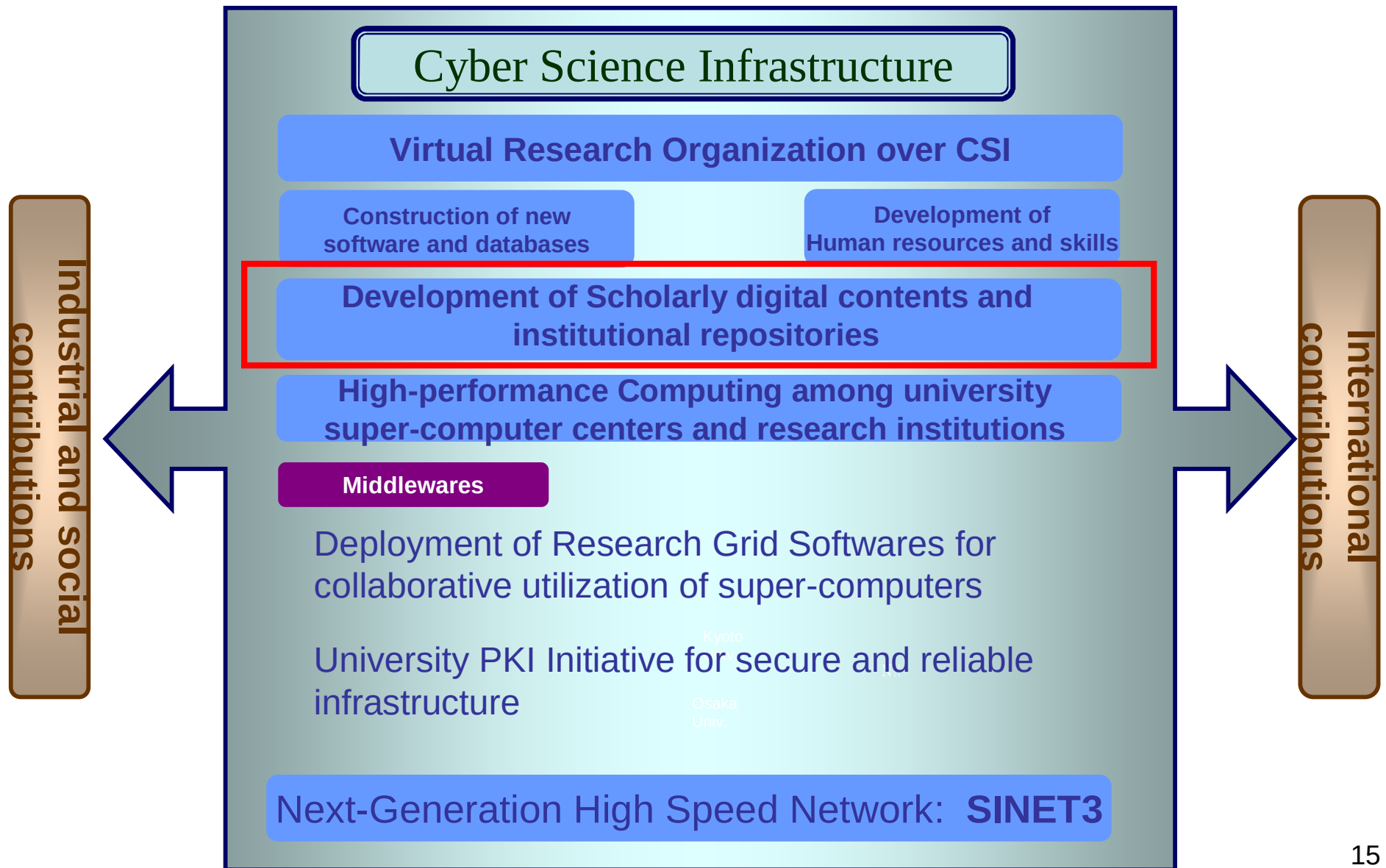
Multilayered Virtual Organizations



Sample e-Science Research Activities on CSI

- ❑ Widely distributed data analysis system
 - High-energy Physics Institute and universities
- ❑ Multi-scale simulation system
 - Utilization of the “Earth Simulator” super-computer
- ❑ Large-scale design simulation system
 - For nano-technology researches
- ❑ Widely distributed database system
 - Genetics database sharing
- ❑ Remote analysis of observation data
 - VLBI in astronomy

CSI Architecture



Digital Scholarly Information and NII

Aiming at:

Stable dissemination, preservation and provision of information produced by academic communities

Precise and timely service of information that academic communities require.



Academic communities

(researchers , universities, academic societies, etc)



supporting

Cyber Science Infrastructure: C S I

Networking + Digital contents + Research Activities

Scholarly digital contents are common property among universities and NII

NII-ELS

- E-journals of Japanese academic societies and university bulletin
- Retrospective digitization as an e-archive
- 2,900 titles including 2.8 million articles

NII-REO

- Archive server for university library consortium
- “last resort” of scholarly e-journals
- 1,500 titles including 3.4 million articles, such as Springer and OUP

Institutional Repositories

- Dissemination of university information and research results
- Linkage with REO and ELS
- Linkage to open access journals and green journals

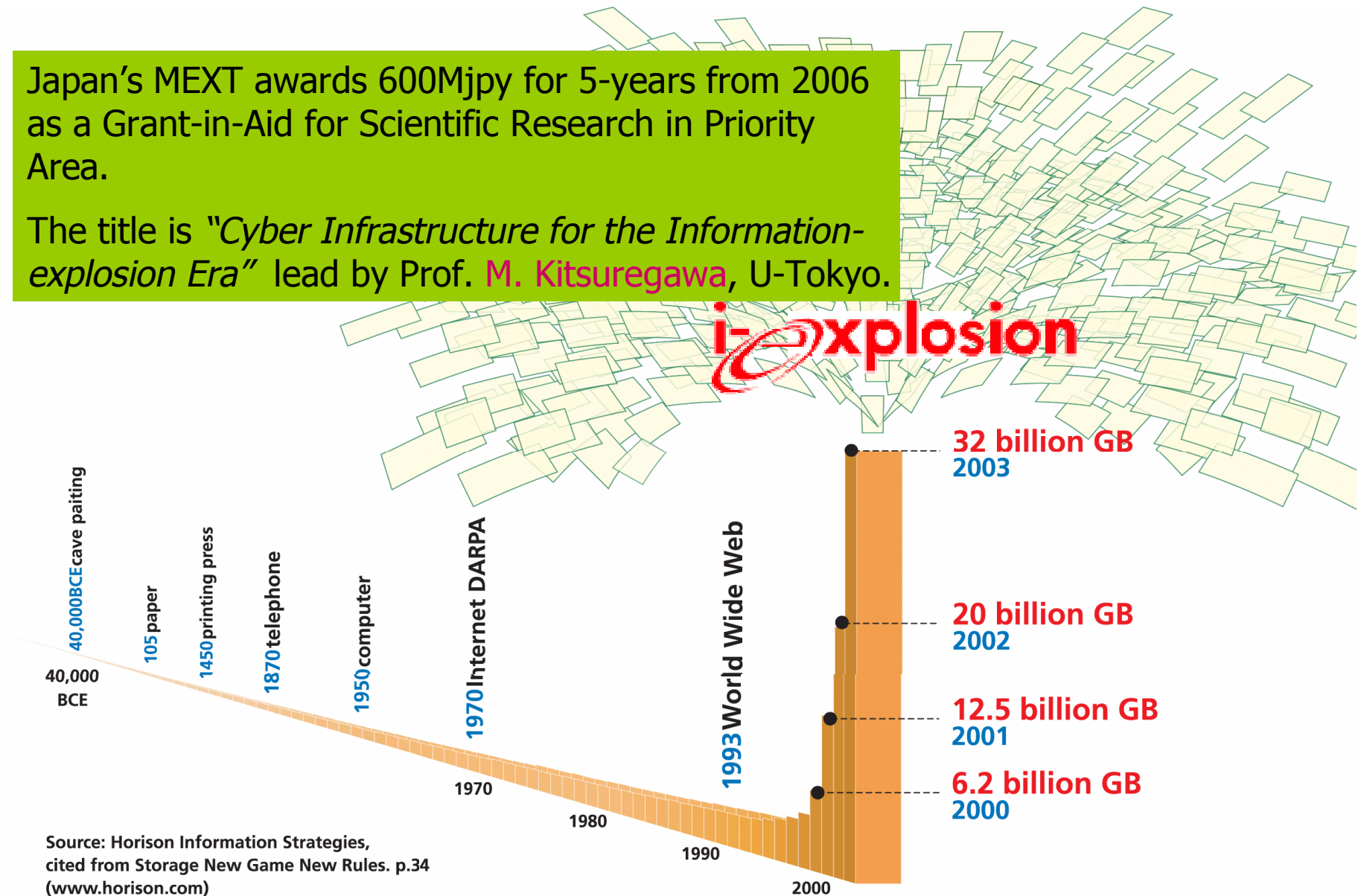
NII

NII + Universities

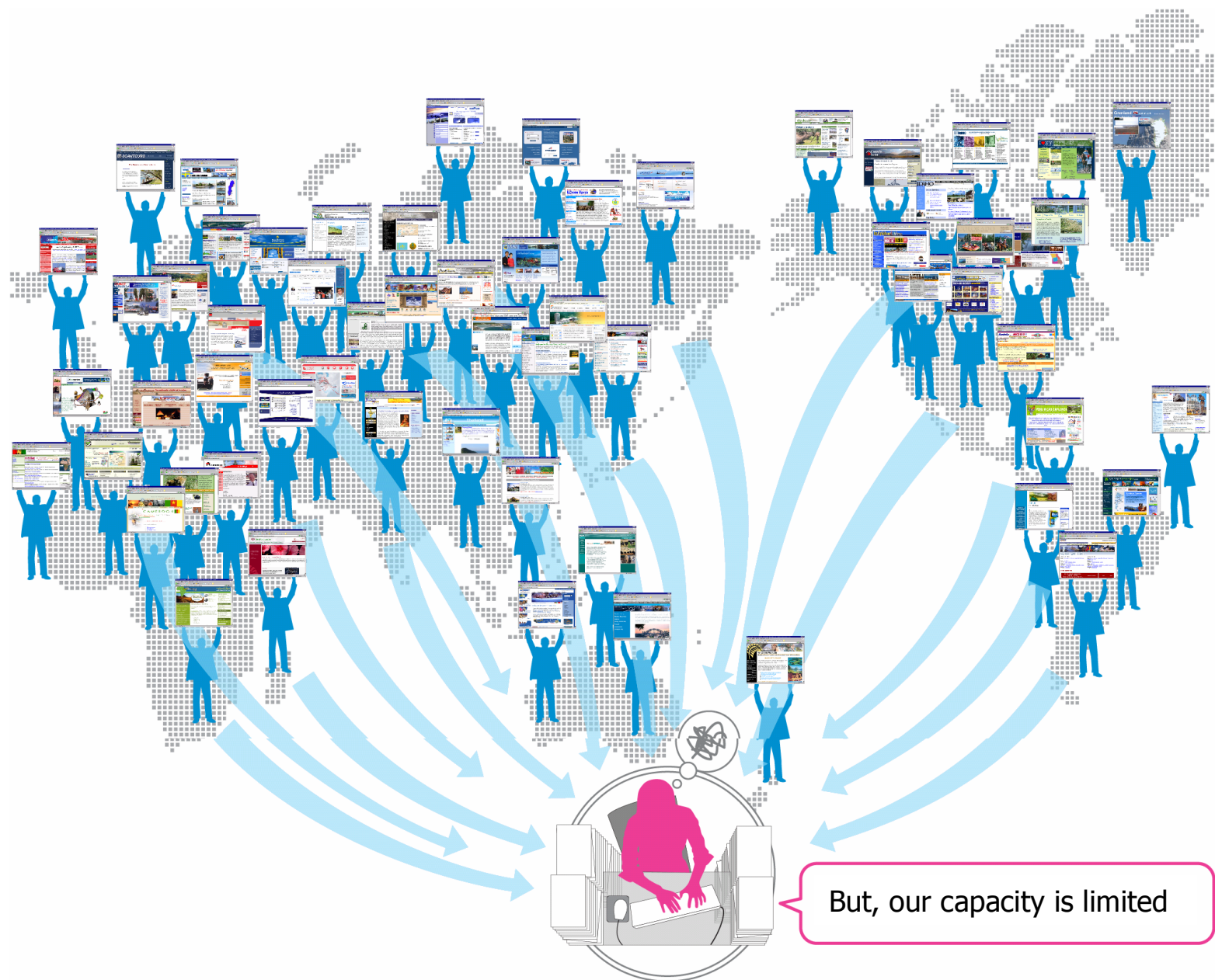
The Beginning of “Information Explosion” Era

Japan's MEXT awards 600Mjpy for 5-years from 2006 as a Grant-in-Aid for Scientific Research in Priority Area.

The title is “*Cyber Infrastructure for the Information-explosion Era*” lead by Prof. **M. Kitsuregawa**, U-Tokyo.

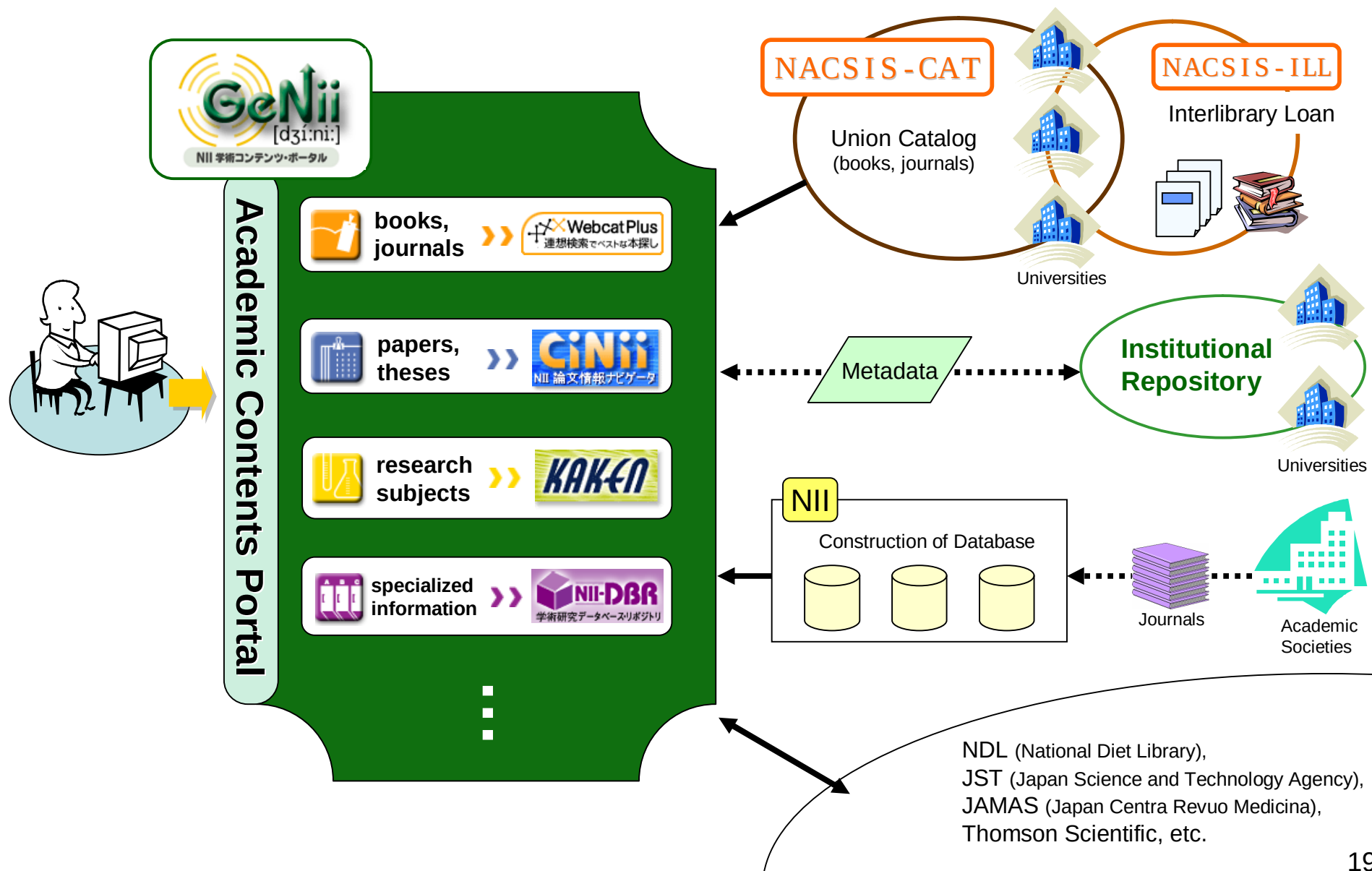


Millions of Humans Disseminate Information



But, our capacity is limited

GeNii: Scholarly Information Portal



Scholarly Information – Books & Journals Search Service

Webcat Plus

Webcat Plus enables you to;

- ◆Narrow search to more relevant info.
- ◆Search books from various areas
- ◆Find unexpected good books



How can I get that book?

Search by
Associative Search



Search by
keyword or
phrase in
theme

Matching Search



Search by title
or Author

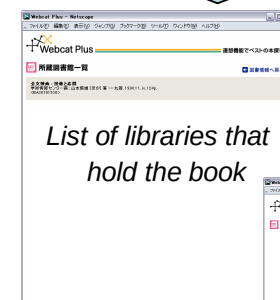
Search by
title, author,
year or ISBN



What is the
content of
the book?



List of libraries that
hold the book



Library info
·Address,
URL
·Terms of use

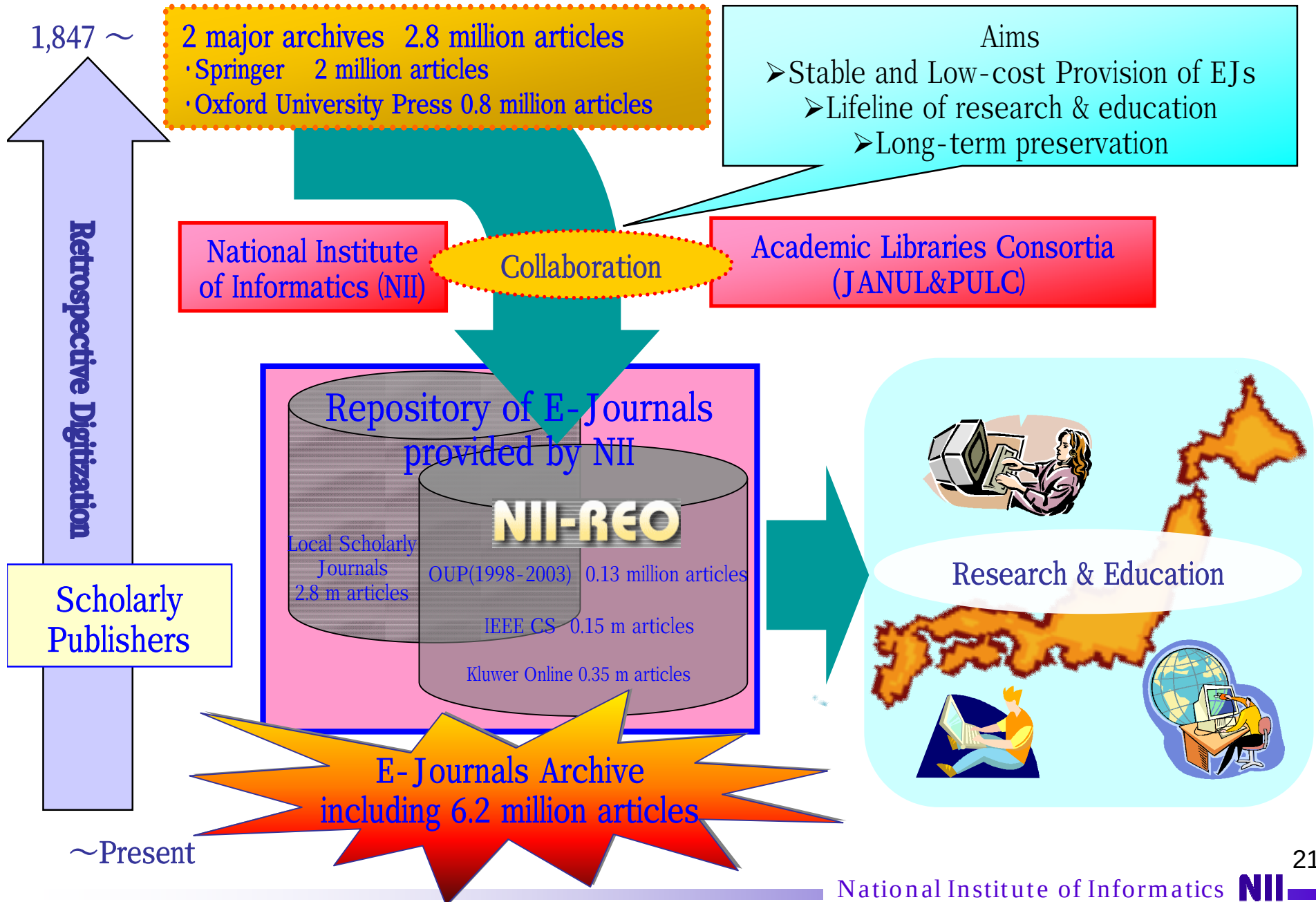
Which library
has the book?

Registered data: about 12,000,000 (as of Nov. 2006)
Page views: about 60million / year (FY 2005)



<http://webcatplus.nii.ac.jp/>

NII-REO: Repository of E-Journals



E-Journals Provided Online by NII

Collections	articles	description
Springer Online Journal Archive	2M	Springer 1847-1996, 800 titles
Oxford University Press	0.8M	OUP 1849-1995, 140 titles
Kluwer Online	0.35M	Former Kluwer 1996-2005, 500 titles
IEEE Computer Society	0.15M	IEEE Computer Society 2003 to current, 25 titles
Oxford University Press 1998-2003	0.13M	OUP 1998-2003, 150 titles
NII-ELS	2.8M	Japanese scholarly articles (society journal and university bulletins, 2,900 titles
total	6.23M	

Collaborative Acquisition of E-Journal Archives

□ In 2005, Japanese university libraries and NII collaboratively purchased the e-journal archives of:

➤ Springer Online Journal Archive

- about 800 titles (1846-1996)
- About 2M articles (16M pages)
- Open for Japanese universities since April, 2006
- LNCS included

➤ Full Online Collection of Oxford Univ. Press

- 137 titles (1849-1995)
- About 0.8M articles (4M pages)

Dissemination of Academic digital contents

CiNii (NII Article Information Navigator)

Large scale database
for articles

[19,000 journals; 10 million papers]

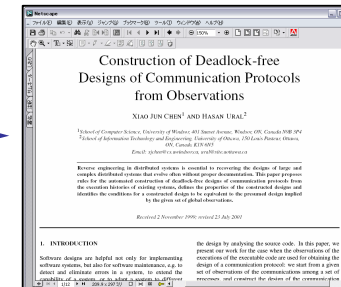
- ◆ Japanese Academic Journals
- ◆ Research Bulletins
- ◆ Japanese Periodicals Index of NDL

Portal site of scientific article
information

<http://ci.nii.ac.jp/>

Full texts (PDF)

[2,900 journals; 2.8 million papers]

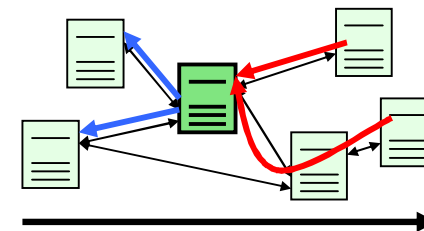


Links to

Collected information
from books and journals

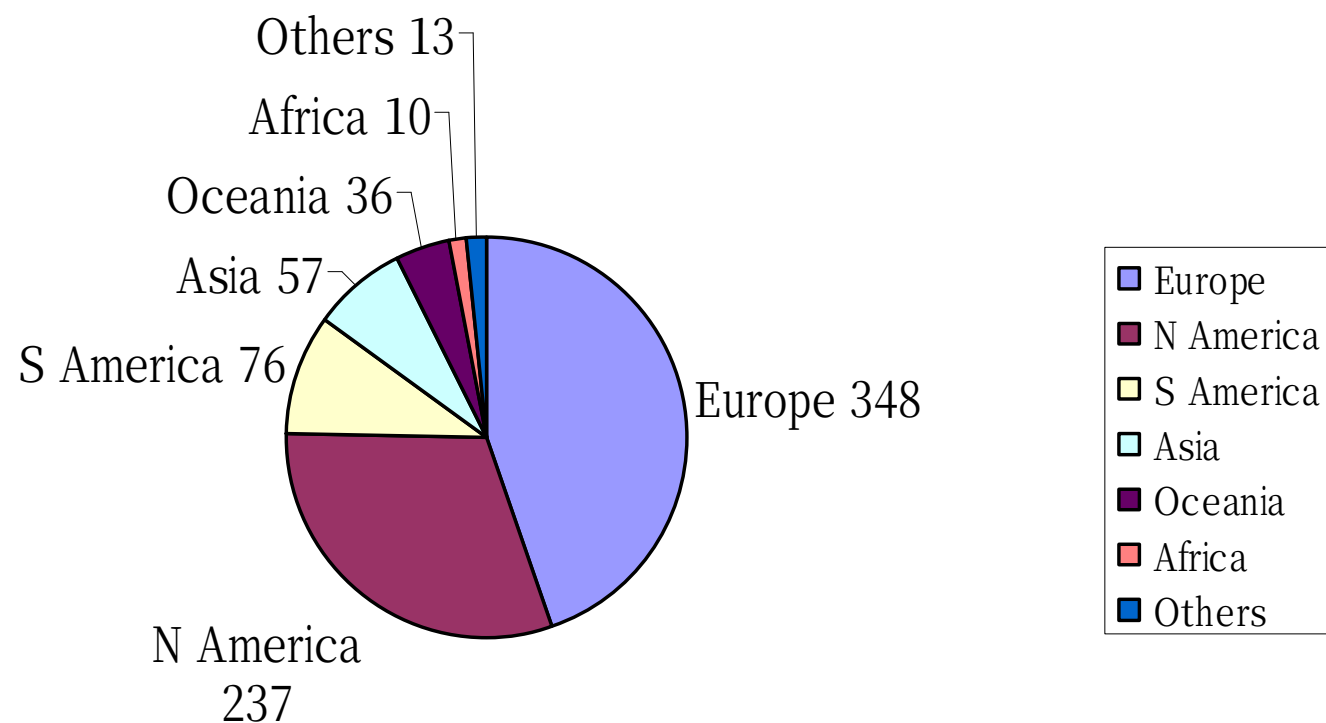
Citation links

[1,700 journals; 1.2 million papers]



Indication of references cited

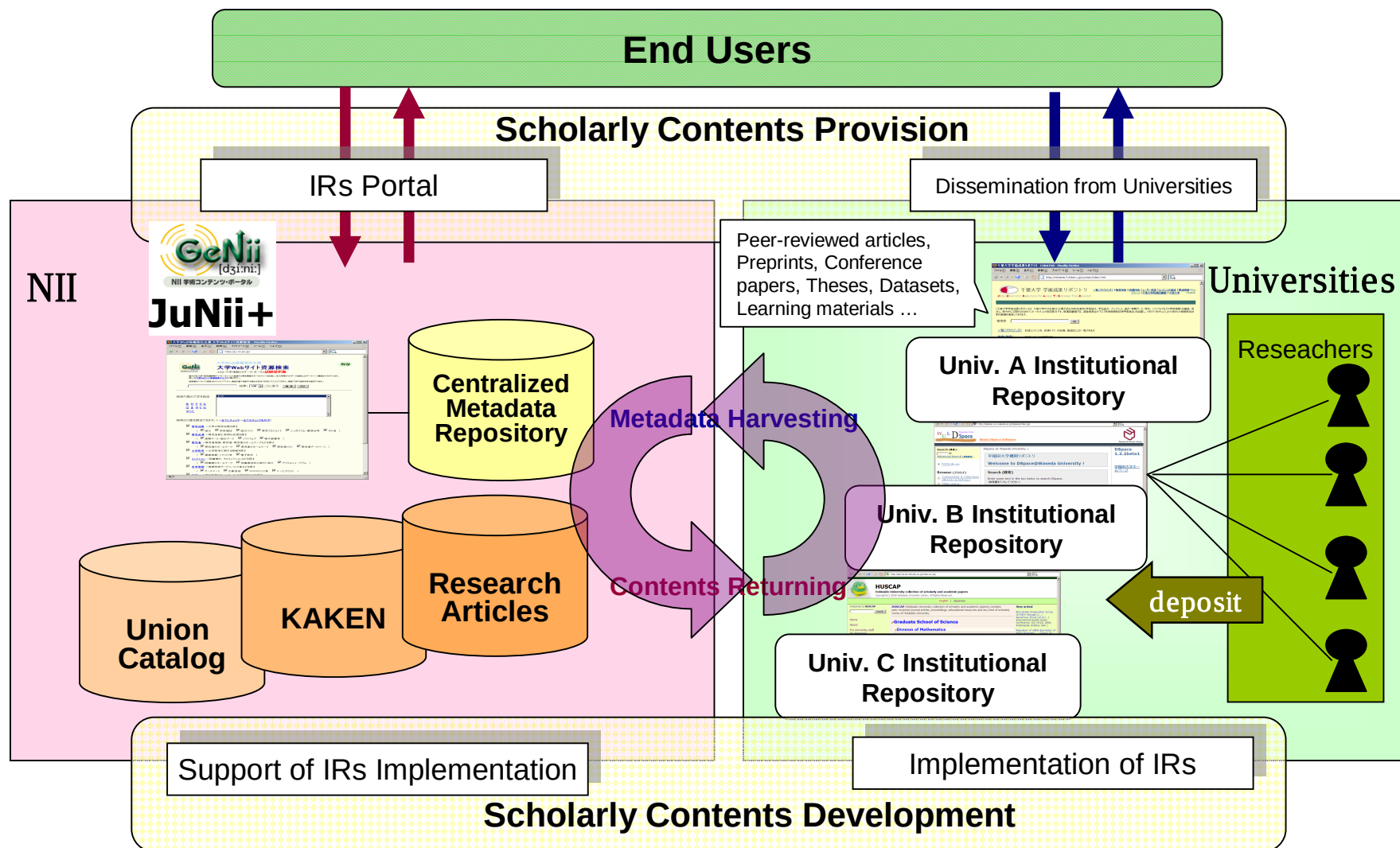
Institutional Repositories World-wide (777)



Registry of Open Access Repositories (2006.11.25)

<http://archives.eprints.org/>

Institutional Repository Collaboration

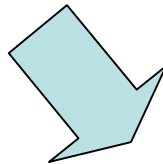


Institutional Repositories in CSI

[What is an **Institutional Repository** ?]

Accumulation and dissemination of electronic intellectual properties that academic institutions produce

- Scholarly / preprinted papers, theses
- Reports on Grant-in-Aid Scientific Research, technical papers
- Presentation slides used at an academic conferences, electronic materials for students
- Scientific databases
- And anything digital,....



NII started a collaborative project for Institutional Repositories with 19 major universities in 2005, then expanded the participating universities upto 57 in 2006.

[Roles of NII]

Support for collaboration of Institutional Repositories

- Automatic harvesting of information in Institutional Repository
- Storage in NII's meta-database
- Provision of integrated retrieval service

Support for creation of Institutional Repositories

- Project with university libraries

Ongoing Projects

□ Integrated Scientific Databases

- Earth Science
- BioScience

□ New Collaborative Project on e-Science

- with 20 Research Institutions
- 2006~

□ IMAGINE

- Innovative Associative Search Service based on GETA (Generic Engine for Transposable Association)

「想」IMAGINE

想いを連ね、想いを映し、その響きに耳を澄まそう



図書館の蔵書DB、古書店の在庫DB、テーマ別新書ガイド、百科事典、博物館収蔵品DB、新聞記事DB、報道写真アーカイブなど、さまざまなジャンルの情報源から、関連する情報を連想計算で収集して一覧表示します。その中からさらに深く知りたい情報を指定して、それを基点に連想を深めることが可能です。従来の検索では不可能だった、多様な情報によって自分の想いを確かめながら深めていける情報探索が経験できます。人間の連想を刺激して発想力を高める、まったく新しい情報利用環境をお楽しみください。



<http://imagine.bookmap.info/>

Summary

- e-Science and Data-centric Science
- CSI: Cyber Science Infrastructure in Japan
- Information Explosion Era
- Digital libraries and Institutional repositories
- Collaborative activities of Universities and NII