



HOKKAIDO
UNIVERSITY



Experimental Database XAFS as an example

Kiyotaka Asakura
Institute for Catalysis
ICAT, Hokkaido University

The presentation is based on the discussion of Q²XAFS 2017 and the following paper
K. Asakura, H. Abe, M. Kimura, Journal of Synchrotron Radiation 2018, 25, 967-971
doi:10.1107/S1600577518006963.

Introduction

Why do we need database?

1. We can guess the structure faster and more correctly.

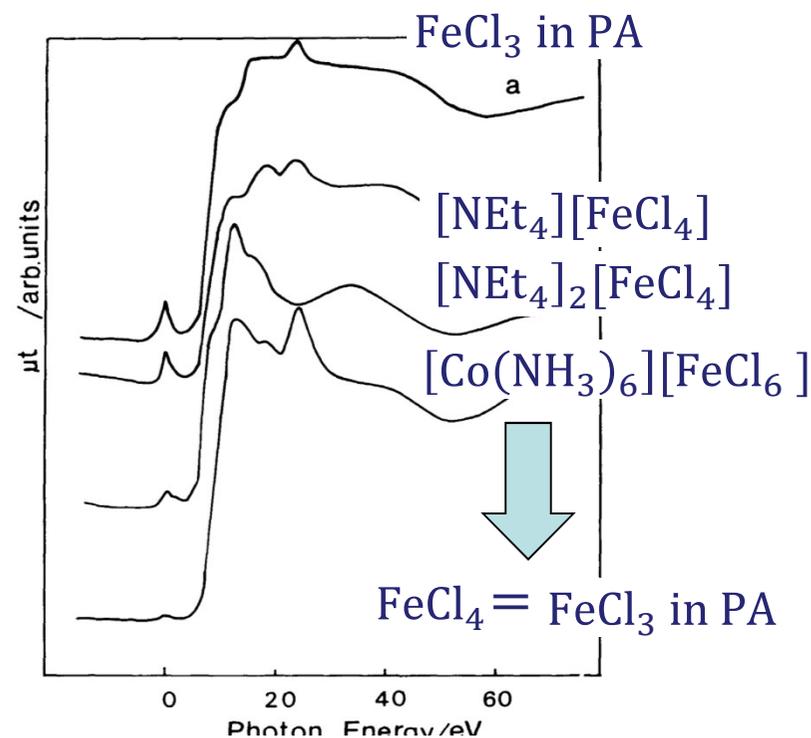
Strong demand from beginners and non-professionals

2. To create a new concept.

3. AI and DEEP LEARNING

will give us directly the structure from the spectra based on database in near future.

Data-driven analysis



Kiyotaka, A.; Isao, I.; Haruo, K.; Takaaki, K.; Hideki, S.
Bulletin of the Chemical Society of Japan **1985**, *58*,
 2113.

Introduction

3

Why do we need database?

1. We can guess the structure faster and more correctly.

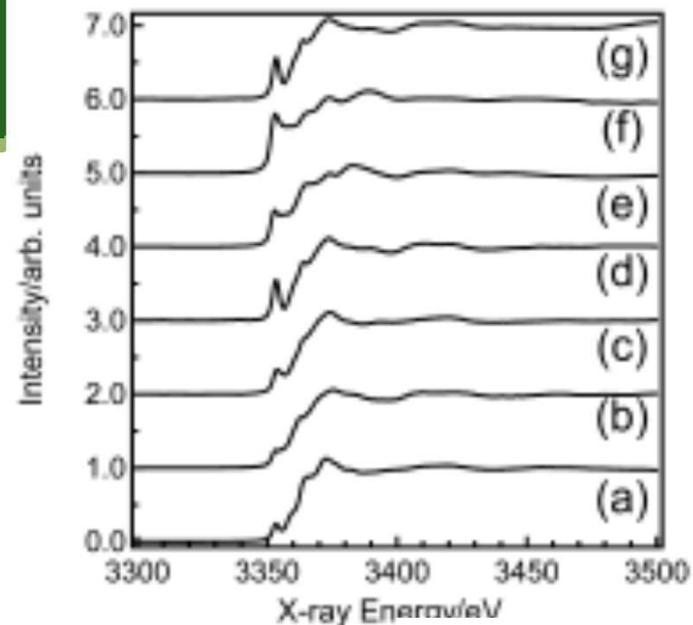
Strong demand from beginners and non-professionals

2. To create a new concept.

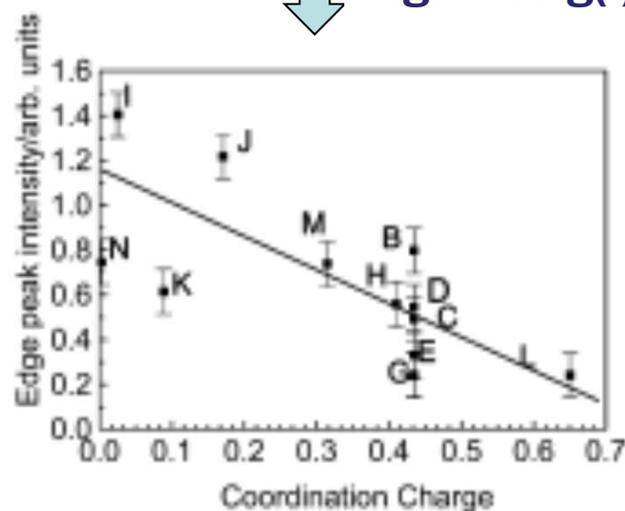
3. AI and DEEP LEARNING

will give us directly the structure from the spectra based on database in near future.

Data-driven analysis



Covalency and L3 edge peak height of Ag(I) are in proportion.



Introduction

Why do we need database?

1. We can guess the structure faster and more correctly.

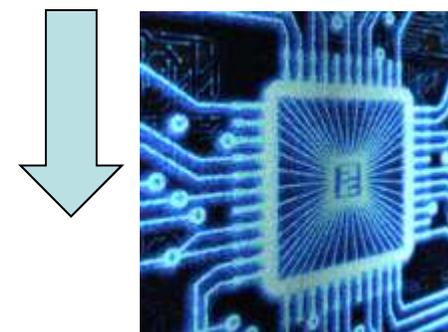
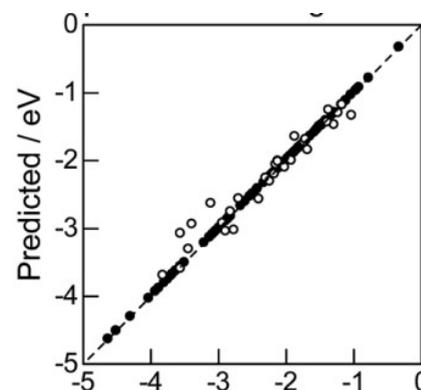
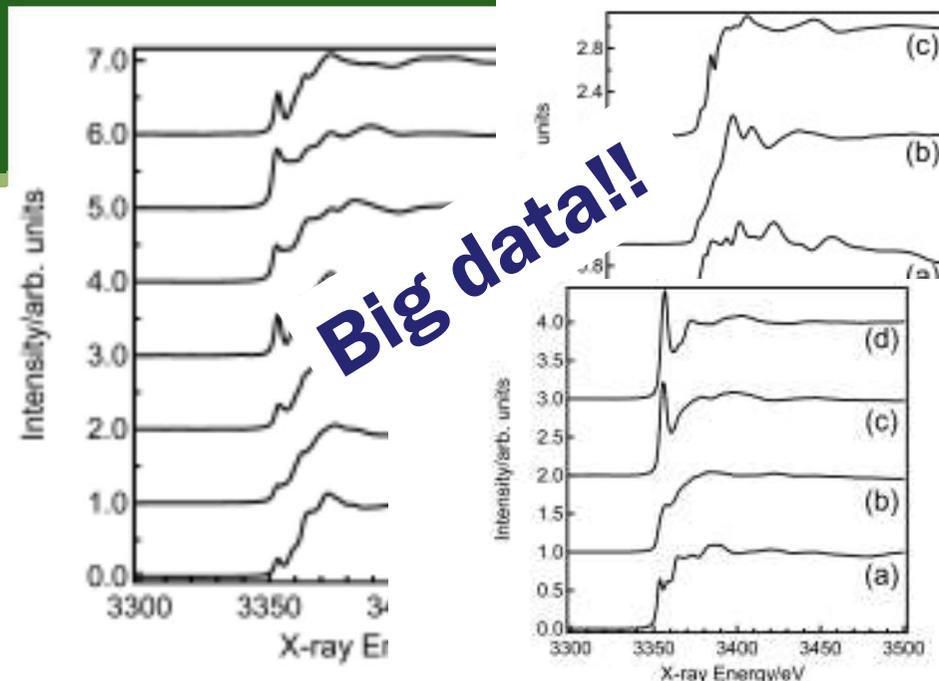
Strong demand from beginners and non-professionals

2. To create a new concept.

3. AI and DEEP LEARNING

will give us directly the structure from the spectra based on database in near future.

Data-driven analysis



AI and DEEP LEARNING
Create knowledge

XAFS DATABASE

<http://cars.uchicago.edu/xaslib/search>

Graphical

<http://ixs.iit.edu/database/>

Text base

SPring-8 BL14 Data repository <https://sp8dr.spring8.or.jp/portal/dspace>

<http://www.esrf.eu/home/UsersAndScience/Experiments/XNP/ID21/php.html>

Sulfur database

https://www.cat.hokudai.ac.jp/catdb/index.php?action=xafs_login_form&pnid=2

XAFS SOCIETY OF JAPAN DATABASE

Database	URL	Features
F.W. Lytle database	http://ixs.iit.edu/database/	Text format with periodic table interface All data obtained by F.W. Lytle 19,307 records covering 74 elements
CARS	http://cars.uchicago.edu/xaslib/search	Text (XDI) format with periodic table interface and graphical output of spectra (Newville et al., 2015) 171 records covering 13 elements Measured at the Advanced Photon Source, SSRL, and National Synchrotron Light Source Includes suites (tagging-related spectra) and user ratings Only available to SPring-8 users
SPring-8 BL14	https://sp8dr.spring8.or.jp/portal/dspace	725 records Compressed text file (ZIP) format with metadata that contains vendor and lot number
European Synchrotron Radiation Facility, ID21 group	http://www.esrf.eu/home/UsersAndScience/Experiments/XNP/ID21/php.html	Sulphur database 41 (inorganic) and 26 (organic) records for S compounds Graphical output, but text format is available User name is included
Hokkaido University ICAT and JXS	https://www.cat.hokudai.ac.jp/catdb/index.php?action=xafs_login_form&opnid=2	209 records and 24 elements Text format with metadata Includes user name, beam line, and facility Open access, but data uploading is restricted to users with ID and password

History

Catalysis Research Center(now ICAT) started to build the database of XAFS for future usage of data science in 2011, collaborating with XAFS Society of Japan(XSJ) .

System

Web is prepared by ICAT and XSJ members upload their data.

Fundamental principles

- **Volunteer base** = anyone who is registered can deposit his/her data .
- **Open database** = anyone in the world can see and use it with citation.
- **Low cost** = no one manage the data because of no budget.
- **Simple dataset** =text base. Only energy and μt of raw spectrum.
- **Compatibility** = Foil data simultaneously measured should be deposited.
- **Target** = Mainly standard compounds. But some repository for the published data.
- **Guarantee for accuracy = majority rule.** We allowed the many data depositions of the same samples.



Majority rule = basic idea of the XAFS Database in XSJ

Quantity overcomes quality

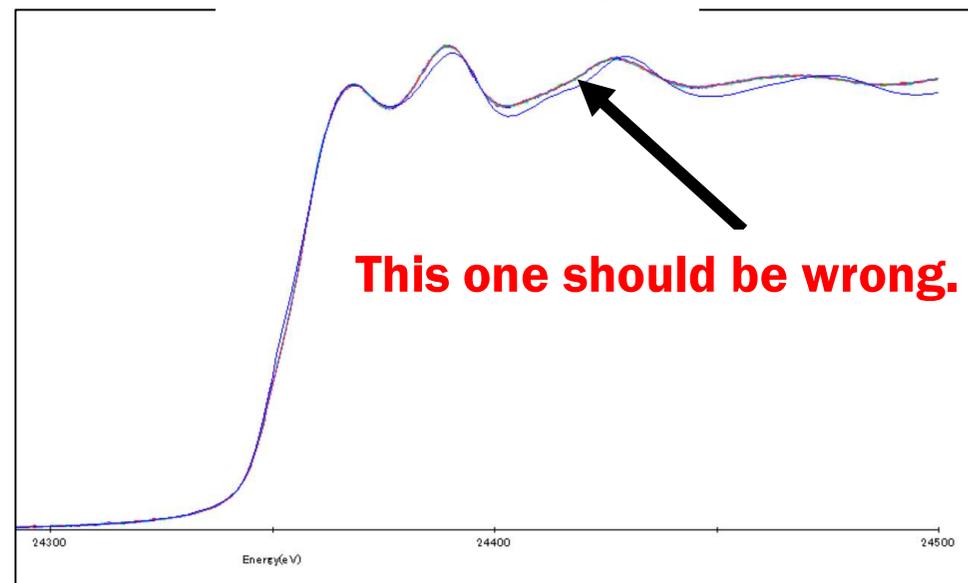
8

If we have 60 data with same quality(S/N ratio) of the same standard compounds.

59 data sets are the same spectra and one is different from the other. The one should be rejected.

It requires number of data of the same compounds.

Rh9Co1 foil



Inside database

https://www.cat.hokudai.ac.jp/catdb/index.php?action=xafs_login_form&opnid=2

北海道大学触媒科学研究所 Institute for Catalysis, Hokkaido University
触媒科学研究データベース

XAFS database

View data

Input/change the data

mail address: askr@cat.hokudai.ac.jp

password:

I store an e-mail address and a password for a browser

login

[The one where forgets a password.](#)

XAFS database top XAFS database

View data

Atom: periodic table

Edge: K L₃ L₂ L₁ M

Facility: PF PFAR Spring-8 RITS SAGA-LS UV-SOR EELS in-lab others

Correspondence:

Date of measurement:

search

Search

Download Click

PF	T	Sample	Date	Facility	Method
		Kiyotaka Asakura	1984.06.20	ATHENA	DETAIL
				REX	
PF	T	Kiyotaka Asakura	1984.06.20	ATHENA	DETAIL
				REX	
PF	T	Kiyotaka Asakura	2017.02.25	ATHENA	DETAIL
				REX	

<< < 1 2 3 4 >> Total to 194

Atom	Edge	Sample JCR/XAFS	crystal	Facility	Method	Correspondence	Date	Download	Detail
Nb	K	NbO2	Si(311)	PF	T	Kiyotaka Asakura	1984.06.20	ATHENA REX	DETAIL
		JXAFS-1703150002							
Nb	K	Nb2O5	Si(311)	PF	T	Kiyotaka Asakura	1984.06.20	ATHENA REX	DETAIL
		JXAFS-1412230001							
Fe	K	Fe2O3 XANES	Si(111)	PF	T	Kiyotaka Asakura	2017.02.25	ATHENA REX	DETAIL
		JXAFS-1705190002							
Co	K	CoTPP_01	Si(111)	PF	T	Kiyotaka Asakura	2017.02.24	ATHENA REX	DETAIL
		JXAFS-1705190001							
Rh	K	Rh foil	Si(311)	PF	T	Kiyotaka Asakura	2017.05.03	ATHENA REX	DETAIL
		JXAFS-1705080004							
		RhCl3	Si(311)	PF	T	Kiyotaka Asakura	2017.05.03	ATHENA	

Inside database

https://www.cat.hokudai.ac.jp/catdb/index.php?action=xafs_login_form&opnid=2



View data

Atom

Edge

Facility K L₃ L₂ L₁ M

Correspondence

Date of measurement

<< < 1 2 3 4 > >> Total to 194

Atom	Edge	Sample JCRXAFS	crystal	Facility	Method	Correspondence	Date	Download	Detail
Nb	K	NbO ₂	Si(311)	PF	T	Kiyotaka Asakura	1984.06.20	<input type="button" value="ATHENA"/>	<input type="button" value="DETAIL"/>
		JVAFEC-1703150003						<input type="button" value="PDF"/>	

Downloaded data

```
DATE=84.U6.2U
PLACE=PF
BEAMCURRENT(mA)=100
BEAMENERGY(GeV)=2.5
MONOCHROMATOR=Si(311)
OPTICS=
METHOD=Transmission(T)
*GAS IO/I = Ar(50)N2(50)/Ar(100)
MEASUREMENT TEMPERATURE=RT
REFERENCE=K_Nb205_Si311_19840620.txt
E(eV)     $\mu$ t
18507.21312 -1.281775946
18518.04279 -1.284004024
18528.88541 -1.285567492
```

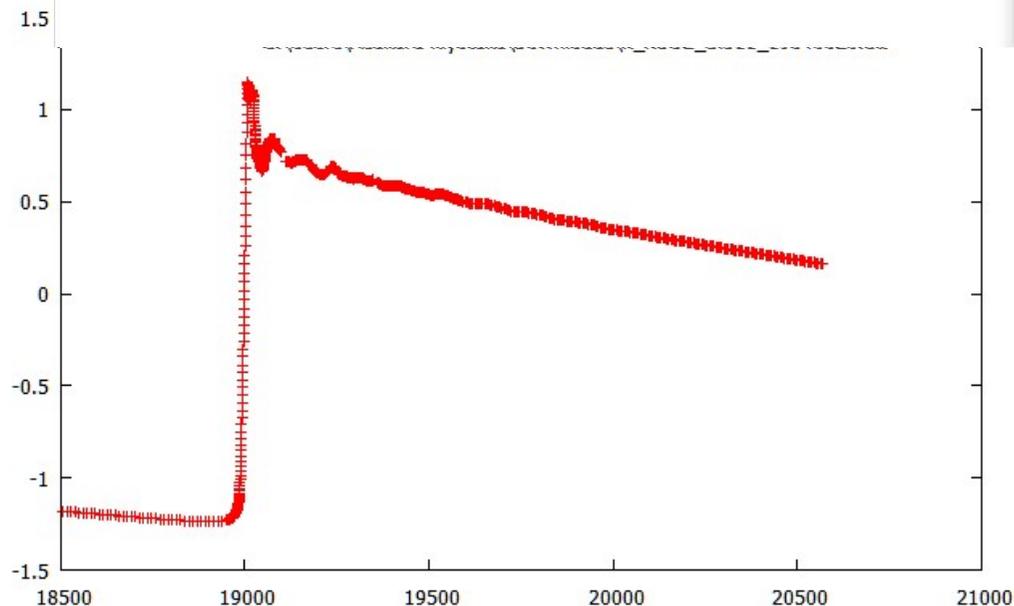
```
18539.74101 -1.288068686
18550.60962 -1.288701973
18561.49126 -1.290779827
18572.38594 -1.292854447
18583.2937 -1.294284416
18594.21456 -1.295542402
18605.14854 -1.297676177
18616.09567 -1.299018371
18627.05596 -1.30066751
18638.02944 -1.302029982
18649.01614 -1.304040825
18660.01608 -1.306237246
18671.02929 -1.307050868
```

```
GNUPLOT
Version 4.6 patchlevel 6 last modified September 2014
Build System: MS-Windows 32 bit
```

```
Copyright (C) 1986-1993, 1998, 2004, 2007-2014
Thomas Williams, Colin Kelley and many others
```

```
gnuplot home: http://www.gnuplot.info
faq, bugs, etc: type "help FAQ"
immediate help: type "help" (plot window: hit 'h')
```

```
Terminal type set to 'wxt'
gnuplot> load 'C:\Users\Asakura Kiyotaka\Downloads\K_Nb02_Si311_19840620.txt'
gnuplot> 18507.21312 -1.179182569
"C:\Users\Asakura Kiyotaka\Downloads\K_Nb02_Si311_19840620.txt", line 1: invalid
gnuplot> plot "C:\Users\Asakura Kiyotaka\Downloads\K_Nb02_Si311_19840620.txt"
warning: Skipping unreadable file "C:\Users\Asakura Kiyotaka\Downloads\K_Nb02_Si311_1
No data in plot
gnuplot> help
gnuplot> plot 'C:\Users\Asakura Kiyotaka\Downloads\K_Nb02_Si311_19840620.txt'
gnuplot>
```



What are the problems?

- **Now August 12 2018. We have 218 datasets uploaded.**
 - Too small and too biased.
 - Duplication and multiplication must be necessary.
- Autodepository system and agreement of data owner.
- **Sample name and data file rule** which is internationally available should be decided unified for easy link to other Material data base. Not Local data file name but no universal one.(something like **Data DOI**)

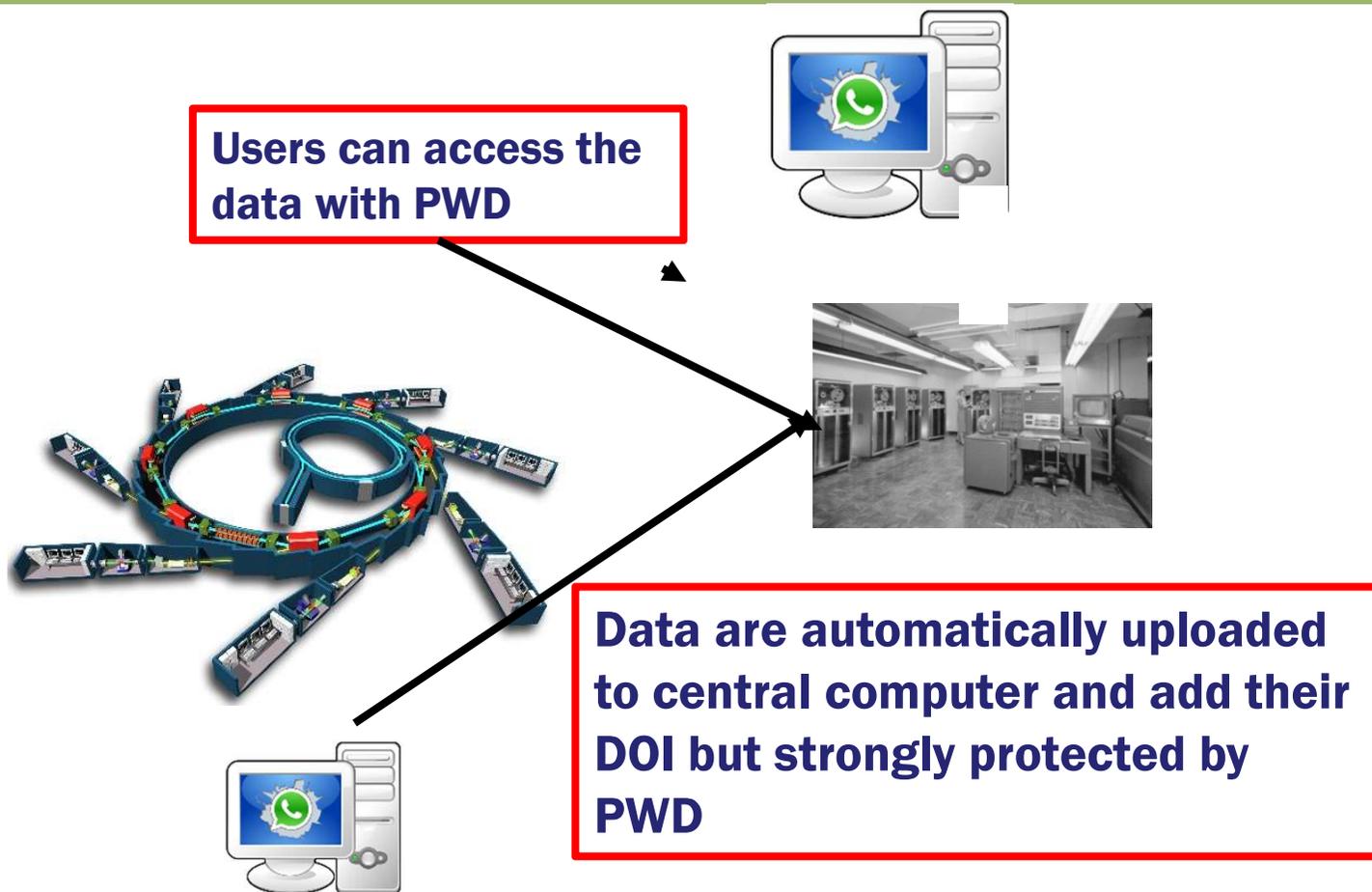


Proposals

- 1. All data measured at facility should be deposited automatically that should be named with an appropriate file name and metadata.**
- 2. First these data must be strongly protected.**
- 3. Open the data under some rule.
Publications, 10 years after deposition.**
- 4. Integrate all XAFS database**
- 5. Connect the XAFS database with other material database.**

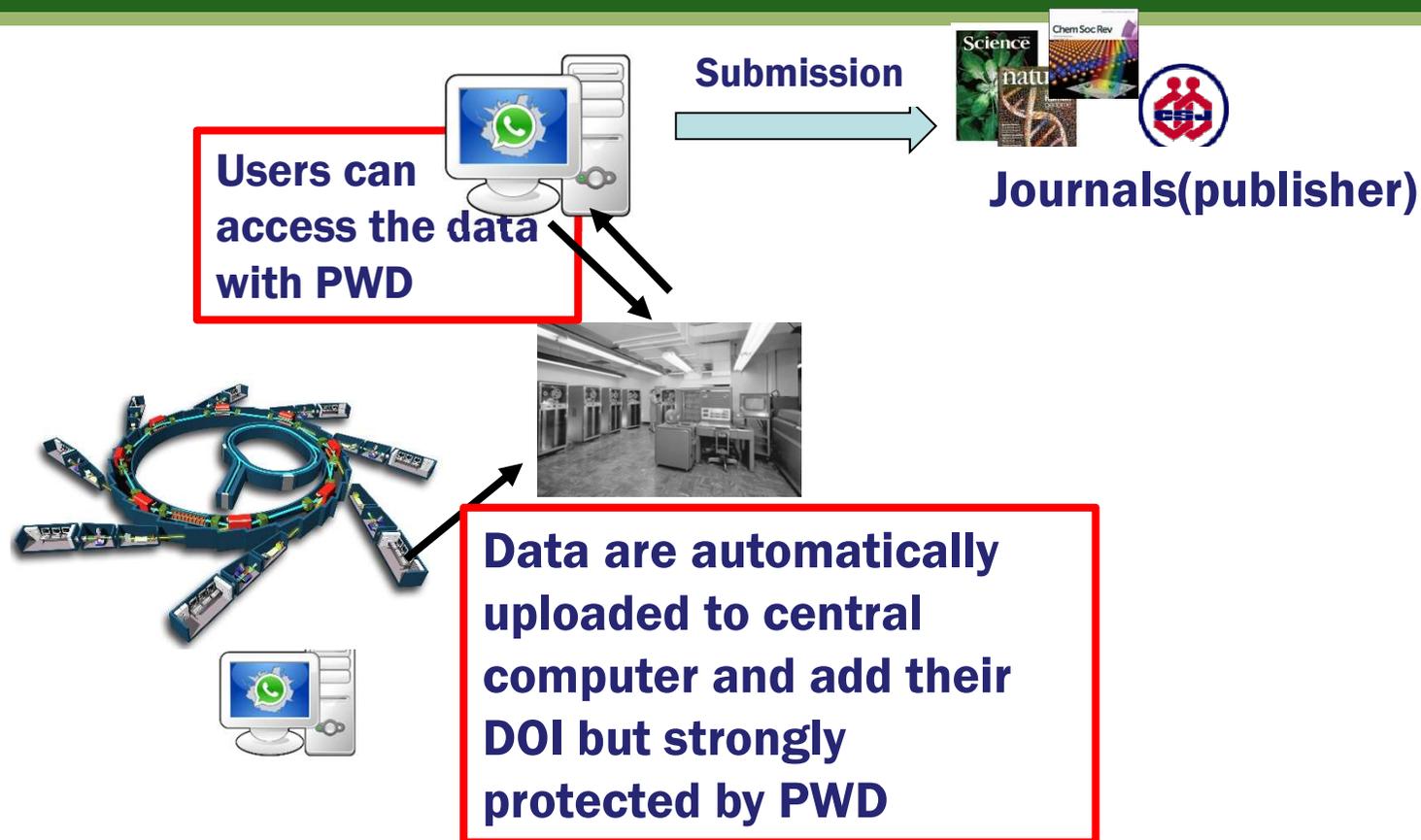
How to get more data = automatic depository and open after publication with DOI

14



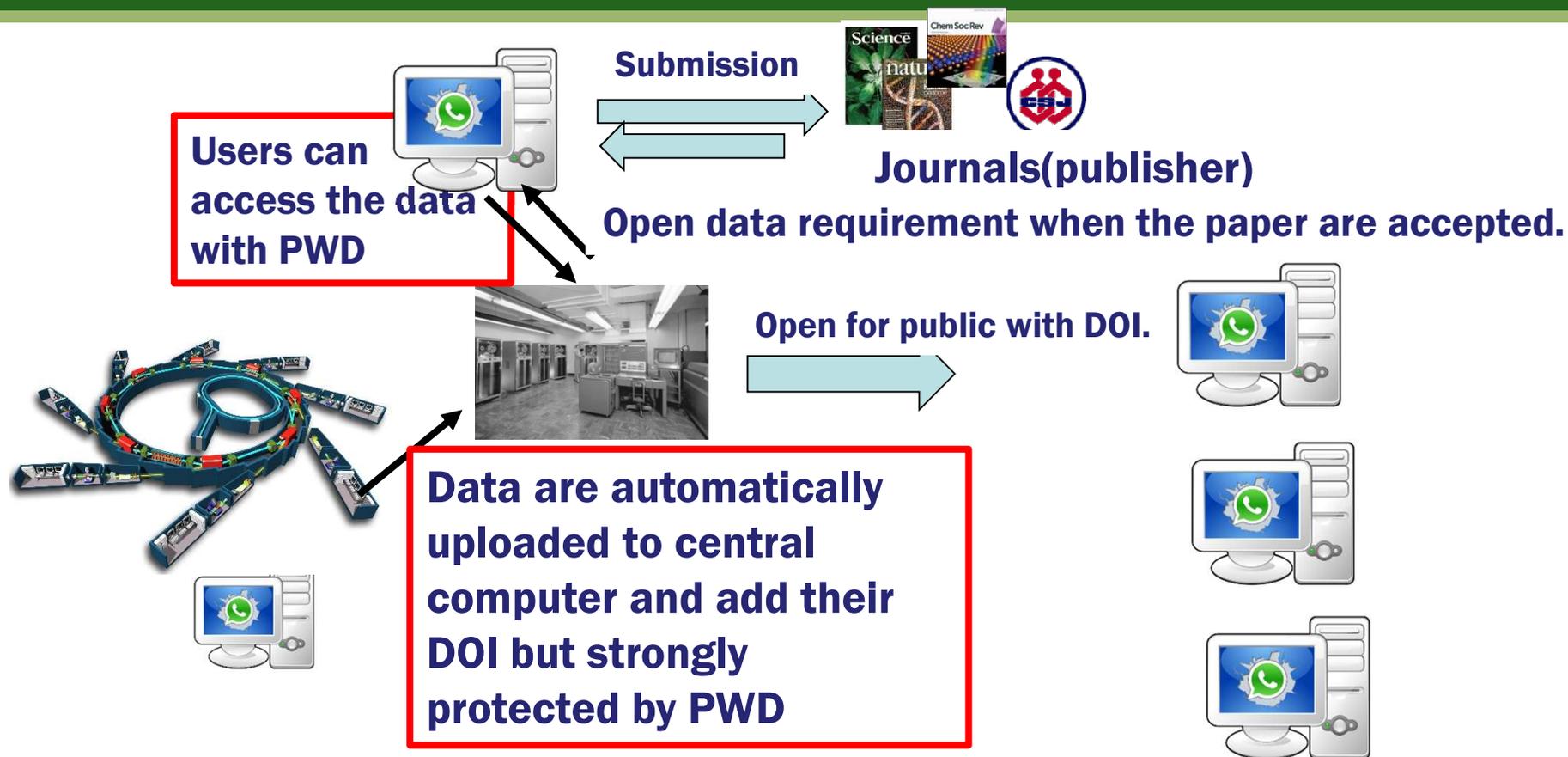
How to get more data = automatic depository and open after publication with DOI

15



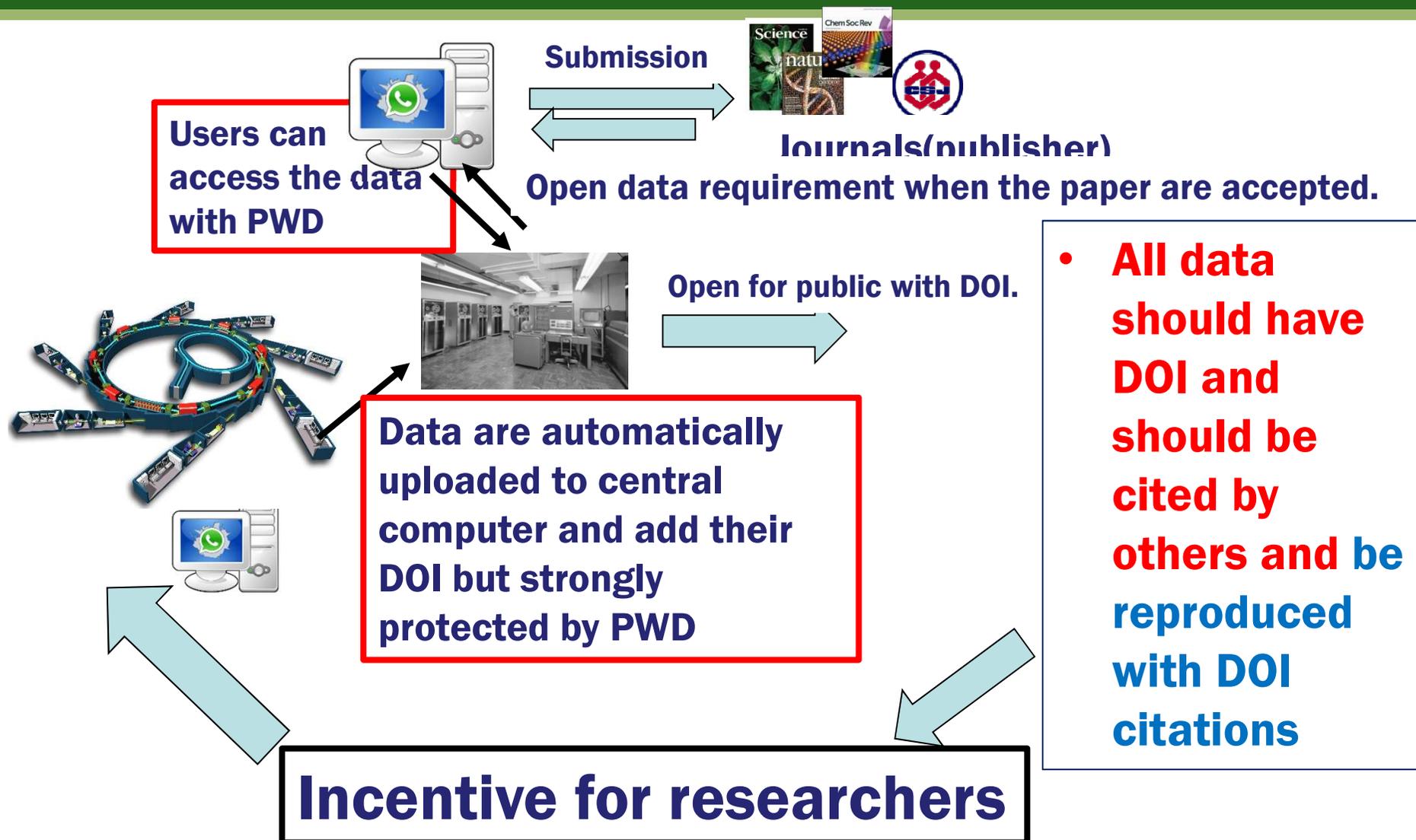
How to get more data = automatic depository and open after publication with DOI

16



How to get more data = automatic depository and open after publication with DOI

17



What we need.

- **To determine the international rule for open database.**
- **To determine the international agreement for metadata and file name(DOI).**
- **Metadata should contain**
 - Sample name(if standard CAS No.; if Unknown Unique Sample identifier)**
 - Measurement method and condition details Date, Facility and ContactPersonName**
- **file name FacilityBeamLineYDTContactPersonName_BranchNo**
- **We have started discussion on this September.**

Report on XAFS database

Sept 5 2018, Hokkaido University

Chair Prof Masao Kimura(PF,KEK)

**19 members from SP8, PF, AICHISR, RITS, UVSOR,
SAGALS and 3 users.**

1. Discussion about AutoDepository:

**Technically possible and some facilities started(SP8,
PF)**

**But there is a security problem (AICHISR is
negative)because many company users**

Ownership problems.

Report on XAFS database

2. How to Guarantee the data quality.

Authority measures and analyzes the standard compounds. Otherwise it is not reliable.

Many data without reliable factor are just gabbage.

(Many members opinion except me.)

3. Who manages the database.

Unified authority < SPring-8 /NIMS>

4. Is DOI of DATA really incentive? Just gathering and depositing the data will not be highly evaluated.

Ownership problems

5. Who is the owner of the data?

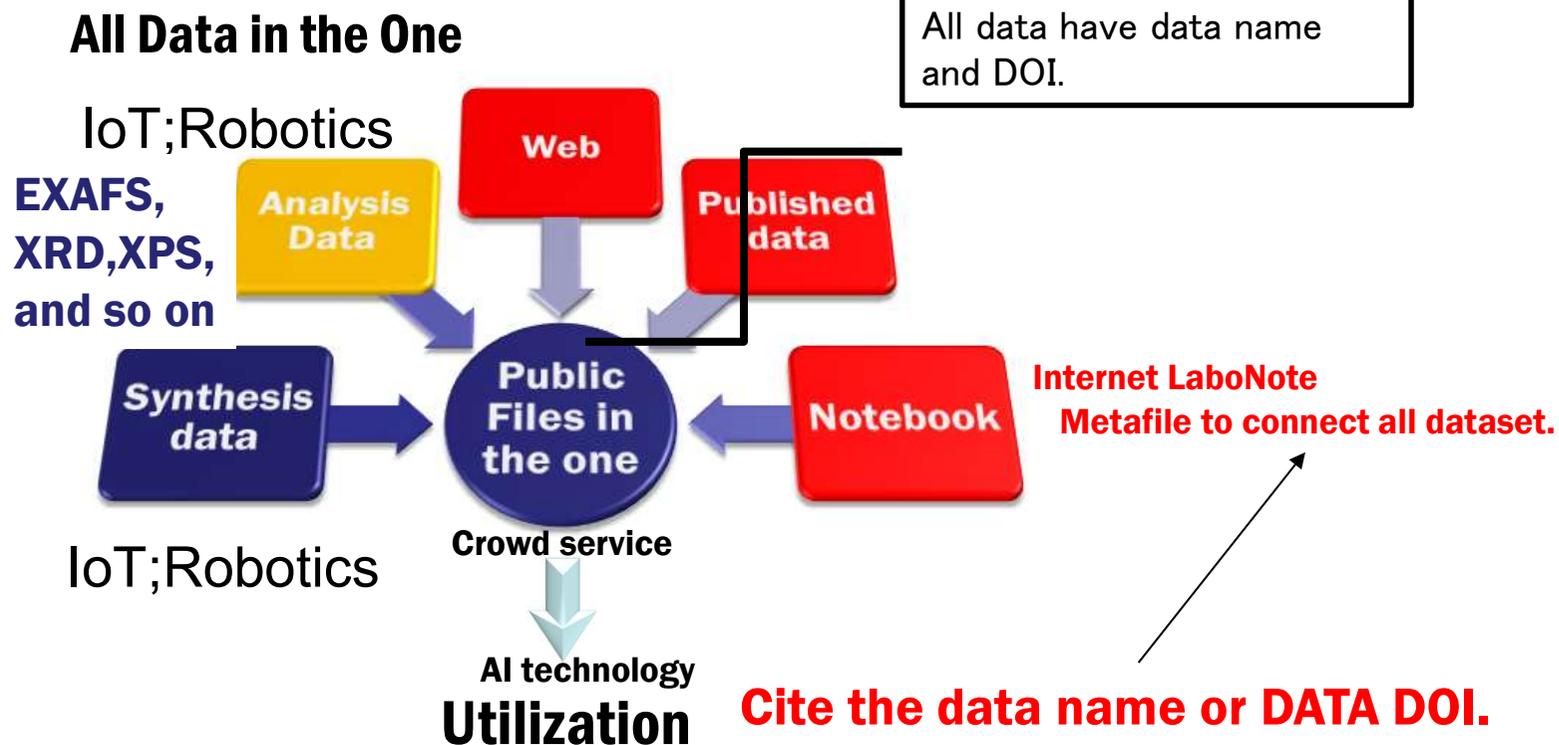
Operator? PI? Facility? Nation? Humans?

KEK thinks data belong to facility.

Astrophysicists think data belong to all.

Some agreement for open data must be necessary or culture change.

All data should be deposited on crowd and linked



acknowledgement

The presentation is based on the discussion with following people

**H. Abe(PF), M.Kimura(PF), Y.Niwa(PF), Y.Nitani(PF), Y.Tamenori(SP8),
T.Honma(SP8), S.Ito(NIMS), Tanaka(NIMS),
K.Takahashi(NIMS),T.Chigyo(NIMS), K.Funatsu(Univ. Tokyo),
J.Hasegawa(ICAT), M.Nishida(ICAT), J.Lauterbach(USC)
K.Shimizu(ICAT), W.Ueda(Kanagawa),H.Takaya(Kyoto U)
XAFS Society of Japan**