

オープンサイエンスに 関するアイディア

真板英一(NIES)

maita.eiichi@nies.go.jp

2015-09-19

トピックス

- ・ 人材も「オープン」に
- ・ 失敗事例の共有
- ・ 実務担当者の交流

「オープン」の中身

オープンサイエンスでは何がオープンなのか？

- ・ 文献（オープンアクセス）
- ・ 研究データ（データシェアリング）
- ・ 人材
 - ・ 人材へのオープンアクセス
 - ・ 人材のシェアリング

人材確保の責任

- ・ 「人材確保の必要性」は共通認識
- ・ どこがそのコストを負担するのか？
- ・ 今は負担を押し付け合っている状態
- ・ 単独負担が無理なら、協力してはどうか

人材コストのシェア

- ・ 人材確保のコストのシェアは可能か
 - ・ スペシャリスト、テクニシヤンのレンタル、派遣
- ・ DIAS が人材派遣で収益を得ることができる？
- ・ 「オープンサイエンスと労働者派遣法」という観点から法整備を議論する必要がある？

DIASと派遣業

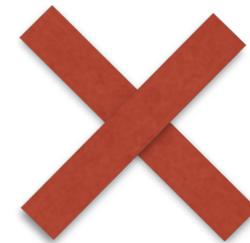
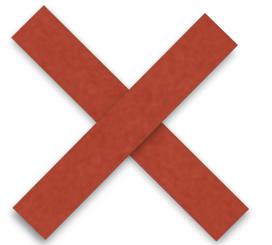
- ・ DIAS が派遣業を手がけることに拒否感があるかも知れない。
- ・ 実運用 = 「研究の論理」から「札束の論理」へ
- ・ 人材そのものがソリューションになり得るのなら、DIAS がそれを提供しても構わないはず。

失敗学

- ・ 共通認識を持つため、失敗事例を共有してはどうか
- ・ いわば「オープンサイエンスの失敗学」
- ・ データシェアリングについて言うと…

データシェアリングの 戦略的破綻の歴史

- ・ ICT 万能論、ITシステムさえ作れば全て解決
- ・ 研究者にトレーニングを
- ・ スペシャリストに頼む



ICT万能論の破綻

研究者トレーニングの破綻

NEWS FEATURE DATA SHARING

NATURE | Vol 462(9) September 2009

Empty archives

Most researchers agree that open access to data is the scientific ideal, so what is stopping it happening? **Bryn Nelson** investigates why many researchers choose not to share.



In 2003, the University of Rochester in New York launched a digital archive designed to preserve and share dissertations, preprints, working papers, photographs, music scores — just about any kind of digital data the university's investigators could produce. Six months of research and marketing had convinced the university that a publicly accessible online archive would be well received. At the time of the launch, the university librarians were worried that a flood of uploaded data might swamp the available storage space.

Six years later, the US\$200,000 repository lies mostly empty.

Researchers had been very supportive of the archive idea, recalls Susan Gibbons, vice provost and dean of the university's River Campus Libraries — especially as the alternative was to keep on scattering their data and dissertations across an ever-proliferating array of un-integrated computers and websites. "Some spent all this money, we spent all this time, we got the software up and running, and then we said, 'OK, here it is. We're ready. Give us your stuff,'" she says. "And that's where we hit the wall." When the time came, scientists could not find their data,

or didn't understand how to use the archive, or lamented that they just didn't have any more hours left in the day to spend on this business.

As Gibbons and anthropologist Nancy Fried Foster observed in their 2005 postmortem, "The phrase 'if you build it, they will come' does not yet apply to IRs [institutional repositories]."

A similar reality check has greeted other data-sharing efforts. Most researchers happily embrace the idea of sharing. It opens up observations to independent scrutiny, fosters new collaborations and encourages further discoveries in old data sets (see pages 168 and 171). But in practice those advantages often fail to outweigh researchers' concerns. What will keep work from being scooped, poached or misused? What rights will the scientists have to relinquish? Where will they get the hours and money to find and format everything?

Some communities have been quite open to sharing, and their repositories are bulging with

data. Physicists, mathematicians and computer scientists use arXiv.org, operated by Cornell University in Ithaca, New York; the International Council for Science's World Data System holds data for fields such as geophysics and biodiversity; and molecular biologists use the Protein Data Bank, GenBank and dozens of other sites. The astronomy community has the International Virtual Observatory Alliance, geo-

scientists and environmental researchers have Germany's Publishing Network for Geoscientific & Environmental Data (PANGAEA), and the Dryad repository recently launched in North Carolina for ecology and evolution research.

But those discipline-specific successes are the exception rather than the rule in science. All too many observations lie isolated and forgotten on personal hard drives and CDs, trapped by technical, legal and cultural barriers — a problem that open-data advocates see only just beginning to solve.

One of those advocates is Mark Parsons at

"We got the software up and running and said 'Give us your stuff'. That's when we hit the wall."
— Susan Gibbons

研究に必要な
スキルセットと
データ管理に必要な
スキルセットが
違いすぎて、
付け焼き刃では
どうにもならない。

THE LIBRARY REBOOT

As scientific publishing moves to embrace open data, libraries and researchers are trying to keep up.

BY RICHARD MONASTERSKY

A few passing students do a double take as Sayeed Choudhury waves his outstretched right arm. In his crisply pressed dress shirt and trousers, the engineer looks as if he is practising dance moves in slow motion. But he is really playing with astronomical data.

Standing in a US\$32-million library building opened last year at Johns Hopkins University in Baltimore, Maryland, Choudhury faces a 2-metre-by-4-metre 'visualization wall' of television screens. Pointing with his arm, he selects a picture of the Ring Nebula out of 40 images from the Hubble Space Telescope. Choudhury spreads his hands in a welcoming gesture and the nebula's rim of glowing orange gas fills the frame.

This wall is the brainchild of computer scientist Greg Hager and Choudhury, who directs digital research and curation at the library. For \$30,000, they and their team patched together monitors, processors and the Microsoft Kinect system that recognizes arm and body gestures. They placed the wall in the library last October as an experiment, allowing students and researchers to explore a few of the university's data sets, from star



Sayeed Choudhury demonstrates the visualization wall, part of Johns Hopkins University's drive to transform how its libraries and researchers deal with data.

「餅は餅屋」戦略

例えば、
「データ管理を図書館に集約する」
という動きが発生

というのが、私の
データシェアリングの
歴史認識

失敗学の利点

- ・ 失敗を繰り返さない
- ・ 議論の繰り返しを避けられる
 - ・ 「その話は何年も前に終わってる」

実務者コミュニティ

- ・ 集会にはだいたい代表者や研究者しか来ない
- ・ プロジェクトの目的や成果などの話しか出ない
- ・ システムのオペレーションや開発などの実務的なノウハウの共有が進まない
- ・ 実務者が集まるイベントがあってもいいのではないか

研究系データベース

Dev+Op 会議

- ・ 開発・運用担当者が実務的な話題を出す
 - ・ 予算、人員、ワークフロー
 - ・ 機材、更新、開発
- ・ DIAS の運用のヒントが得られるけど