Biocuration 2010
The Conference of the International Society for Biocuration
October 11-14, 2010 Odaiba, Tokyo, Japan
Fourth International Biocuration Conference

Nature Precedings : doi:10.1038/npre.2010.5078.1 : Posted 21 Oct 2010

MetaCuration Standards and Minimum Information about a Bioinformatics Investigation (MIABi)

TAN Tin Wee

Asia Pacific Bioinformatics Network (APBioNet)

NUS National University of Singapore

DEPARTMENT OF BIOCHEMISTRY

APBioNET

The Asia Pacific Bioinformatics Network

Endorsed By APAN TEL APBioNG ISCB

Pet Peeve 1

- Disappearing Databases and Web sites Services
- Instant access to Information:
  "When you see a citation, you should be able to read a pdf file of the paper in seconds"

Nature Precedings : doi:10.1038/npre.2010.5078.1 : Posted 21 Oct 2010

Pet Peeve 2

Nature Precedings : doi:10.1038/npre.2010.5078.1 : Posted 21 Oct 2010

"Now how did he get that conclusion?"
*"Nice graphic, but how the h*ll did he plot that?"*

- Insufficient information and data for reconstructing in silico experiments
- Full Disclosure in materials and methods, supplementary materials and all – Scientific Reproducibility

Pet Peeve 3

Nature Precedings : doi:10.1038/npre.2010.5078.1 : Posted 21 Oct 2010

"Compilation Error"
"Missing Library File"
"Loads of gibberrish"

- Cannot find software or test it, installation errors, missing dependencies
- Bioinformatics software should be downloadable and installable and work instantly

Pet Peeve 4

Nature Precedings : doi:10.1038/npre.2010.5078.1 : Posted 21 Oct 2010

陳定煒

Tan Tin Wee

"Is this guy
TW Tan or TT Wee?
When is he Tin Wee Tan?
Even the Chinese can't tell!"

Geof Bilder story
Cross Ref Contributor ID
STM Innov London 2009)
http://river-valley.tv/media/conferences/stm-innovations-2009/0102-Geoffrey_Bilder/

- Name Order, Name Variants, Name Changes
- Name Collisions
- Problem with Citations and H-Index
- Multiple Scopus IDs; Hopefully one ResearcherID
- Any number of OpenID, ORCID ?

Nature Precedings : doi:10.1038/npre.2010.5078.1 : Posted 21 Oct 2010

WIBNI
"Wouldn't it be Nice If"

- Reproduce everything the paper did instantly?
- State the author name and have all his papers correct citations etc all pop up?
- Review a paper and "play" back the authors' workflow and poke and prod the system at any step
- Take the software and plug in my data and pop up my own nice result?

WIBFI projects

"Wouldn't It Be Fun If"

Nature Precedings : doi:10.1038/npre.2010.5078.1 : Posted 21 Oct 2010

1984 Punchcards to do Chou Fasman secondary structure plots

1991/1992
Biocomputing, WAIS, Gopher and Web:
Amos Bairoch story, Prosite on WAIS
(Time before Gopher and the Web browser)

1994: Internet for Hearing Impaired
and Visually Handicapped

1996: 3D Photocopying: Internet2 and
Rapid Prototyping:

1998: Internationalized Domain Names (IDNs)

2001: S* Alliance: global bioinformatics teaching

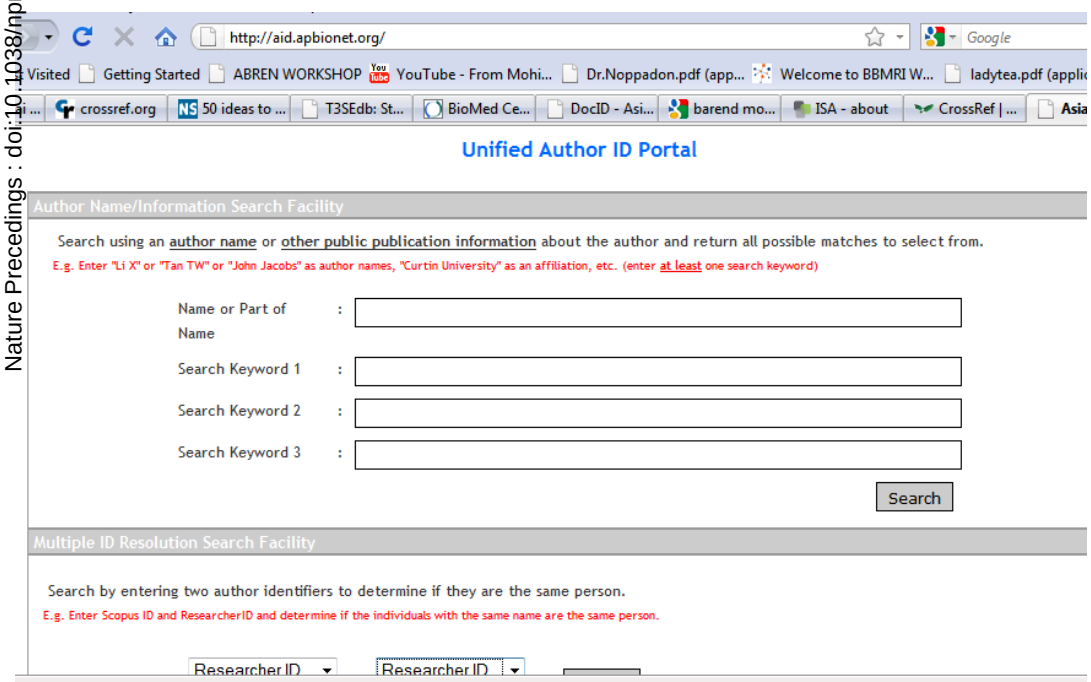
2005: Integrated Workflow and Manufacturing Systems
for BioSurveillance

2011: Synthesizing Left Handed Molecules

Nature Precedings : doi:10.1038/npre.2010.5078.1 : Posted 21 Oct 2010

Disambiguation: Unique identifiers

- Author identifiers
  equivalencing ResearcherIDs, ScopusIDs etc
- Unique identifiers in Scientific Terminology
- Accreditation; Accountability; Non-Repudiability



APBioNet's Prototype For Disambiguating Author IDs In http://aid.apbionet.org
- Self-Assertion Mode or Hybrid Assertion Mode
- Distributed system based On Internet Domain Name System

Reachabilitity:
Digital Object Deposition

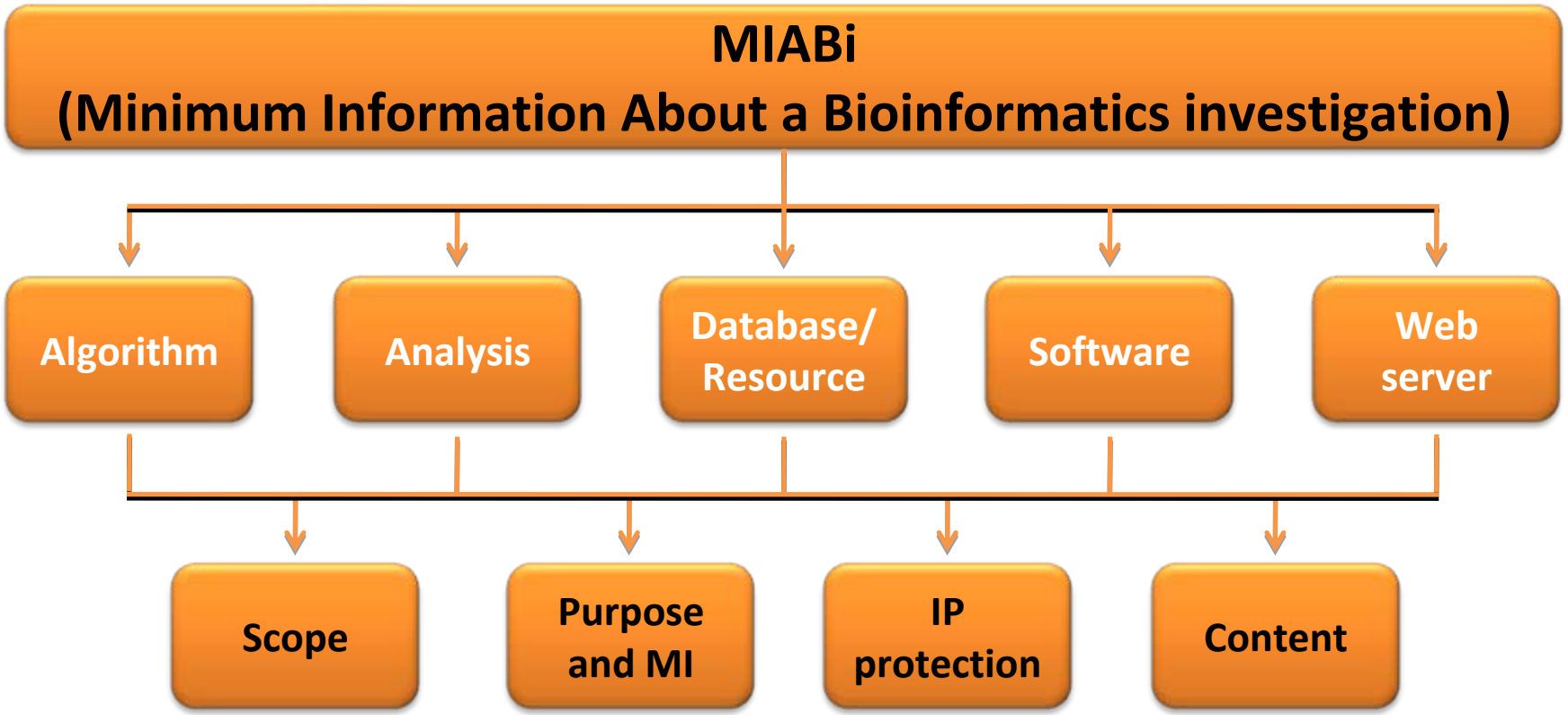
Nature Precedings : doi:10.1038/npre.2010.5078.1 : Posted 21 Oct 2010

- **Open Source**
- **Open Access**
- **Self-Deposition**
- **Escrow**
- **DocID**
- **Granularity**

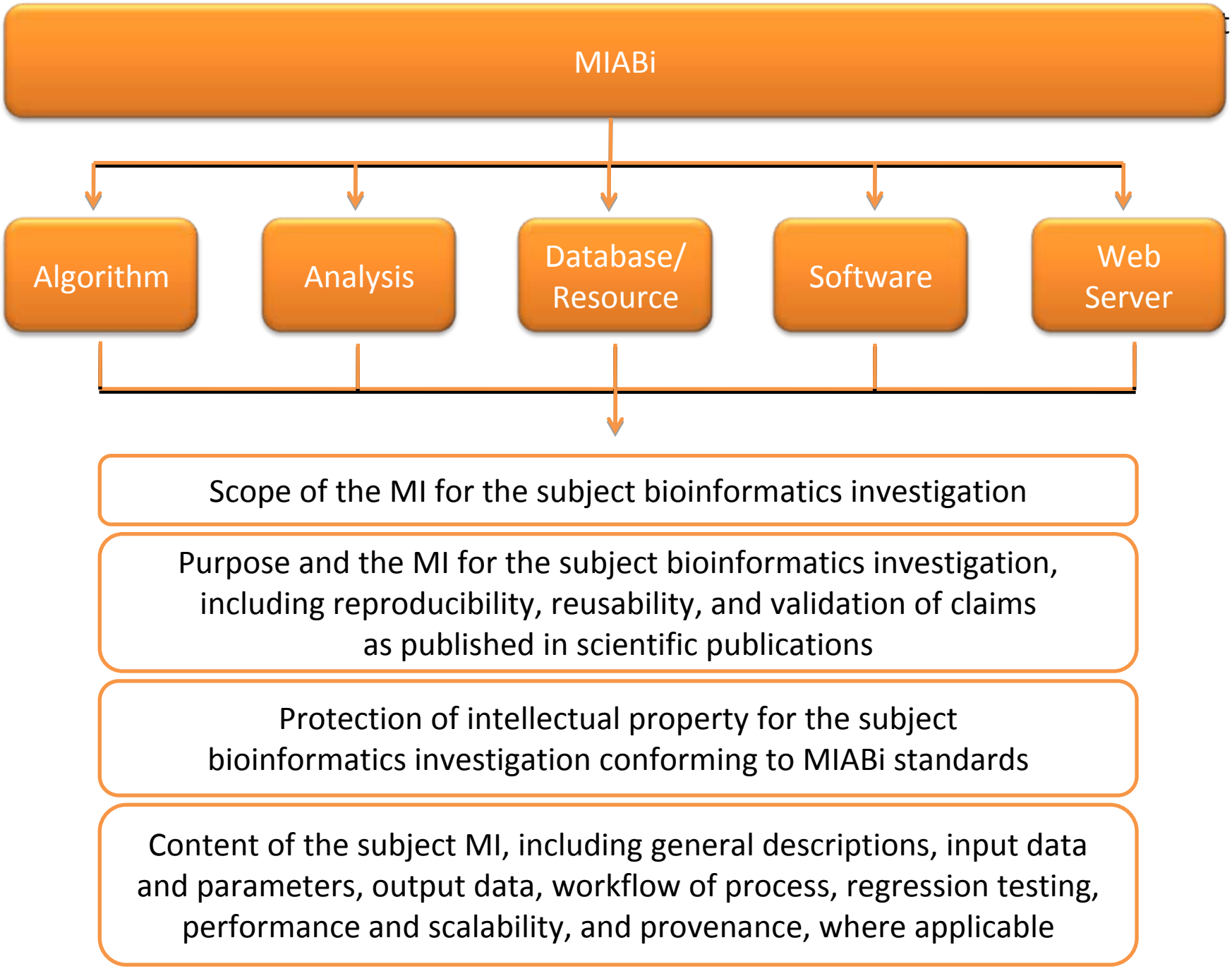
http://docid.apbionet.org

Sufficiency of Disclosure: MIABi

Nature Precedings : doi:10.1038/npre.2010.5078.1 : Posted 21 Oct 2010



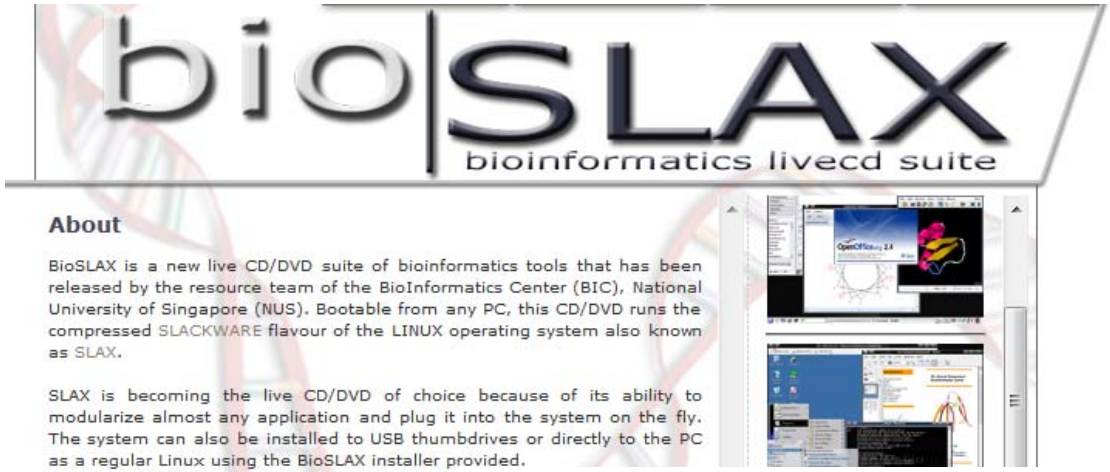
Supporting Reproducibility

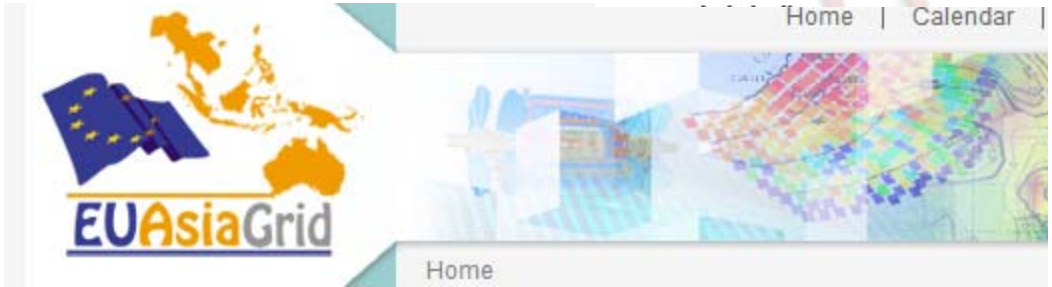


Nature Precedings : doi:10.1038/npre.2010.5078.1 : Posted 21 Oct 2010

Reinstantiability

- BioSlax ”LiveOS” images
- Lzm compressed files of deltas
- Citrix Xen Hypervisors and full hardware virtualisation
- Grid Computing on ***EUAsiaGrid.org***
- Cloud Computing on Amazon AWS





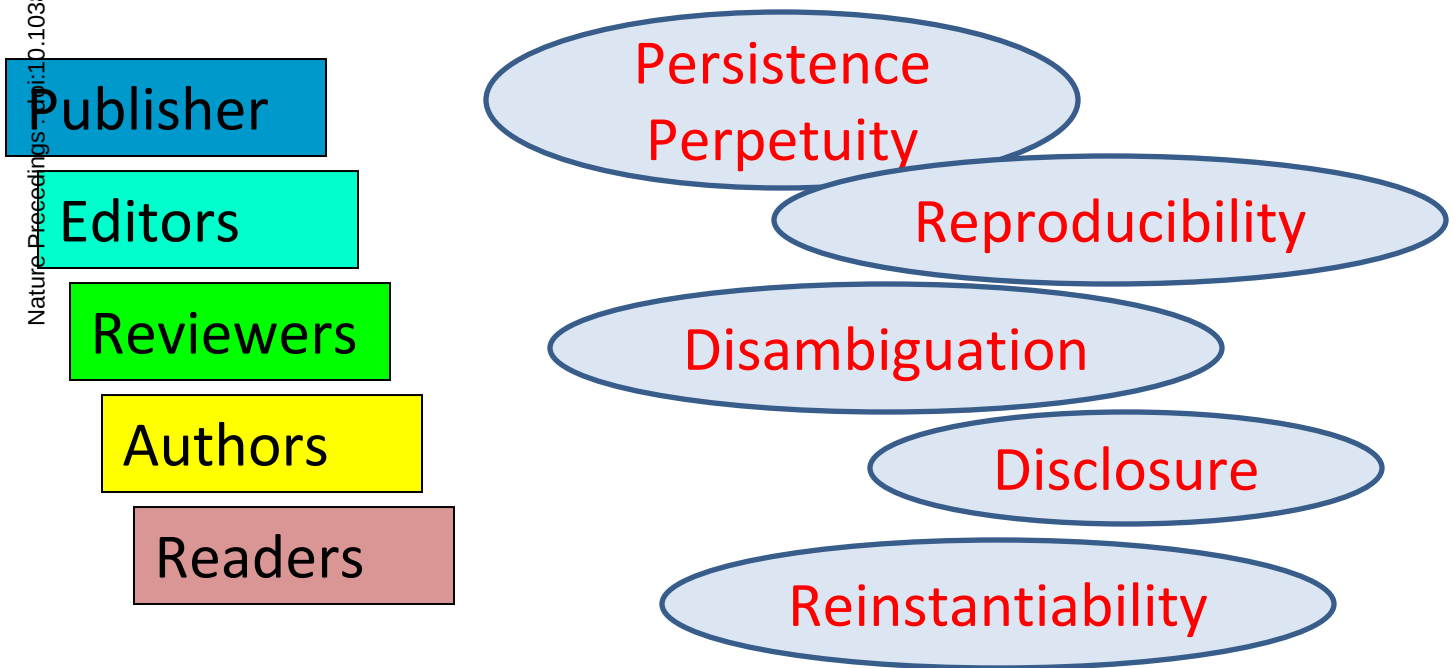
Open Science Grid

Nature Precedings : doi:10.1038/npre.2010.5078.1 : Posted 21 Oct 2010

People-Papers Platform: InCoB

- International Conference on Bioinformatics since 2002
- 2006: Publication of top papers in BMC Bio
- 2010: BMC Bioinformatics; BMC Genomics; Immunome Research; Bioinformation; IPSJ Transactions

Nature Precedings : doi:10.1038/npre.2010.5078.1 : Posted 21 Oct 2010



WIBMFI

Nature Precedings : doi:10.1038/npre.2010.5078.1 : Posted 21 Oct 2010

Wouldn't it be MORE fun if:

- BioDB100 exemplar of how this will all work together
- MetaCurate all kinds of databases:
  Maintain Meta-Information About Databases which curate biological knowledge
- Basis for Interoperability
- Building the BioInformation Infrastructure for future generations
- Cleaning up our act
  TCP/IP and IETF: Internet Engineering Task Force

The BioDB 100 Opportunity

- InCoB Paper submission process
- Editors and Reviewers
- MIABi-Compliant Disclosures
- Authors-disambiguated; maybe Terminologies? Ontologies?
- Deposited Databases, Datasets, Webservices
- Reinstantiability with virtual machines
- Reproducibility on grids and clouds
- Regression testing with bioinformatics software
- Microcitation and Nanocitation and Metacuration
- To the future and beyond!

Nature Precedings : doi:10.1038/npre.2010.5078.1 : Posted 21 Oct 2010

Contact

Tan Tin Wee
Bioinformatics Centre Lab/
Dept of Biochemistry
National University of Singapore
Email: tinwee@bic.nus.edu.sg

Nature Precedings : doi:10.1038/npre.2010.5078.1 : Posted 21 Oct 2010