

出國報告（出國類別：國際研討會）

2014 科學資料研討會 （SciDataCon 2014）

服務機關：國立中興大學 土木工程學系

姓名職稱：楊明德 教授

派赴國家：印度新德里

出國期間：10 月 31 日至 11 月 6 日，2014

報告日期：11/22，2014

摘要

本報告紀錄本人赴印度新德里參加 11 月 2 日至 11 月 6 日之 2014 科學資料研討會 (SciDataCon 2014)，並協助舉辦 11 月 2 日黑客松 (Hackathon)。研討會中討論許多開放資源共享的議題，網路上的開放資源 (Open sources) 約可分為開放資料 (Open data) 與開程式 (Open program)，而其特性皆基於網路世界之分享、合作、進化。就資料庫而言，開放資料的好處在於資料可被更多不同領域的人利用，而產生更大的效益，甚至發展出新的資料，共同分擔維護資料庫的成本與技術支援，個人也可因此獲社群認可，並參與國際合作聯盟。

目次

一、	目的-----	4
二、	過程-----	4
三、	心得及建議-----	7
四、	附錄-----	8

1. 目的

赴印度新德里參加 11 月 2 日至 11 月 6 日之 2014 科學資料研討會 (SciDataCon 2014)，並協助舉辦及參與 11 月 2 日黑客松 (Hackathon)。此一國際會議由「科學與技術資料委員會」(The Committee on Data for Science and Technology, CODATA)主辦，CODATA 是「國際科學理事會」(International Council for Science)所屬的跨領域科學常設委員會，目前有 17 席科學聯合會會員、26 個國家會員 (包括台灣)。本次會議的目的在討論並促成開放資源 (Open sources)，尤其是開放資料 (Open data)，以利網路世界之分享、合作、進化。就資料庫而言，開放資料的好處在於資料可被更多不同領域的人利用，而產生更大的效益，甚至發展出新的資料，共同分擔維護資料庫的成本與技術支援，個人也可因此獲社群認可，並參與國際合作聯盟。

2. 過程

日期 工作事項

- 10/31 台灣--> 印度新德里
- 11/1 黑客松 (Hackathon) (早期生涯資料專家工作小組，ECDP) 會前會
- 11/2 協助舉辦及參加黑客松 (Hackathon)
- 11/2-5 參加「2014 科學資料研討會 (SciDataCon 2014)」國際研討會
- 11/5-6 印度新德里-->台灣

2014 科學資料研討會 (SciDataCon 2014)

1966 年創設的「科學與技術資料委員會」(The Committee on Data for Science and Technology, CODATA) 是「國際科學理事會」(International

Council for Science) 所屬的跨領域科學常設委員會，目前有 17 席科學聯合會會員、26 個國家會員（包括台灣）。2014 年由印度國家科學院（Indian National Science Academy）在新德里主辦科學與技術資料委員會（CODATA，以下行文以 CODATA 代表）會員大會暨科學資料研討會（SciDataCon 2014，以下行文以 SciDataCon 代表）。

黑客松 Hackathon（早期生涯資料專家工作小組，ECDP）會前會

黑客松（Hackathon）會前會做事前準備工作，包括討論黑客松（Hackathon）執行工作項目，與即將參與黑客松（Hackathon）的台灣科學家們（如中研院資訊所莊庭瑞研究員、中研院資訊所與荷蘭屯特大學地球觀測及地理資訊科學學院博士候選人鄧東波先生、中研院生物多樣性中心研究助理麥館碩先生，及中興大學李士傑先生）在會議前有一個較長的時間討論與交換資訊，讓大家提出問題與討論。

協助舉辦及參加黑客松（Hackathon）

黑客松（Hackathon）是一個重要的挑戰，這次跟 16 位來自 7 個國家的與會代表，以將近四個小時的時間一同討論 CODATA 社群中，科學資料共享的議題。在很短時間內，達成了對 CODATA 來說重要的目標，也是一個重大突破。事前籌備的會前會已整理了一些不同研討會的黑客松（Hackathon）成功案例，以便因應非程式設計者、對於資料處理工具概念也有限的參與成員。過去幾年，本人執行教育部科學人文跨科際人才培育計畫（SHS）中所導入的設計思考（Design Thinking），其中「設計思維」與舉辦工作坊分組實作的經驗提供不少建議。

在大會開始之前先用情境設計的方式，整理了不同的應用情境來幫助參與者熟悉黑客松（Hackathon）的動腦實作「暖身」。這個過程也幫助了我們其他的參與者可以多次不同程度的模擬、溝通與調整最適合的進行方式。然而會議開始進行後，就面臨到「拆鷹架」的關鍵時刻：原本設計用來輔助會議進行的工具，必須要挪開，來讓參與者充分表達自己是誰、關心什麼問題，

以及想要帶進什麼樣的議題與「偏見」，如何參與與進行。

我們花了大約一個半小時來充分介紹與討論參與者的背景，以及所想關切的議題。這些議題在整理之後，可以放到 CODATA 的未來執行脈絡中來發展。這次會議的討論議題現場調整後，分成四組做不同議題討論：

1. 如何發起/籌備黑客松（Hackathon）？
2. 資料整合案例研究
3. 公民科學資料生產
4. CODATA 任務群組（Task Group）產出成果應用：資料經紀與中介服務（brokering）

這次的黑客松（Hackathon）舉辦經驗也還有相當大的成長空間。Dr. Tiwari 特別針對國際科學理事會（ICSU，以下行文以 ICSU 代表）未來地球（Future Earth）計畫相當重視，希望早期生涯資料專家工作小組（ECDP）能夠發展出可以在未來地球（Future Earth）計畫中能夠提出的項目來參與之。單純就黑客松（Hackathon）而言，我們的主要議題有兩個主角：資料「例如：CODATA 任務群組數據集（Task Group Datasets）」、與人（國際會議的科學家們）。這兩者都需要更長的準備時間才能夠準備齊全。這次透過台灣科學家們（中研院資訊所、生物多樣性中心與中興大學等與會者）的全力協助，才讓現場研究者的精彩主題與重要方向，成為所有人共同挖掘探索的重點。

SciDataCon 大會議程

這次有幾十位包含會長在內的中國學員被拒絕簽證（或者被延遲發證），讓大會一開始就在做危機處理。專題演講邀請講者的內容都相當的精彩並且發人深省。這些講者名單 包括：Paul Berkman "全球永續知識發現（Knowledge Discovery for Global Sustainability）"，PLoS 的前主編

Theodora Bloom "British Medical Journal"; ICSU 未來地球(Future Earth) 計畫的科學委員會印度科學家 Bina Agarwal, "未來地球 (Future Earth)" 等。個別議程的部份都相當紮實, 可以學習如何讓台灣的學界更快速掌握到關鍵問題、少走冤枉路? 更重要的是, 安排這些講者, 同時也是在替 CODATA 自身所關心的議題尋找足以「解套」的解決方案; 例如 Ubiquity Press 出版社就有可能協助 CODATA 所出版的 資料科學期刊(Data Science Journal) 建議調整組織運作方向。公民科學 (Citizen Science) 是熱門的題目, 同時也在轉型成為新的領域; 這次 SciDataCon 遂規劃了兩個聯合的議程: 民眾搜集的資料提供大眾使用-鼓勵再利用與協作(Data from the People for the People - encouraging Re-use and Collaboration)。除此之外, 大會還安排了 "資料鏈進展程度? (How Well Is the Data Chain Going?)", 邀請資深天文科學家 Dr. Elizabeth Griffin 主持對談, 讓與會老中青的不同特色科學家, 從自己的角度來檢視科學資料分享的核心譬喻:「資料鏈」的銜接與運作是否順暢。這樣的反思其實很像是用傳統的形式, 但是卻做到近年來社群媒體雲端共同筆記服務(Hackpad)之類的效果。另外, 本校李德財校長也在本次「科學與技術資料委員會」會員大會中, 獲選連任 8 席執行委員之一。

本人參加研討會與參訪過程相關照片如附錄。

3.心得及建議

網路上的開放資源 (Open sources) 約可分為開放資料 (Open data) 與開放程式 (Open program)，而其特性皆基於網路世界之分享、合作、進化。就資料庫而言，開放資料的好處在於資料可被更多不同領域的人利用，而產生更大的效益，甚至發展出新的資料，共同分擔維護資料庫的成本與技術支援，個人也可因此獲社群認可，並參與國際合作聯盟。然而，目前開放資料的障礙是在基於私心欲獨自佔有資料庫，以便在研究上領先或商業上獲利，需花更多時間整理資料成特定型式才得以分享，資源的投入不保證回收等。為了鼓勵眾人分享資料，提高分享行動的價值是必要的，例如，已有一些期刊專以發表資料庫為主，短短一、二頁介紹資料庫的內容、型式、可應用方向等，使用資料庫而發表的期刊論文者，同樣可引用此短文為參考文獻，若是建置有用的資料庫，可能更容易變成為高引用作者 (High cited researcher)，受到學術界肯定。

台灣因政治因素，常常無法成為國際組織的正式會員。多年來都因個別多年來都因個別學者的努力以連繫國際組織與台灣學術間的關係。這些年來，由於中國的強大（包括政治、經濟與科技）相對更壓縮台灣的國際學術地位與空間，尤其在台灣經濟愈來愈差，而年輕學者愈來愈關起門來寫 SCI 論文，與國際組織逐漸脫鉤，讓台灣學術界離國際舞台愈來愈遠。未來應把握跟各國代表細緻交換意見的機會，越認識與會的代表與他們所來自的機構、文化，就越能夠有機會進一步合作。團結國內學者的研究成果，一起貢獻於國際舞台，讓國際組織難以平略台灣的學術成就，也許是我們可以努力的方向。

4. 附錄



研討會會場



研討會報到處



Keynote 演講



與美國郭教授合影



餐會



黑客松（Hackthon）會議

SciDataCon 2014

International Conference on
Data Sharing and Integration for Global Sustainability



Get Involved!	Scientific Programme	Side Events	Local Information	National Committee	News	Contact
---------------	----------------------	-------------	-------------------	--------------------	------	---------

Brew your own CO-DATA Hackathon

This **FREE**, pre-conference workshop is open to anyone interested in data science, data sharing, and data integration. It is organized by the **Career Data Professionals Group** and kicking about science.

Register for the Workshop

To help us gauge the attendance, **please register for the workshop** by clicking on the link below.

Note: It is essential to bring a laptop to do the hands-on exercises.

About the Workshop

The CODATA ECDP Working Group would like to invite early career / senior scientists and data professionals to participate in this CODATA Hackathon / Datathon. Discussions will centre around the research issues pertaining to CODATA Task Group datasets, as well as their re-use. However, BYOD (Bring Your Own Data) is encouraged in this community event. Preliminary information will be provided on <http://CODATA.hackpad.com> by late October. Following brief introductory sharings and an introduction to the various datasets, participants will be grouped according to the datasets in which they are interested. After approximately one hour of discussion (facilitated as necessary with appropriate tools), these groups will share their discoveries online and continue the dialogue.

Come along and explore the opportunities for data analysis and extraction of information with likeminded colleagues from around the world!

Get Involved!

Date: 2-5 November 2014
Venue: JNU Convention Centre

Register online!
Sponsor the Conference!
Draft Programme

News & Updates

NEW Special Day Registration offer: special registration fee will be applicable: INR 3000 (for Delegate) & 20/10/2014

Because of online payment difficulties, the early bird registration is exceptionally extended to 8 October 2014. 01/10/2014

Useful Information

- Conference Venue
- Accommodation
- Transportation
- Support for Indian Young Scientists
- Visas
- Tours

SciDataCon2014 on Twitter

Tweets

<http://www.scidatacon2014.org/associated-events>

黑客松（Hackthon）會議公告

SciDataCon 2014 : Final Programme (2/11/2014)				
	Track A. Dynamic Planet	Track B. Global Development	Track C. Transformations towards Sustainability	Track D: Data Science and Services
Sunday 2 November (Side Events and Inauguration)				
9.00-16.00	Registration			
9.00-12.00	Data Papers and their Applications: examples from Nature Publishing Group's Scientific Data; Training Room 107		Text and Data Mining Workshop; Lecture Hall I	WDS Members Forum; Lecture Hall III
12.00-13.00	Lunch			
13.00-16.00	SeMDAT: International Workshop on Semantics and Metadata for Data; Training Room 107	Early Career Data Professionals Hackathon; Committee Room 108	Text and Data Mining Workshop; Lecture Hall I	WDS Members Forum; Lecture Hall III
16.30-18.10	Opening Ceremony (JNU Convention Centre): Auditorium I			
18.15-18.30	Tea Break			
18.30-19.30	Bus Transfer			
19.30-	Welcome Dinner, Cocktail and Cultural Programme			
Monday 3 November				
9.00-17.00	Registration: Convention Centre Foyer			
9.00-9.45	Keynote I: Paul Berkman (Co-chair Programme Committee): Auditorium I			
9.45-11.15	Interoperability, Standards and Integration of Multidisciplinary Data; Lecture Hall II		Open Data Initiatives; Lecture Hall III	Open Data Infrastructure and Best Practices Auditorium I
11.15-11:45	Poster Viewing Coffee/Tea Break			
11.45-13.15	Data Rescue and Data at Risk; Lecture Hall I	Open Government Data and Smart Cities; Lecture Hall III		Materials Science Data; Auditorium I
13.15-14.30	Lunch			
14.30-16.00	Our Scientific Heritage: A vanishing yet essential resource for global development; Lecture Hall I	Towards a Global Metadata Interoperability Network: Experiences and Technologies for Linking Catalogues; Lecture Hall II	Early Career Data Professionals—today and future; Lecture Hall III	Development and sustainability of data services; Auditorium I
16.00-16.30	Poster Viewing Coffee/Tea Break			
16.30-17.00	Poster Madness Session: Auditorium I			
17.00-17.45	How Well is the Data Chain Working 1?: Auditorium I			
17.45-18.30	Keynote II: Theodora Bloom, British Medical Journal: Auditorium I			

研討會行程表（1）

Monday 3 November				
9.00-17.00	Registration: Convention Centre Foyer			
9.00-9.45	Keynote I: Paul Berkman (Co-chair Programme Committee): Auditorium I			
9.45-11.15	Interoperability, Standards and Integration of Multidisciplinary Data: Lecture Hall II	Open Data Initiatives: Lecture Hall III	Open Data Infrastructure and Best Practices Auditorium I	
11.15-11:45	Poster Viewing Coffee/Tea Break			
11.45-13.15	Data Rescue and Data at Risk: Lecture Hall I	Open Government Data and Smart Cities: Lecture Hall III	Materials Science Data: Auditorium I	
13.15-14.30	Lunch			
14.30-16.00	Our Scientific Heritage: A vanishing yet essential resource for global development: Lecture Hall I	Towards a Global Metadata Interoperability Network: Experiences and Technologies for Linking Catalogues: Lecture Hall II	Early Career Data Professionals—today and future: Lecture Hall III	Development and sustainability of data services: Auditorium I
16.00-16.30	Poster Viewing Coffee/Tea Break			
16.30-17.00	Poster Madness Session: Auditorium I			
17.00-17.45	How Well is the Data Chain Working 1?: Auditorium I			
17.45-18.30	Keynote II: Theodora Bloom, British Medical Journal: Auditorium I			
Tuesday 4 November				
9.00-17.00	Registration: Convention Centre Foyer			
9.00-9.45	Keynote III: Chaitanya Baru, Senior Advisor for Data Science, CISE Directorate, US National Science Foundation NSF: Auditorium I			
9.45-11.15	Data Infrastructures and Data Mining: Lecture Hall I	Building trust through certification of quality data services: Lecture Hall II	Data from the People for the People—encouraging Re-use and Collaboration I: Lecture Hall III	Belmont Forum: E-Infrastructure And Data Management: Auditorium I
11.15-11:45	Poster Viewing Coffee/Tea Break			
11.45-13.15	Development of Earth and Space Science Data Networks: Lecture Hall I	Sustaining data centers for sustainable research: Lecture Hall II	Data from the People for the People—encouraging Re-use and Collaboration II: Lecture Hall III	International Collaboration to Advance Data Science and Foster Collaboration: Auditorium I
13.15-14.30	Lunch			
14.30-15.15	Future Earth Keynote: Bina Agarwal, Future Earth Science Committee: Auditorium I			
15.15-16.00	Future Earth: quality and availability of data for global sustainability: Auditorium I			
16.00-16.30	Poster Viewing Coffee/Tea Break			
16.30-17.15	How Well is the Data Chain Working 2?: Auditorium I			
17.15-18.15	Prizes and Awards: Auditorium I			
18.30-	Conference Banquet: Jaypee Vasant Hotel			

研討會行程表（2）

Wednesday 5 November				
9.00-17.00	Registration: Convention Centre Foyer			
9.00-9.45	Keynote IV: Murali Mohan, Head of Big Data Initiative, Ministry of Science & Technology, India: Auditorium I			
9.45-11.15	Linkages for development and integration of population-specific data bases in healthcare: Lecture Hall I	Data for Development: Lecture Hall III		Citing Data to Facilitate Multidisciplinary Research: Auditorium I
11.15-11:45	Poster Viewing Coffee/Tea Break			
11.45-13.15	Linkages Workshop / Policies and Practices in Public Health and Food Sustainability: The Data Challenge: Lecture Hall I	Challenges in Developing Countries: Reports and Perspectives: Lecture Hall II	Challenges and Benefits of Open Science Data and International Data Sharing: Lecture Hall III	Publishing Data to Facilitate Multidisciplinary Research: Auditorium I
13.15-14.30	Lunch			
14.30-16.00	Challenges and Opportunities for Open Data in the Health Sciences: Lecture Hall I	Strategies towards Open Science: Lecture Hall III	Future Initiatives in Data Citation and Data Publication: Auditorium I	
16.00-16.30	Poster Viewing Coffee/Tea Break			
16.30-17.30	Closing Panel: How Well is the Data Chain Working 3?: Auditorium I			
17.30-18.00	Closing Ceremony: Auditorium I			

研討會行程表（3）