

# 2012 Digital Earth Workshop WS-3

## Novel Geospatial Information Architecture and Methods

### 創新的地理空間資訊架構與處理方法

---

#### Date& Time

2012-05-16/ 9:10-16:20

#### Venue /地點:

3F Computing Room, the Building for  
Humanities & Social Sciences, Academia  
Sinica

中央研究院人文館 3 樓遠距會議室暨電腦  
教室

#### Organized by/主辦單位



Center for GIS, RCHSS,  
Academia Sinica

中央研究院人社中心地理資訊  
科學研究專題中心



Digital Earth Research Center,  
Chinese Culture University  
中國文化大學數位地球研究中  
心

#### Background/緣起背景:

Linking and processing of large sets of geospatial data has becoming a critical issue faced by many researchers. This workshop aims to present geographic information science frontier research which included geospatial ontology, semantic link, and grid and cloud computing to address this issue. Our goal is to develop new methods and to provide effective computational services and resources to help researchers from different disciplines to use complex spatiotemporal datasets in their respective domains of research.

In this workshop, we will present the research approaching of geospatial research program in Geography Department, San Diego State University and Academia Sinica. We expect our research results to be materialized as services to other. Be it restructured geospatial datasets, Web APIs, or dependable computational resources, we plan to provide our results to others to use. We will demonstrate an ontology-based approach to unify spatiotemporal information and to integrate restructured geographic information datasets to other geospatial resources. We will also be building the underlining computational infrastructure to support the reuse of geospatial resources over the networks. Our goal is to provide to researchers of

dependable data services and computational infrastructures for large-scale geospatial and spatiotemporal information processing.

大量地理空間資訊的連結以及處理，是目前許多研究者所共同面臨的議題。本工作坊目的在於介紹目前地理資訊科學前沿研究，包括：網路知識本體、語意連結、雲端計算以及 e-Science 等資訊發展，而透過這些新技術整合研發處理多種來源、型態繁複、資料龐大的地理空間資訊的方法，並提供可供外界使用的服務介面與運算資源，有助於不同 GIS 應用環境的開發。

本工作坊中將展示美國聖地牙哥州立大學地理系以及中央研究院研究團隊的研究成果，無論是結構化的資料、可程式的介面、或是可依靠的計算資源，都能提供給本團隊以外的研究者使用，也尋求各種合作可能。嘗試面對複雜型態的地理資訊，提出以「知識本體」(Ontology) 為基礎的時空資訊組織架構，並以此設計為對外的「資訊網服務」(Web Services) 介面；另一方面，結合各種資源（資料、運算、儲存、頻寬等），也都能整合一起提供給外界使用，方便各學科領域的研究者開發應用程式與從事學術研究。

### Objectives/學習目標

This workshop contains one keynote speech and four topics, which aim to provide new prospects on GIScience research and technology. The topics of workshop are:

#### ■ Keynote Speech

Geospatial cyberinfrastructure is “a combination of distributed high performance geospatial computing resources, geo-referenced information, geo-processing web services, and geographic knowledge to facilitate the advancement of Geographic Information Science (GIScience), geospatial technology, and geographic education” (Zhang and Tsou 2009, p.605). Geospatial cyberinfrastructure provides valuable geospatial data, location-based information, and geographic knowledge for citizens, scientists and researchers. It also integrates or interconnects different types of web portals, mashups, and web mapping services.

This presentation will highlight the research challenges and opportunities for geospatial cyberinfrastructure in four domains: grid computing, cloud computing, web portals, and mashups. A few case studies (Spatial Demographic Simulation with Schelling’s segregation model, Metropolitan Area Pluralism Study (MAPS) website <http://geoinfo.sdsu.edu/maps/>) and ongoing research projects (NSF funded

CyberGIS project <http://cybergis.cigi.uiuc.edu/> and Mapping Cyberspace (<http://mappingideas.sdsu.edu/>) will be introduced in the presentation. The recent Big Data Initiative by the White House Office of Science and Technology Policy (<http://www.whitehouse.gov/blog/2012/03/29/big-data-big-deal>) and the research activities in the AAG Cyberinfrastructure Specialty Group (<http://www.cigi.illinois.edu/cisg/index.php>) will also be included in the presentation. One major challenge in geospatial cyberinfrastructure is to develop effective parallel computing algorithms for high performance GIS models. Most traditional GIS tools and programming libraries are created for stand-alone computers rather than parallel computing platforms. Current parallel processing GIS projects are mainly created for solving specific domain problems and these parallel algorithms are not reusable. By adopting grid computing and cloud computing frameworks, we may be able to develop more sustainable parallel algorithms for GIS models. Another key challenge is the semantic interoperability of GIS related web services and web mapping functions. We need to create comprehensive and sharable ontology databases, keywords,

and knowledge bases for the GIScience communities and transform geospatial web to intelligent “semantic” web. The possible resolution will require not only the standardization of domain keywords and concepts, but also the standardized translation methods between different domain knowledge bases.

The development of geospatial cyberinfrastructure is critical for the progress of GIScience and the innovation of geospatial technology. We hope that the development of geospatial cyberinfrastructure will facilitate a quantum leap of GIScience and enable more people to discover the value of geographic knowledge and geospatial technology.

### ■ **Topic1- An Ontology-based Approach for Handling Volunteered Geographic Information**

Volunteered Geographic Information (VGI) is a grassroots collection of geographic Web-content that is being produced increasingly by Web 2.0 applications. As VGI is regarded as a potential resource for institutional Geographic Information (GI), integrating it with institutional GI is becoming an important research topic in the field of GIScience. However, because of different data production methods, there exists a semantic gap between VGI and institutional GI. In this research, we propose an ontology-based information

formalization approach for handling this semantic gap, including five phases: Data collection, Structure Extraction, Semantic Formalization, Ontology Mapping & Integration, and Ontology Deployment. This approach can continually disambiguate implicit geographic semantic of VGI and bridge its semantic to GI's. This thereby facilitates the integration of VGI and institutional GI. As deploying this integration into Web 2.0 applications, users can have assistants from the applications and create the explicit VGI along with formal geographic knowledge. This research applied the ontology-based information formalization approach to extract place information from Flickr and to extract disaster relief information of March 11, 2011 Japan earthquake from Twitter.

### ■ **Topic 2- Ecological Data Management and Sharing Using Semantic Link: A New Approach**

Ecology is a discipline emphasizing integrative and collaborative approaches. In terms of data management and sharing it hasn't reached that state. In order to facilitate the integration of ecological data, we attempt to build a new approach by using Open Linked Data technique. The main purpose of this study is to demonstrate the difference of this approach from metadata and relational database methods of data management

and sharing adapted currently. The practice uses Resource Description Framework (RDF) to make data linkable and machine-readable in the finest level of granularity, interweaving data silos into the Web of Data. The new-proposed framework for sharing ecological data is so named Linked Open Data of Ecology (LODE). There are 5 datasets including forest fire database, plant and insect specimen databases, the Nanshi forest dynamics plot census and the catalogue of life in Taiwan tested in this practice.

### ■ **Topic 3- The Comparison between Cloud Computing and Grid Computing prospected on GIS Application**

Cloud Computing has become another buzzword after Web 2.0. However, there are dozens of different definitions for Cloud Computing and there seems to be no consensus on what a Cloud is. On the other hand, Cloud Computing is not a completely new concept; it has intricate connection to the relatively new but over 10-years established Grid Computing paradigm, and other relevant technologies such as utility computing, cluster computing, and distributed systems in general. This talk strives to compare and contrast Cloud Computing with Grid Computing from various angles and give insights into the essential characteristics of both. The talk will show the outlook of Cloud Computing

with Grid Computing in the future field of GIS application development.

### ■ **Topic 4- Data Management and Research Infrastructure**

Effective large-scale data analysis is much essential for wider scientific disciplines today. The distributed computing infrastructure (DCI) by Grid technology has been proved to be the most feasible solution for big amount data sharing and collaboration today. Reconfigurable services and separation of workflow with computing infrastructure accelerate the development of service-oriented sciences. Shared services and automation of research tasks can reduce the burden of big volume data processing. Taking advantage of emerging technologies such as virtualization and Cloud to offload the data analysis life cycle is the best strategy. Academia Sinica Grid Computing Centre (ASGC) acts as the only Worldwide LHC Computing Grid (WLCG) Tier-1 Centre in Asia and the e-Science Asia coordinator in European Grid Initiative. In this presentation, the current largest world wide Grid DCI as well as the e-Science applications and technology will be introduced. The data management infrastructure in the DCI is also explained in conjunction with some exemplar e-Science applications.



此研習會包含一個專題演講及四個主題，旨在介紹創新的地理資訊科學研究主題與技術發展，以下為個別講題：

- **以知識本體方法處理自願性地理資訊**  
自願性地理資訊是一個草根性的網路地理內容，其內容並不斷地經由 Web 2.0 的應用程式持續的增加，大量的自願性地理資訊被視為一個官方地理資訊的潛在資源，如何整合自願性地理資訊和官方地理資訊成為地理資訊科學的重要研究課題之一。然而，不同的資料生產方式，使得二種資訊之間存在語意的隔閡，研究中提出一個基於知識本體的資訊形式(正規)化方法來消除語意隔閡，橋接二個世界的地理資訊。這個方法包含 5 個步驟：(1)資料收集、(2)結構擷取、(3)語意形式(正規)化、(4)知識本體映對與整合、和(5)知識本體布署，這個方法可以持續地消除自願性地理資訊的語意模糊，並橋接自願性地理資訊的語意至官方的地理資訊，使得二者可以整合，若將這個混合正規和非正規地理知識的知識本體布署於 Web 2.0 的應用程式，使用者可以經由這個知識本體得到幫助，而使所生產的自願性地理資訊更為結構，易於整合於官方的地理資訊。本研究並應用這個知識本體方法處理以 Flickr 中的照片和標籤，以擷取地方知識以豐富地名資訊，以及應用於擷取推特(twitter)報導日本 311 大地震的訊息，以有系統地提供救災訊息。
- **生態資料管理與分享新嘗試：語意連結**  
生態學是一門重視領域合作與資料分享整合的科學，但傳統的生態研究並未

達到這樣的境界。嘗試以鍵連資料技術，以實作的方式，將現有生態資料庫加以語意連結。研究主要目的在展示鍵連資料技術有別於現行元數據或關連式資料庫之生態資料管理與分享框架，以達成較理想的生態資料管理、整合與分享需求。使用了森林火災、植物標本、昆蟲採集、森林動態樣區調查以及台灣的物種名錄等五個資料庫，以資源描述框架作為資料標準連結資料，讓原本分散難以分享的資料在最細的單元彼此連結，形成有別於傳統網路的資料管理與分享方式，並讓機器協助資料處理作業，跨越了系統的界線的連結限制。

- **雲端計算與網格計算比較以及在 GIS 應用展望**

雲端計算已經成為 Web 2.0 之後網路流行語。不過，雲端計算卻有幾十種不同的定義，缺乏共識。另一方面，雲端計算不算是一個全新的概念，它與其他高效能計算技術，例如網格計算、公共計算、集群計算等分散式系統都有關聯性。本研究將針對雲端計算與網格計算進行多面向比較，並和回顧以及展望這些技術在 GIS 領用的應用。

- **資料管理與研究基礎環境**

如何有效地進行的大規模數據分析是今日各學科基本問題。基於網格技術分佈式計算基礎設施 (DCI)，已被證明是目前大量的數據共享和協作最可行的解決方案。重構服務和計算基礎設施工作流程分離，可以加快服務導向應用開發。共享服務和科學計算自動化，可以降低大量數據處理負擔。藉由新興技術，如虛擬化和雲端計算，將數據分析

的生命週期進行封裝是最好的策略。中科院網格計算中心 (ASGC) 作為全球唯一的大型粒子對撞機計算網格 (WLCG) 的 Tier-1 中心，以及亞洲 e-Science 和歐盟 Grid 的核心機構。在這專講中，將詳盡介紹分佈式計算基礎設施，以及在 e-Science 的應用。

### Speakers/講者介紹



Prof. Ming-Hsiang Tsou

■ Professor Department of Geography,  
San Diego State University, U.S.A.  
Bioneers 美國聖地牙哥州立大學地理  
學系教授

Dr. Tsou is Professor in the Department of Geography, San Diego State University and served on the editorial boards of the *Annals of GIS* (2008-) and the *Professional Geographers* (2011-). He is co-author of the book, *Internet GIS*, published by Wiley in 2003, the co-chair of the NASA Earth Science Enterprise Data System Working Group (ESEDWG) Standard Process Group (SPG) from 2004 to 2007, the 2007-2008 Chair of the Cartographic Specialty Group, and the 2011-2012 Chair of the Cyberinfrastructure Specialty Group in the Association of American Geographers (AAG). Tsou was appointed by the National Academy of Science in 2006 to serve on the committee on "Research Priorities for the USGS Center of Excellence for Geospatial Information Science". In 2007, he created and maintained an interactive Web-based mapping services

for San Diego Wildfires 2007 (<http://map.sdsu.edu>) and his efforts have been recognized by the AAG and the San Diego Union Tribune (newspaper). In 2008, Tsou served as a senior researcher in the NSF-funded GeoTech Center (<http://www.geotechcenter.org/>) for promoting GIS education in community colleges. In 2010, Tsou served as the Principle Investigator (PI) of the NSF-CDI project, "Mapping ideas from Cyberspace to Realspace", funded by National Science Foundation. This NSF project integrates geographic information systems, computational linguistics, web search engines, and social media to track and analyze public-accessible websites and tweets in cyberspace.

Dr. Tsou 'r Research Interests are Mapping Cyberspace, Web GIS applications, Internet mapping, Mobile GIS, and Cyberinfrastructure with grid computing and cloud computing.





Mr. Dongpo Deng(鄧東波)

- PhD Candidate, International Institute for Geoinformation Science and Earth Observation University of Twente, The Netherlands 荷蘭屯特大學地理資訊科學與地球觀測學院博士候選人
- Research Assistant for Institute of Information Science, Academia Sinica. 中央研究院資訊科學研究所研究助理

Mr. Deng's research topic mainly focuses on semantically integration of user-generated data and institutional geospatial data via geospatial ontology. Over past five years, he also dedicated on open geospatial technologies of OGC such as GML (Geography Markup Language), WMS (Web Map Service), WFS (Web Feature Service) and GeoSPARQL. Moreover, he also concentrates to the development of Open Source GIS. He and some friends of his collaboratively founded a Taiwan chapter of OSGeo in 2007. This local foundation not only concerns open source GIS but also the issues of open geospatial data. Mr. Deng is also an activities mappers of OpenStreetMap and involves many efforts in

Taiwan OSM.

鄧東波先生在地理資訊研究與處理有 8 年以上的工作經驗，主要專長在於地理本體及語意(Geospatial ontology and Semantics)；地理知識發掘與資料探掘 (Geospatial knowledge discovery and data mining);使用者生產地理內容/自願性地理資訊的處理與分析(Process and analysis of User-Generation Geo-Content/Voluntered Geographic Information)；和開放地理資訊協會的標準使用，如 GML、WMS、WFS 和 GeoSPARQL (Use of Open Geospatial Consortium(OGC) standards, e.g. GML, WMS, WFS, GeoSPARQL)。在 2007 年，他與幾位同好共同地在台灣發起「地理資訊開放源碼基金會」台灣分會 (Open Source Geospatial Foundation, Taiwan chapter)，並於近年極積投入「開放街圖」(OpenStreetMap, OSM) 社群與繪圖的工作。



Mr. Guan-Shuo MAI(麥館碩)

生物多樣性資料之分享及整合。他在以鍵連資料技術將現有生態資料庫加以語意連結方面研究上有很豐富的經驗。

- Research Assistant of Biodiversity Research Center, Academia Sinica.  
中央研究院生物多樣性研究中心研究助理
- Bachelor of Computer Science, National Tsing Hua University.  
國立清華大學資訊工程學系學士
- Master of Science in Natural Resources and Environmental Studies, National Dong Hua University.  
國立東華大學自然資源與環境學系碩士

Mr. Mai is the research assistant of Biodiversity Research Center, Academia Sinica. His job is to apply the technology of semantic web and linked data for ecological and biodiversity data sharing and integration. He has A wealth of practice experience using Resource Description Framework (RDF) to make data linkable and machine-readable

麥館碩先生現任中央研究院生物多樣性研究中心研究助理，工作內容與研究主題為應用語意網與鍵連資料技術協助生態與



Mr. Hsiung-Ming Liao(廖泫銘)

- Assistant Research Specialist, , Research Center for Humanities and Social Sciences, Academia Sinica 中央研究院人文社會科學研究中心研究助技師

Mr. Liao is assistant research specialist at the Research Center for Humanities and Social Sciences, Academia Sinica. His research interests are GIS, Remote Sensing, Digital Archive and Computer Science. He has involved the ‘Taiwan e-Learning and Digital Archives Program’ over 10 years and has many experiences in the digitization and application of maps and remote imagery.

廖泫銘先生現任中央研究院人社中心地理資訊科學研究專題中心研究助技師，曾任中研院計算中心 GIS 組組長，從事地理資訊系統相關工作已超過十年，中央研究院任職期間參與本院多項 GIS 專案工作。主要專長包括地理資訊系統、遙感探測及全球定位系統，目前主要工作執行『數位典藏與數位學習國家型科技計畫-地圖與遙測影像數位典藏計畫』、維護及建置中國及臺灣歷史文化地理資訊系統、推廣地理資訊系統於人文社會研究之應用，並從事 GIS 先進技術的

研發及應用研究，近期則在 Google Earth 應用、街景測量車、360 度環景攝影、UAV 與 3D GIS 等技術議題發表多篇技術報告。



Mr. Eric Yen(嚴漢偉)

- Associate Research Engineer , Research Center for Information Technology Innovation, Academia Sinica 中央研究院資訊創新研究中心研究副技師

Mr. Eric Yen is the associate research scientist of Academia Sinica Grid Computing Centre with research focus on distributed computing infrastructure, geospatial information system, and Grid and Cloud technology.

嚴漢偉先生現任中央研究院資訊創新研究中心網格計算專題研究中心研究副技師，他的研究專嘗試分散式計算環境、地理資訊系統、雲端計算與網格計算。

## 2012 Digital Earth Workshop WS-3

### Program/行程表

2012 Digital Earth Workshop WS-3	
Time	Program
09:10~09:30	Registration
09:30~10:50	Keynote: <b>Research Challenges and Opportunities for Geospatial Cyberinfrastructure and Mapping Cyberspace</b> Speaker: Professor Ming-Hsiang Tsou(鄒明祥), Dept. of Geography, San Diego State University Chair: Dr. Tyng-Ruey Chuang(莊庭瑞), 中研院資訊所
10:50~11:10	Coffee Break
11:10~12:00	Topic: <b>An Ontology-based Approach for Handling Volunteered Geographic Information</b> Speaker: Mr. Dongpo Deng(鄧東波), 中研院資訊所
12:00~13:30	Lunch Time
13:30~14:20	Topic: <b>Ecological Data Management and Sharing Using Semantic Link: A New Approach</b> Speaker: Mr. Kuan-Shuo Mai(麥館碩), 中研院生多中心
14:20~15:10	Topic: <b>The Comparison between Cloud Computing and Grid Computing prospected on GIS Application</b> Speaker: Mr. Hsiung-Ming Liao(廖滋銘), 中研院人社中心 GIS 專題中心
15:10~15:30	Coffee Break
15:30~16:20	Topic: <b>Data Management and Research Infrastructure</b> Speaker: Mr. Eric Yen(嚴漢偉), 中研院資創中心