



# PICO start for Research Bazaar

Malcolm Atkinson  
Data-Intensive Research Group  
University of Edinburgh



OSDC Workshop, Informatics, Edinburgh, 18 June 2013

# Plan

- Phase 1
  - Each speaker gets 3 minutes
  - And three slides
- Phase 2
  - Detailed follow up discussions
  - Choose topics/locations
  - Replay to allow everyone to get to two discussions

# Batting order

1. Josh Eisenberg    The Domination Data Structure
2. Isao Kojima                      Lavatube & GEO Grid
3. Christine Harvey                                      MEG
4. Michael J. Lewis    Framework for Workflows
5. Warren Cole                      Retroviral links to Cancer
6. Pedro Bello-Maldonado    UvA Data Service

# The Domination of Three Dimensional Chessboards by Bishops

Presentation by: Joshua Eisenberg



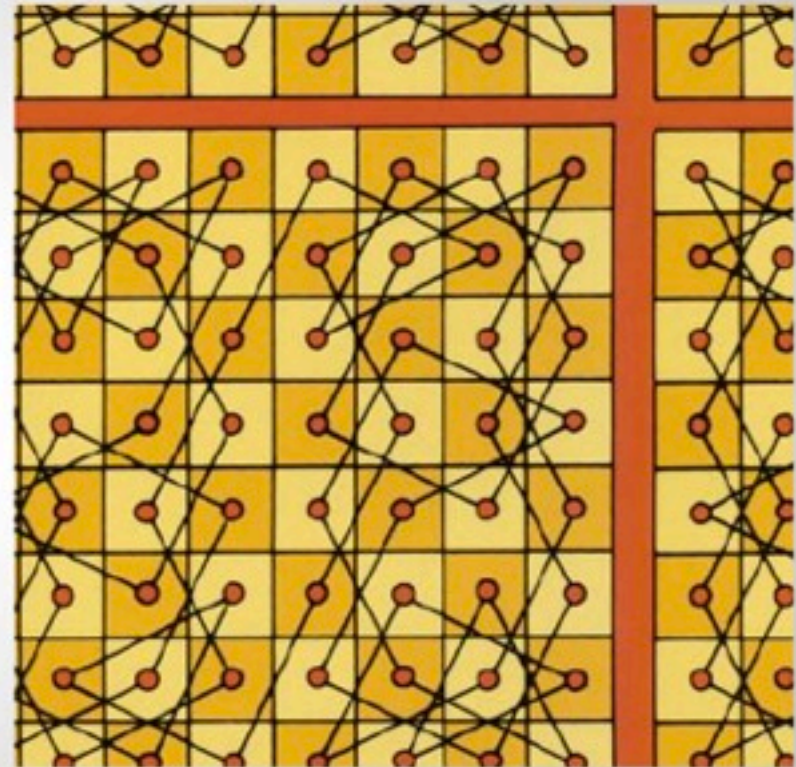
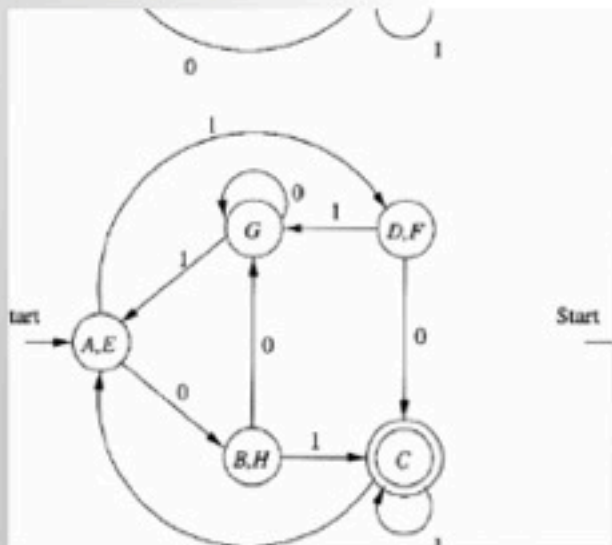
# Graph Theory G

## Introductory I

Chessboard domination is a problem within the area of graph theory

Graph theory is about the relationship of vertices that are connected by edges

- It is the foundation for the theory of computation, for example: Turing machines, Finite Automata



Note: To the left is a discrete finite automata and above is a traversal of an 8x8 chessboard by a knight

# Graph Theory Continued

- How to to depict chessboards as graphs
  - Every square is a vertex
  - Edges are generated by the paths a particular chess piece travels from one square to another



# Isao Kojima

isao.kojima@aist.go.jp

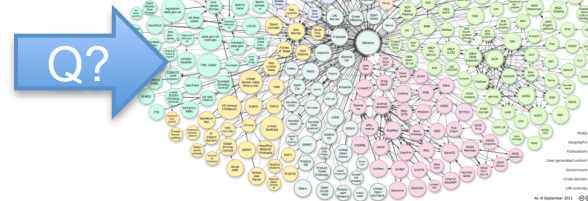


- **Research:** Group Leader of
  - Service-ware Research Group(Research everything as a Service 😊)
    - **ITRI**(Information Technology Research Institute)
      - **AIST**(National Institute of Advanced Industrial Science and Technology)
        - » **Japan**
    - Working with 8 researchers, 2 post-docs and 3-5 engineers
- **Standard Activity**
  - Member of OGC(Open Geospatial Consortium) Catalog WG
  - Chair of OGF(Open Grid Forum) DAIS-WG

# My current projects

## Data Infrastructure R&D

- **GEO Grid Project: (Geospatial Grid/Cloud)**
  - Heterogeneous metadata catalog
  - Workflow engine for geospatial objects
    - Lavatube (will be presented later)
  - Machine Learning of images for Landcover analysis
- **Linked Data Project**
  - LOD query processing
  - Open government data
- **Bioinformatics Project**
  - High Performance & Scalable de-novo Assembling & workflows



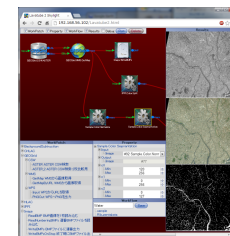
Some detail will be presented on Thursday



# My concern for this session

## Search for some possible joint work for workflow

- Our workflow Lavatube is special-purpose
  - Pros:
    - Real-time various video/image processing
    - Highly-interactive
    - Support of rest based OGC standard service interfaces
  - Cons:
    - No scalability
    - No standard workflow language interface

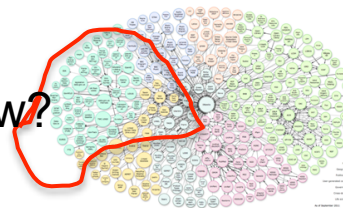


OSDC/PIRE  
student topic

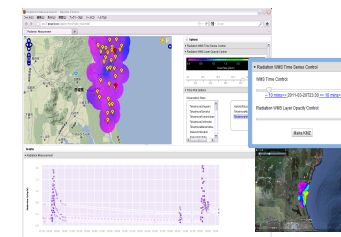
## Share/Exchange the information about:

- GEO Grid related R&D and Applications:
  - Applications
    - Radiation Monitoring
    - Biodiversity
- Linked Data R&D:
  - Open Government Data
- Bioinformatics related R&D:
  - Especially for NGS(next generation sequencer)

How?



GEO Grid Radiation Monitoring  
after 3-11



# The MITRE Elastic Goal-Directed Simulation Framework (MEG)

---

Christine Harvey  
The MITRE Corporation

18 June 2013  
OSDC Workshop  
Edinburgh, UK

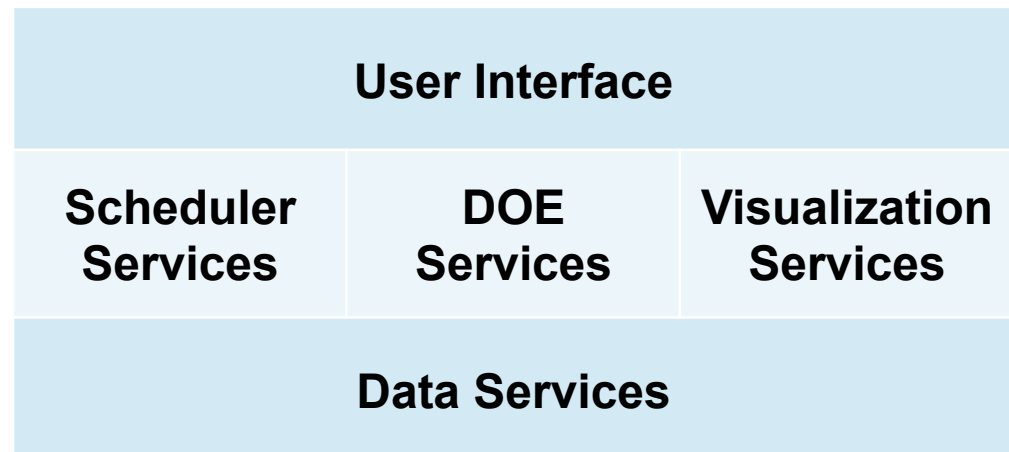
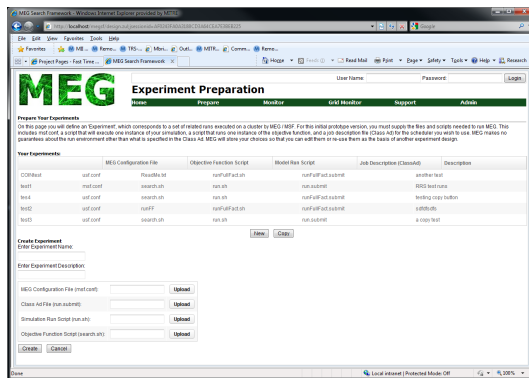
**MITRE**

---

© 2013 The MITRE Corporation. All rights reserved.

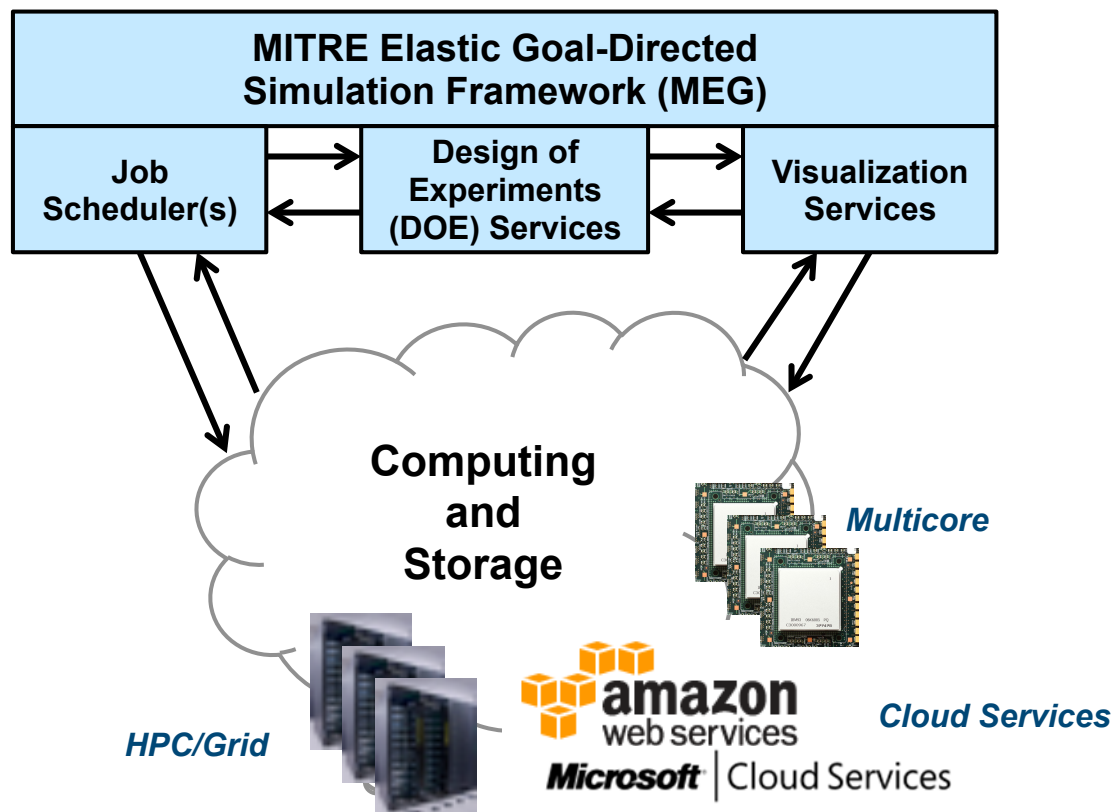
# The MITRE Elastic Goal-Directed Simulation Framework (MEG)

- **Middleware framework to supplement existing simulation applications**
- **Provides access to three capabilities:**
  - Cloud-based or grid-based computing resources
  - Advanced Design of Experiments (DOE) methodologies
  - Robust data processing and visualization tools



# The MITRE Elastic Goal-Directed Simulation Framework (MEG)

- Run experiments on multiple grids
- Simulation Adaption:
  - Command line input
  - No hard coded paths



**The MITRE Elastic Goal-Directed Simulation Framework (MEG) is first and foremost an engineering activity – our goal is to develop a practical, useful tool.**

# Michael J. Lewis

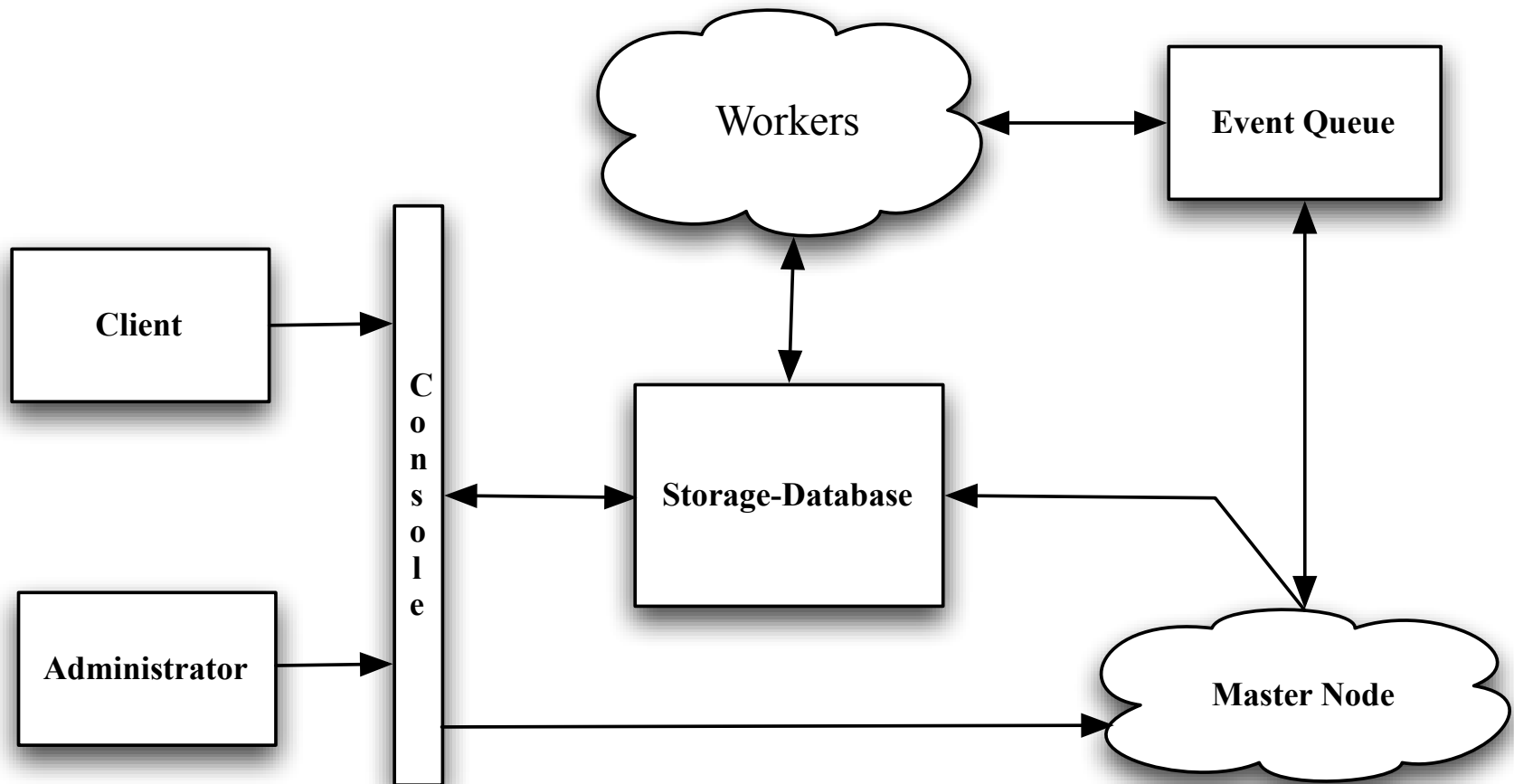
## (University of Illinois At Chicago - UIC)

- Ph.D. Candidate at UIC.
- Area of interest: Cloud Computing & Large Scale Distributed Systems.
- Current research: Classification of large scale distributed workflows.

# Worker Role

- Research goals:
  - To define workflow characteristics.
  - Create a customizable framework for workflows.
  - To experiment with different workflows on different systems.

# Distributed workflow



# Retroviral Links to Cancer

Warren Cole

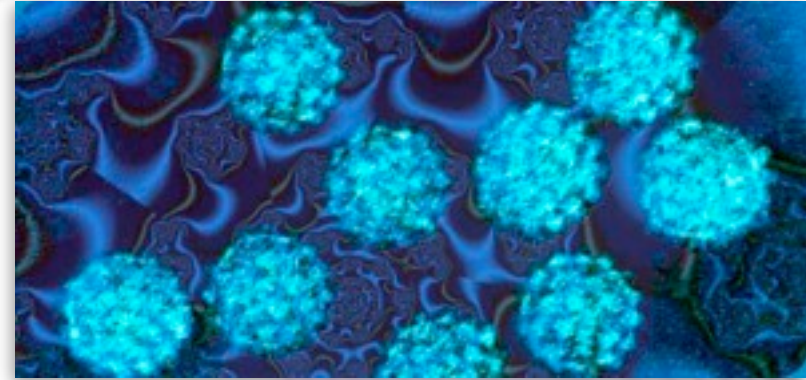
University of North Carolina Charlotte



UNIVERSITY OF NORTH CAROLINA



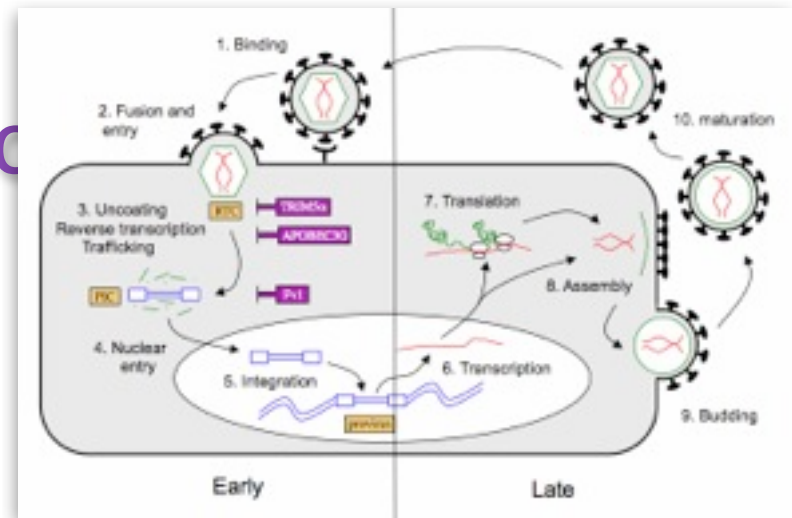
# Retroviral Links to Cancer



## • Research Questions

- 1) Is there evidence of HPV in Cervical Cancer cells?
- 2) Are retroviruses generally represented in cancer cells
- 3) Are retroviruses carrying regulatory signals(microRNA) that could be disrupting regulatory networks within cell resulting in cancer cells

# Retroviral Links to Cancer Research Methods



- OSDC Bionimbus Database
- Search Cervical Cancer cell data within Bionimbus database for evidence of HPV
- Search the Retroviral database against the Bionimbus database for evidence of retroviral expression within cancer cells
- Search for retrovirus' with incorporated micorRNA within cancer cells.

# UvA Data Service:

## Towards an Unified Architecture for Scientific Data Aggregation

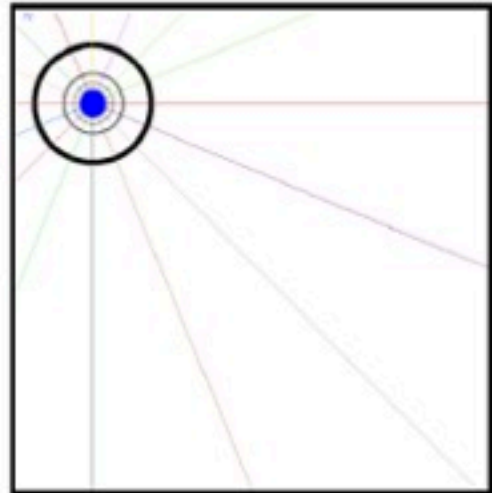
Pedro D. Bello-Maldonado, Ana Opreescu, Paola Grosso



# Problem – Description

- ❑ Data generated by experiments
- ❑ Insufficient tools for data content discovery
- ❑ Different database paradigms -> Insufficient tools for data aggregation

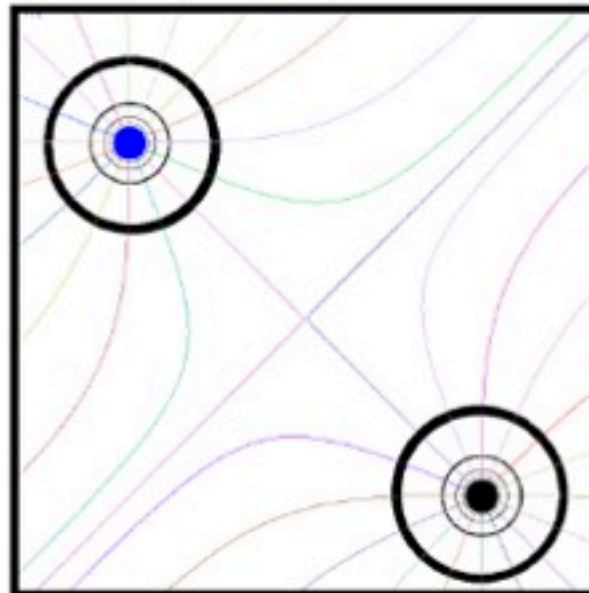
# Problem – Example



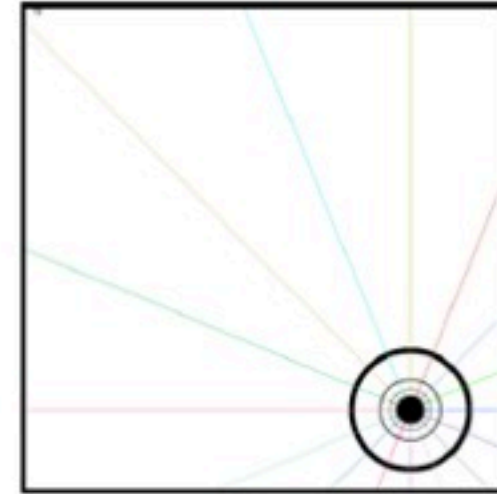
Event A

Has information  
of Event A but  
not Event B

Has information of  
Event A and Event B



Event A + B



Event B

Has information  
of Event B but  
not Event A

# Phase 2 This Room

1. Pedro Bello-Maldonado      UvA Data Service
2. Michael J. Lewis      Framework for Workflows
3. Isao Kojima      Lavatube & GEO Grid

**20 minutes discussion & 10 minutes change over time**

# Phase 2 Room 4.02

1. Warren Cole      Retroviral links to Cancer
2. Christine Harvey      MEG
3. Josh Eisenberg      The Domination Data Structure

**20 minutes discussion & 10 minutes change over time**

Thank you & Questions