

Modelling the Scientific Workforce: A Progress Report

Dowman P Varn,^{1,2} Katy Börner,¹ Robert P Light,¹
Scott B Weingart,¹ & James P Crutchfield²

*¹Cyberinfrastructure for Network Science Center
School of Library & Information Science
Indiana University*

*²Department of Physics & Complexity Sciences Center
University of California, Davis*

12 April 2012



Questions ...

- What roles do gender and ethnicity play in the evolution of a successful career?
- What is the effect of geolocation on a successful career?
- What are the effects of collaboration and competition on career success?
- What are the most successful scholarly areas and career trajectories across the career landscape?
- How do US scholars react to exogenous shocks?
- Are some institutions sinks or sources for scholars?

Research Efforts

1. First Principles Definitions of:
 - (a) Scholars;
 - (b) Careers; &
 - (c) Events.
2. Data Acquisition & Analysis:
 - (a) NSF-SDR; &
 - (b) Web of Science.
3. Modelling the Scholarly Workforce:
 - (a) Multi-Agent Dynamical Systems.

(1.a) Scholars

Colloquially, a scholar, θ , is someone who engages in the publication of scholarly research, or has a scholarly education.

- Autonomous Agents;
- Perform Certain Actions -- e.g. publish papers, work at institutions;
- Have Attributes -- e.g., gender, ethnicity, geolocation;
- Can Store and Process Information -- remember other researchers and past actions; &
- Can Pursue Goals -- seek recognition from peers.

(I.b) Events

Let an event, ω , be some thing that happens to an agent or its environment, or that the agent does to itself or environment.

- Can Be Active or Passive;
- Must Have a Particular Start Date;
- Can Be of Indefinite Duration; &
- Can Be Endogenous or Exogenous.

(1.c) Career

Let us define a career, Ψ , as the time ordered sequence of events that spans the active life of the scholar. This is just a record of what each scholar did, much like one would find on an academic curriculum vitae.

We can express this as a series of events, $\omega_1, \omega_2, \dots, \omega_L$.

How much can we infer from data?

(2) Data Acquisition & Analysis

We are considering two data sets:

- (a) The National Science Foundation Survey of Doctoral Recipients (NSF-SDR); &
- (b) Web of Science (WoS)

(3) Modelling the Scholarly Activity

(a) Multi-Agent Dynamical Systems (MADS)

Multi-Agent Dynamical Systems

- Agent-Based Modelling
- Game Theory
- Reinforcement Learning
- Nonlinear Dynamics

Agent-Based Modelling

- Agents (Scholars) have attributes and can perform actions

Game Theory

- The reward that an agent receives for an action depends not only on his choice of actions, but also on that of the other agents.

Agent B

		C	D
Agent A	C	(3,3)	(0,5)
	D	(5,0)	(1,1)

C: Cooperate
D: Defect

Reinforcement Learning

- Adjusts choice of actions depending on results of previous actions.
- Competition between “exploitation” and “exploration”
- Two parameter (α and β) model

Nonlinear Dynamics

- Put all this together into a set of coupled, nonlinear, first-order differential equations.
- Have the entire theory of nonlinear dynamics at our disposal.