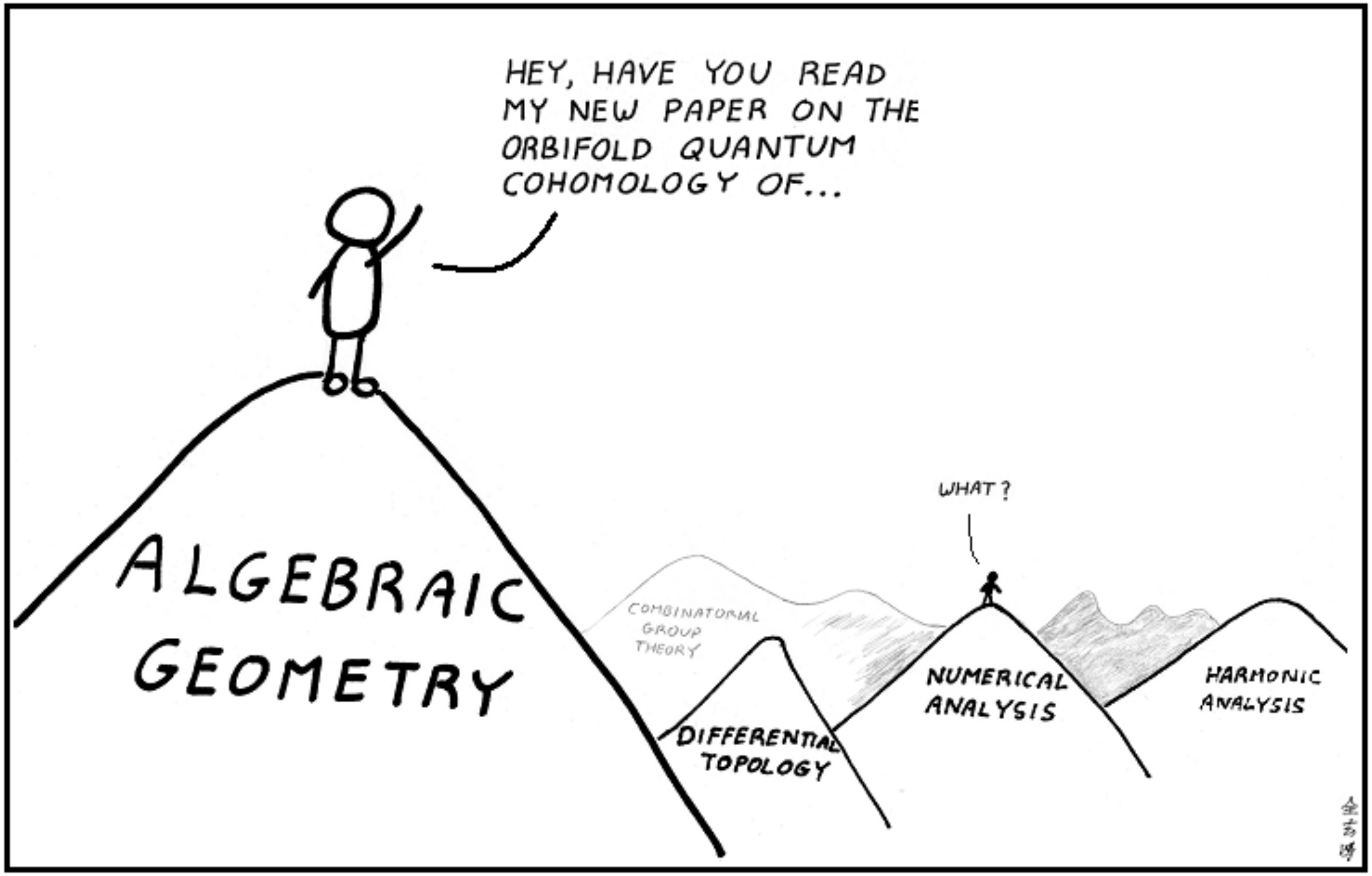
The jargon barriers of science



The Landscape of Modern Mathematics

Jevin West, Information School, University of Washington

Sante Fe Institute Speaker Series, March 17, 2016

The Center for Nonlinear Studies at Los Alamos National Laboratory is pleased to announce its 23rd Annual Conference on

NETWORKS: STRUCTURE, DYNAMICS AND FUNCTION

May 12 - 16, 2003, Hotel La Fonda, Santa Fe, New Mexico, USA

http://cnls.lanl.gov/networks , e-mail: networks@cnls.lanl.gov

Our world is a complex mesh of interacting elements, both natural and man-made. Recent observations suggest that the formation of such complex networks is not random, but rather follows fundamental organizing principles. *The 2003 CNLS Annual Conference* focuses on the search for underlying principles in the structure, dynamics, and function of complex networks. The conference will facilitate cross-disciplinary interactions by bringing together researchers from a diverse set of fields. The emphasis will be analysis of real-world data from information networks (internet, www, data networks), biological networks (in proteomics, gene networks, metabolic networks), social networks (including epidemiological networks) and infrastructure networks (power grid, transportation networks).

Sponsors: Center for Nonlinear Studies, B Division, P Division, T Division, LANSCE, Los Alamos National Laboratory

LIST OF INVITED SPEAKERS

Lada A. Adamic [Hewlett-Packard]	**Shlomo Havlin** [Bar-Ilan]	**Christos Papadimitriou** [Berkeley]

The Game of Leaf

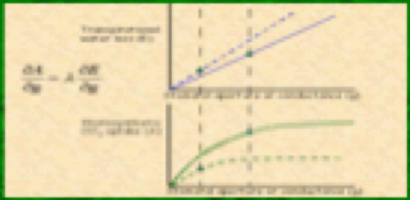
Evidence that Stomatal Networks are Cellular Computers

Jevin West, David Peak, Keith Mott, Susanna Messinger
Utah State University

PLANT'S DILEMMA

- During photosynthesis a plant incorporates CO_2 from and loses H_2O to the atmosphere
- This dilemma can be framed as a **constrained optimization** problem
- Stomata are the hardware the plant uses to resolve this dilemma

CONSTRAINED OPTIMIZATION



How does the plant do it?

STOMATA

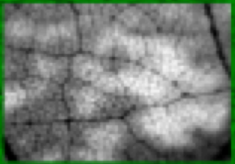
- Tiny pores on the surface of a leaf
- Control H_2O and CO_2 exchange between leaf and atmosphere

•Responsible for 99% of terrestrial carbon fixation and 90% of terrestrial water loss

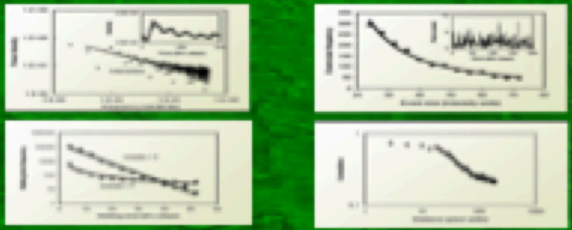
•Aperture size varies in response to light, CO_2, and humidity

STOMATAL NETWORKS

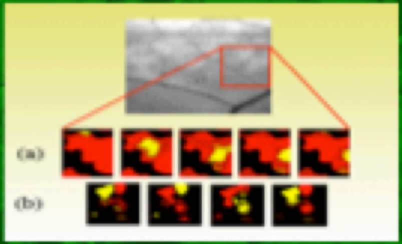
- Stomata interact through short range hydraulic and chemical signals—stomata form a **locally connected network**
- These connected networks show spatially coordinated behavior that change in time
- 512 by 512 grayscale image of chlorophyll fluorescence containing ~10^5 stomata—areas with open stomata are dark and areas with closed stomata are bright



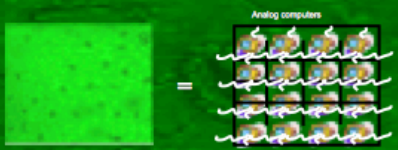
STOMATAL DYNAMICS IS COMPLEX (persistent correlations in time and space)



AND HAS GLIDERS



ARE STOMATAL NETWORKS CELLULAR COMPUTERS?



CELLULAR COMPUTERS

•Network of locally-connected processing information units that perform system wide computation—**emergent, distributed computation**

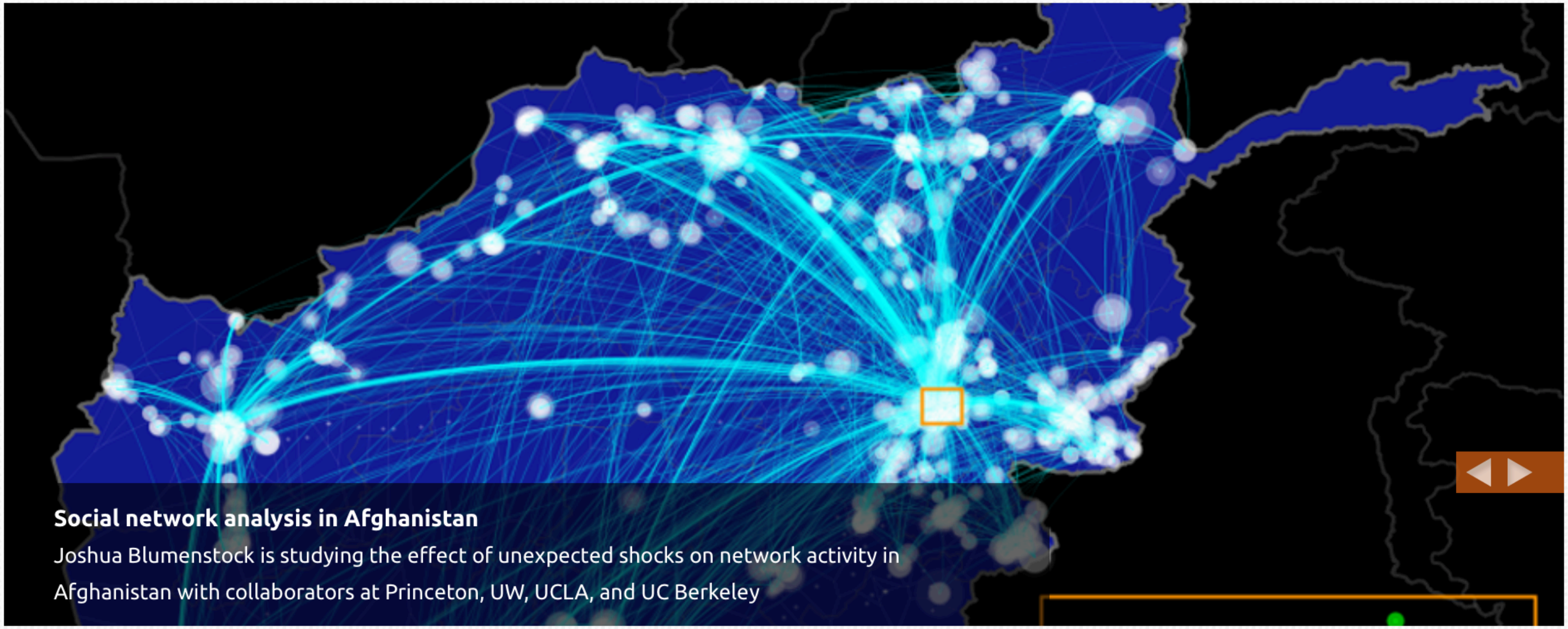
•Example: Density Classifier Automaton

•Dynamics is **complex** and there are particles of information—**gliders**

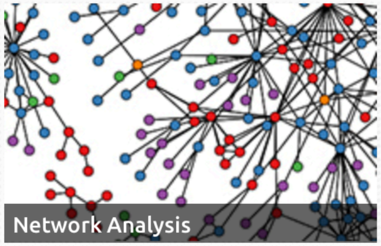

Santa Fe Institute
Cowan Campus
1399

W DataLab
Data Science and Analytics Lab

University of Washington Information School



Research Focus Areas



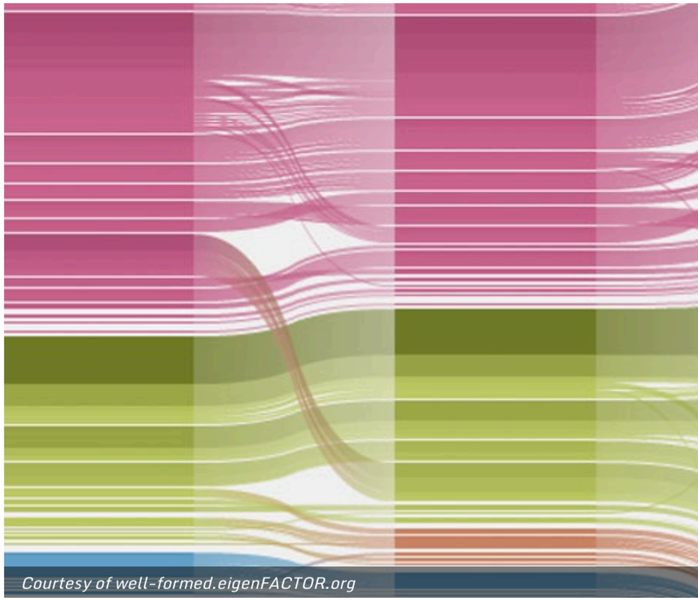
News and Updates

28 **Blumenstock at Population Association of America**

What we do

The DataLab is the nexus for research on Data Science and Analytics at the UW iSchool. We study **large-scale, heterogeneous human data** in an

GORDON AND BETTY MOORE FOUNDATION



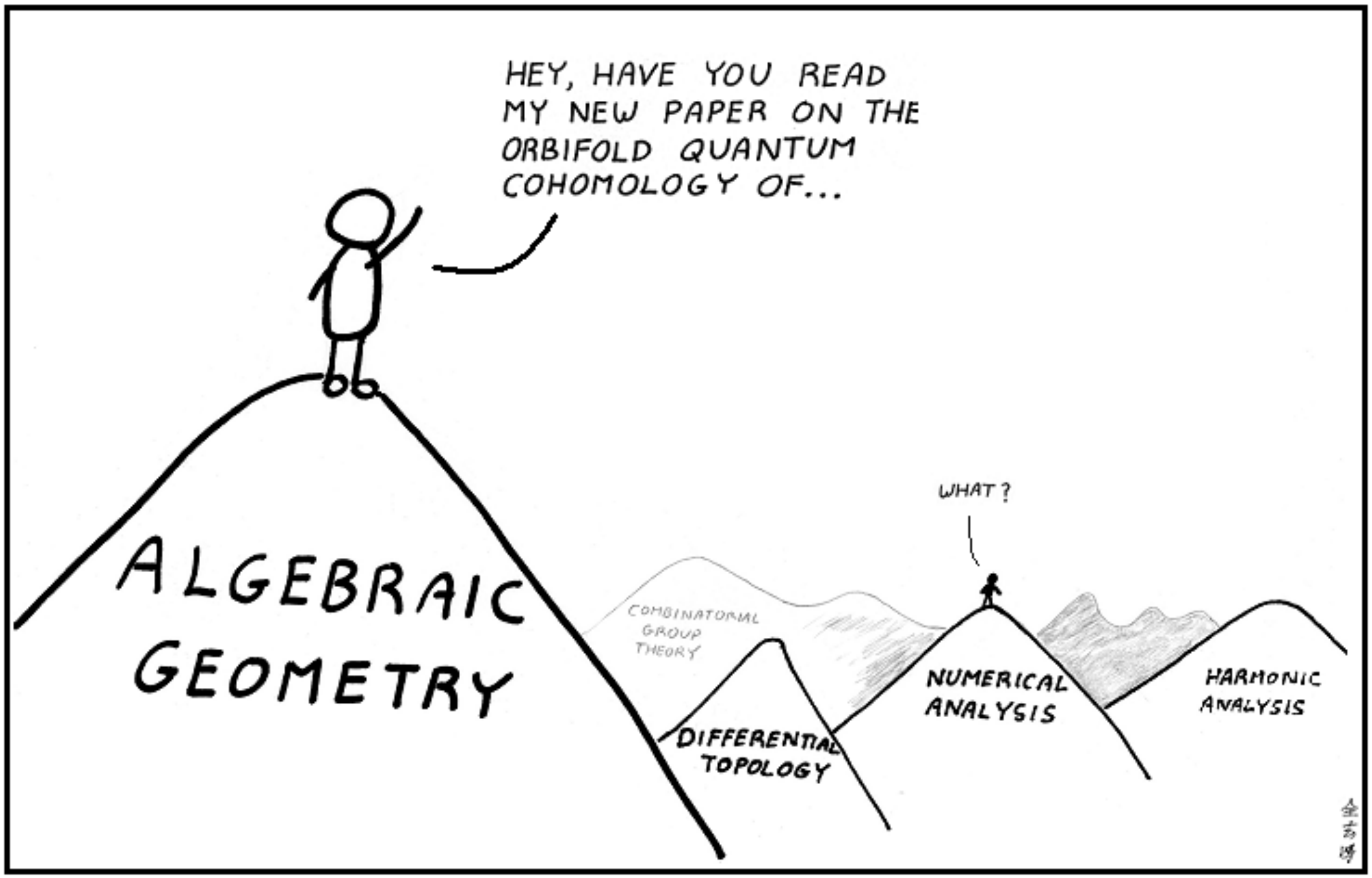
Courtesy of well-formed.eigenFACTOR.org

DATA-DRIVEN DISCOVERY

Data Science

Environments

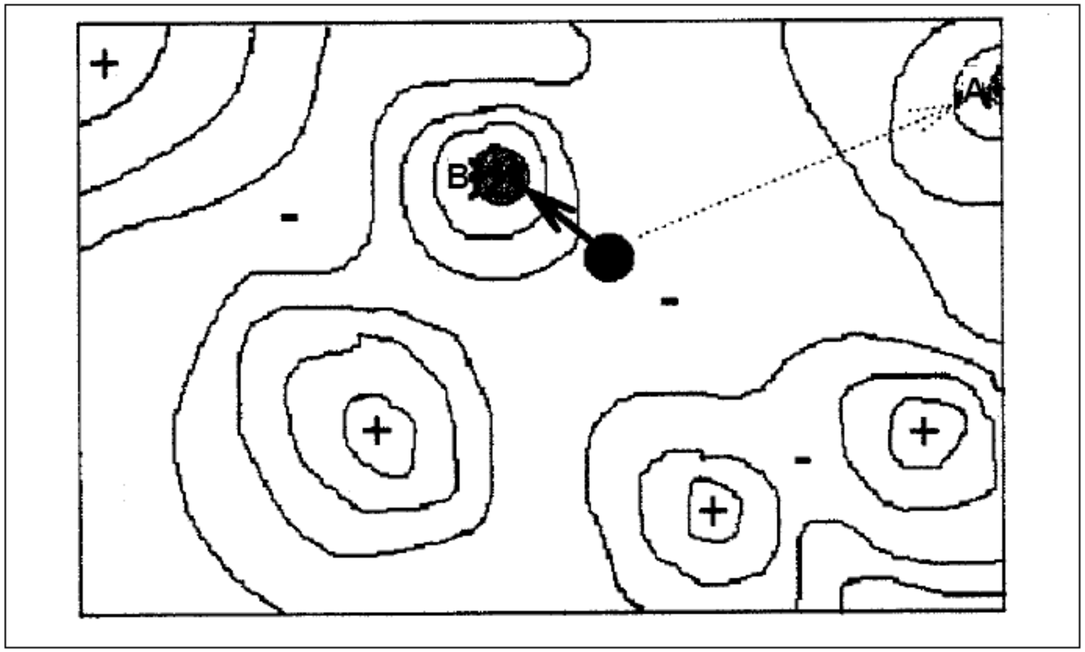
The jargon barriers of science



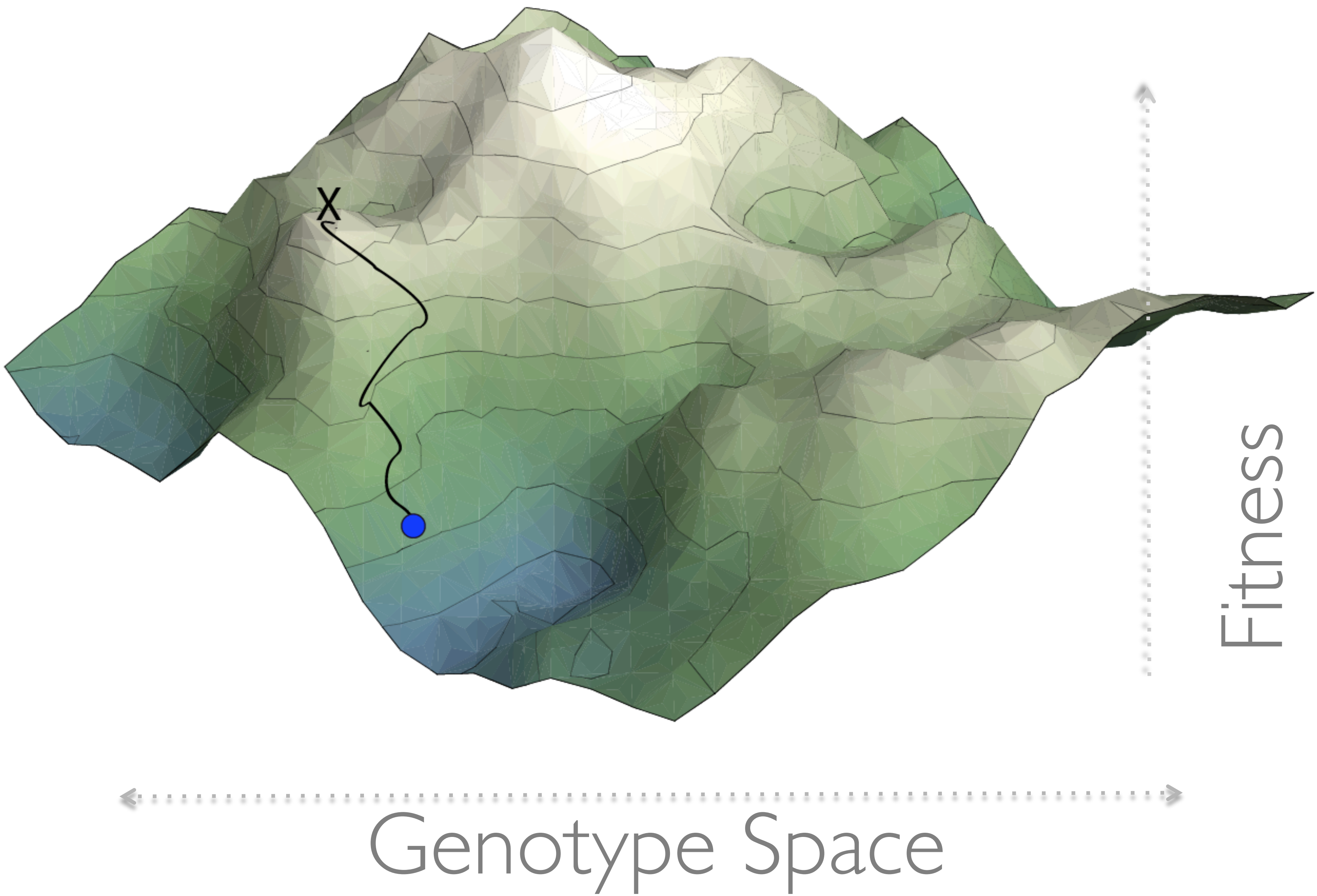
The Landscape of Modern Mathematics

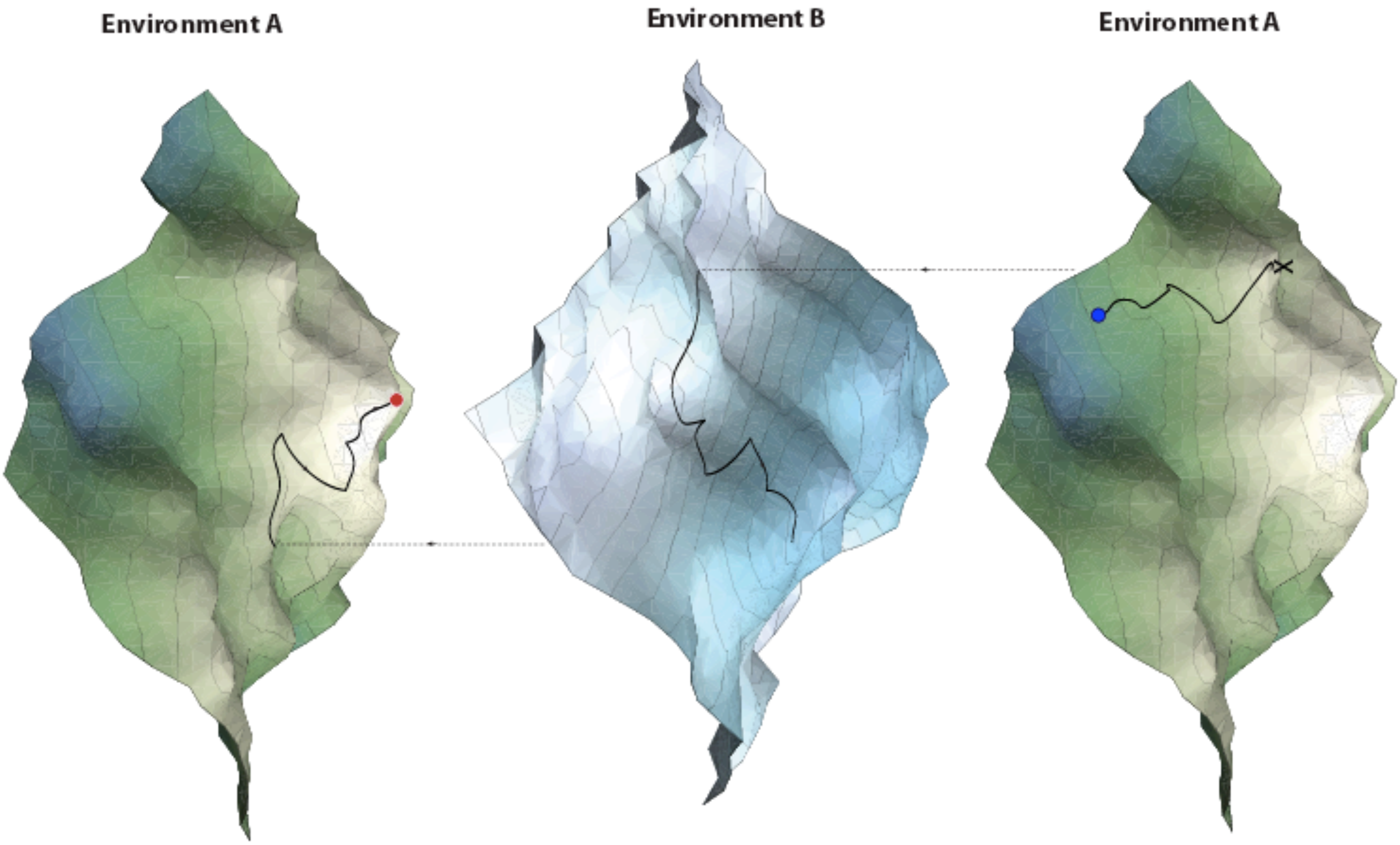
Jevin West, Information School, University of Washington

Fitness Landscapes

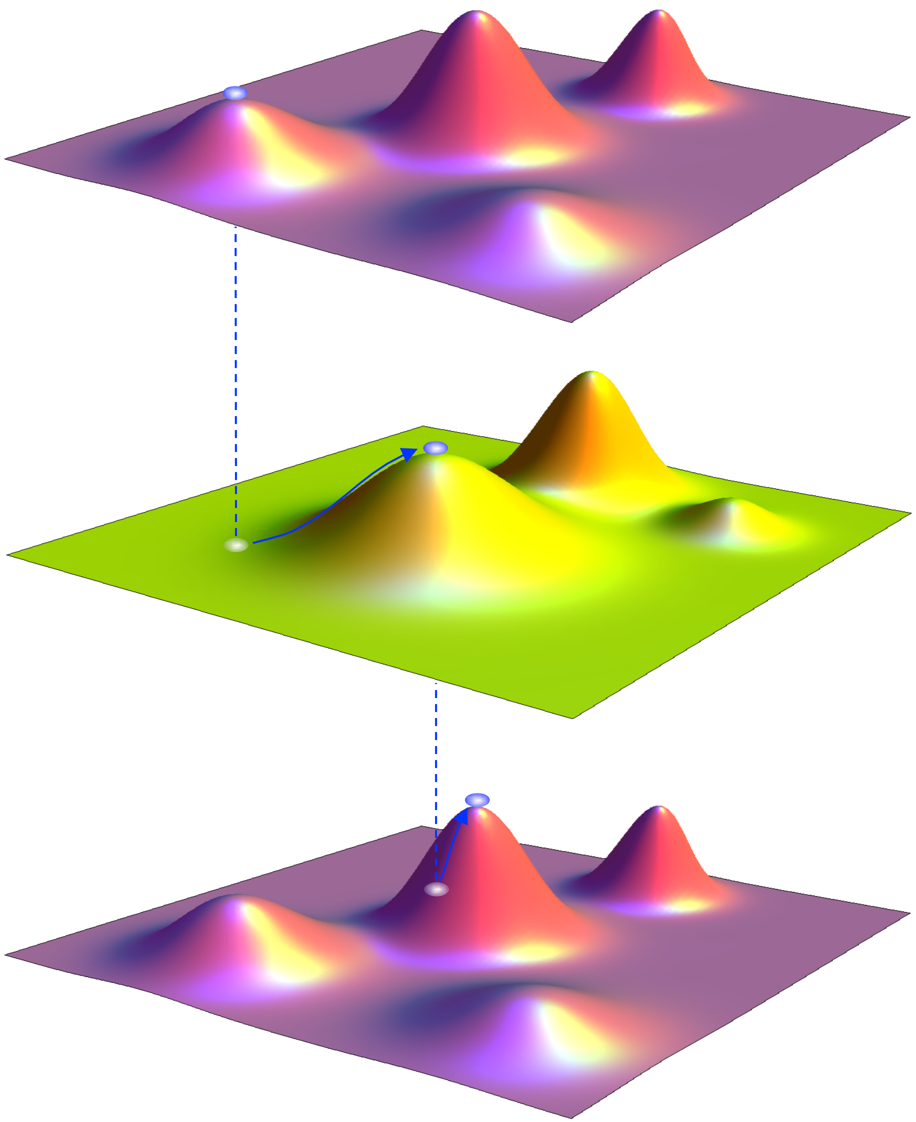


Wright, 1932

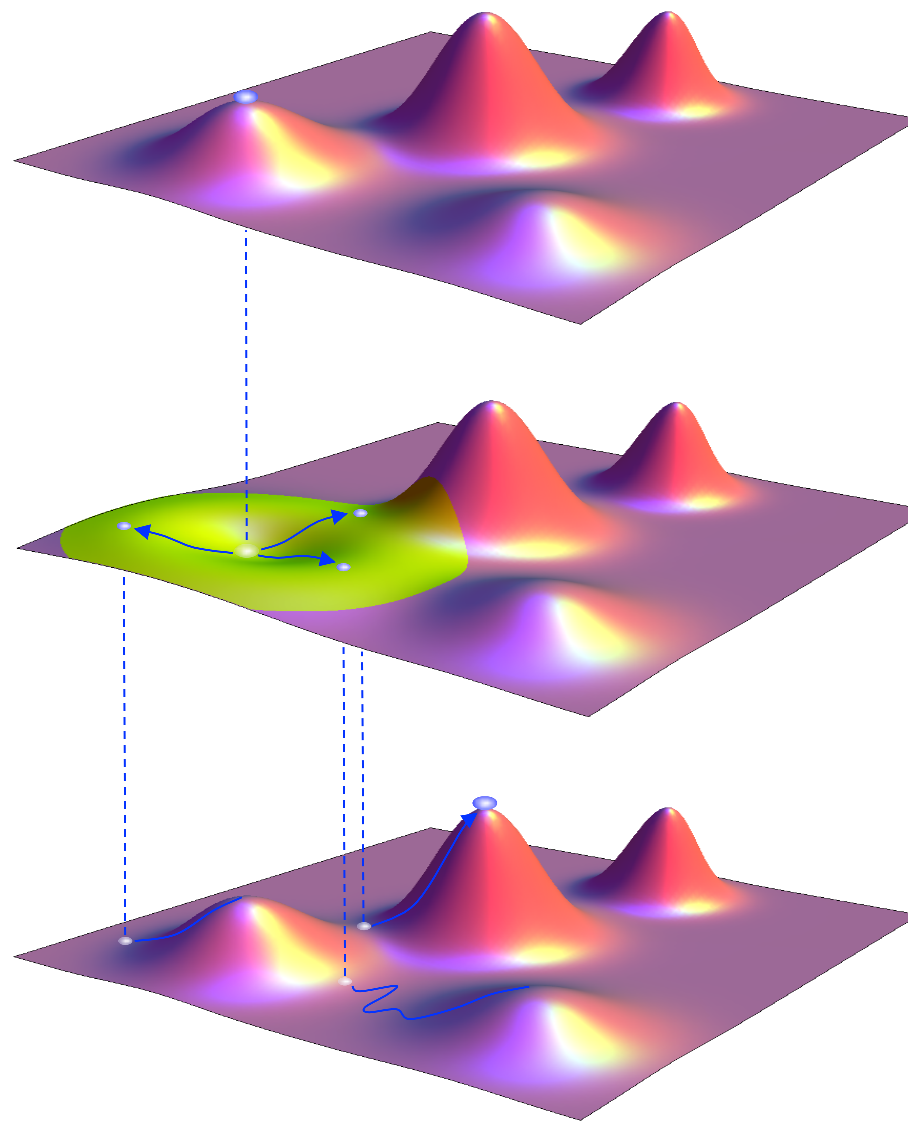

X
Fitness
Genotype Space


Environment A
Environment B
Environment A

Exogenous

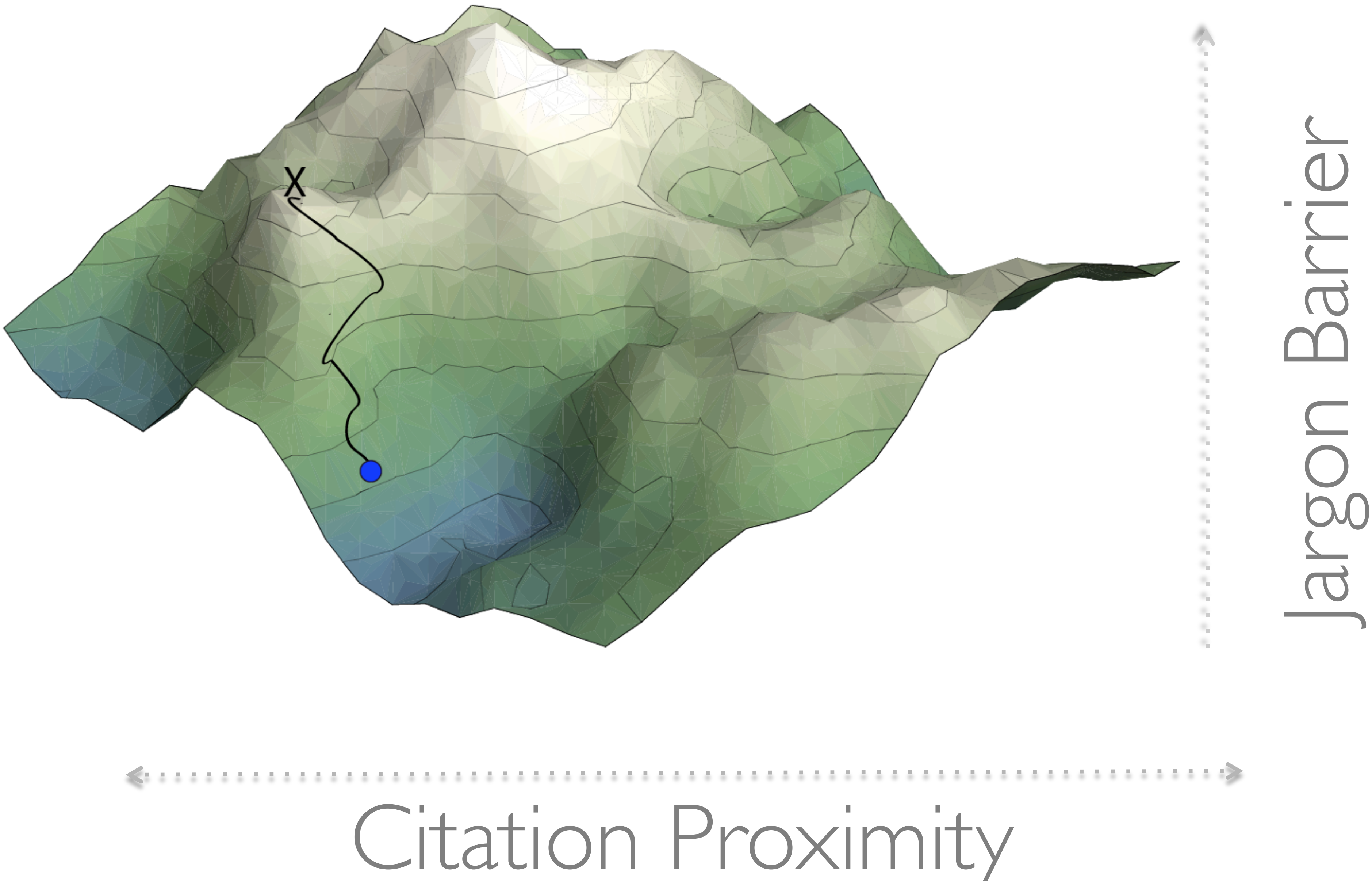


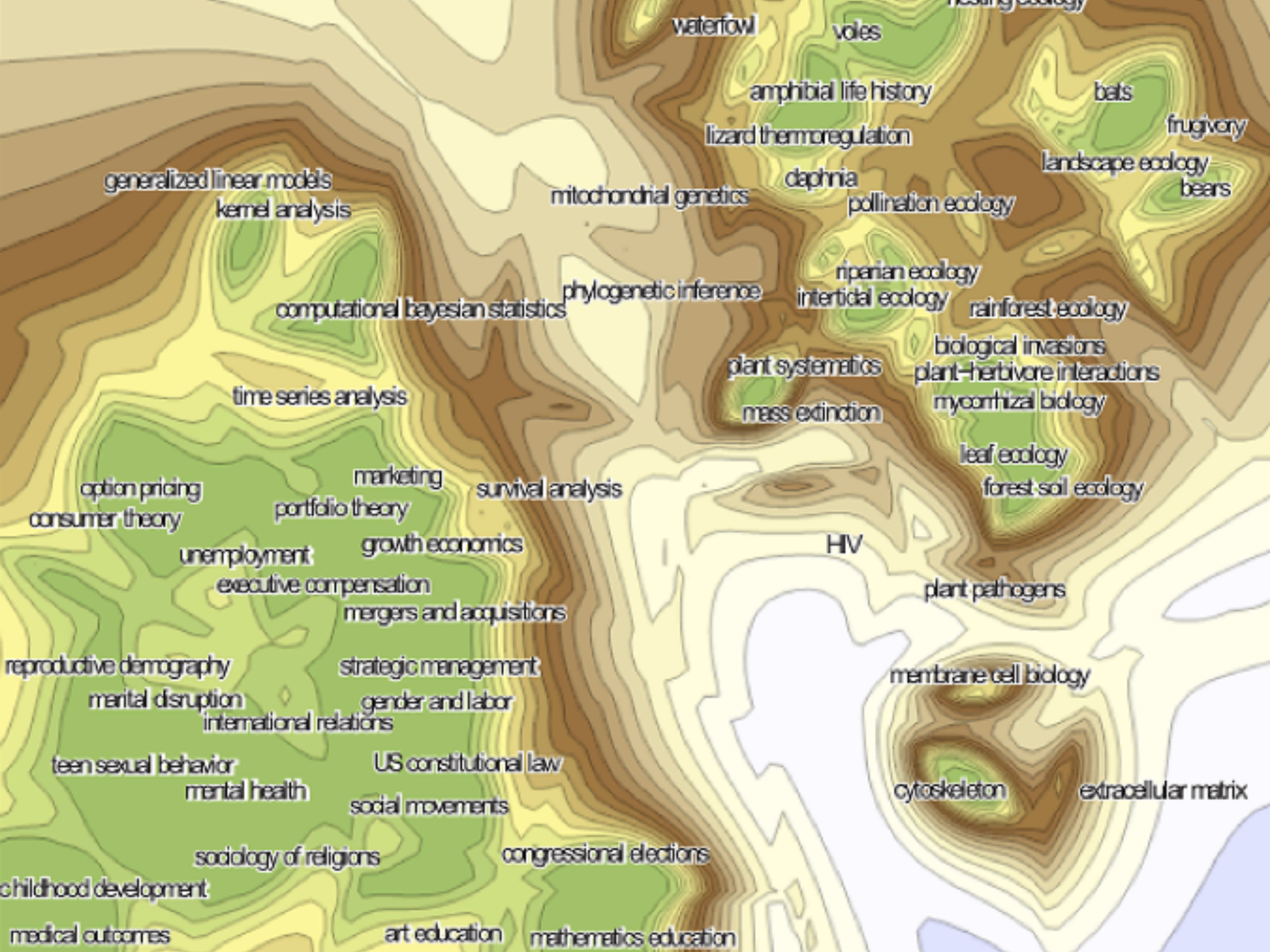
Endogenous

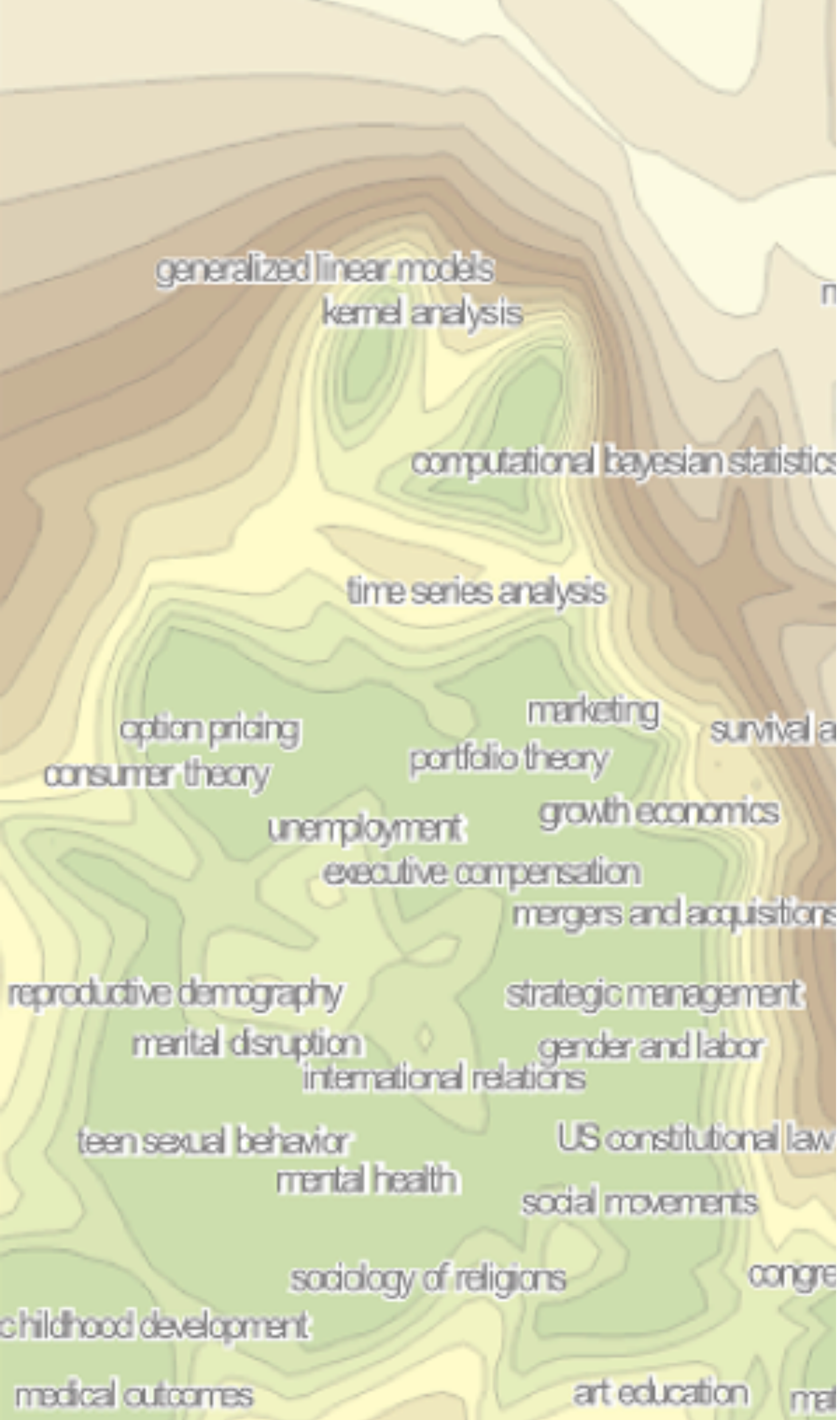


"Every science requires a special language, because every science has its own ideas."

Étienne Bonnot de Condillac, 1752


X
Jargon Barrier
Citation Proximity


nesting ecology
waterfowl
voles
amphibial life history
bats
frugivory
lizard thermoregulation
landscape ecology
bears
generalized linear models
daphnia
mitochondrial genetics
pollination ecology
kernel analysis
riparian ecology
phylogenetic inference
intertidal ecology
rainforest ecology
computational bayesian statistics
biological invasions
plant systematics
plant–herbivore interactions
time series analysis
mycorrhizal biology
mass extinction
leaf ecology
marketing
option pricing
survival analysis
forest soil ecology
portfolio theory
consumer theory
growth economics
HIV
unemployment
executive compensation
plant pathogens
mergers and acquisitions
reproductive demography
strategic management
membrane cell biology
marital disruption
gender and labor
international relations
teen sexual behavior
US constitutional law
mental health
cytoskeleton
extracellular matrix
social movements
sociology of religions
congressional elections
childhood development
medical outcomes
art education
mathematics education



Outline

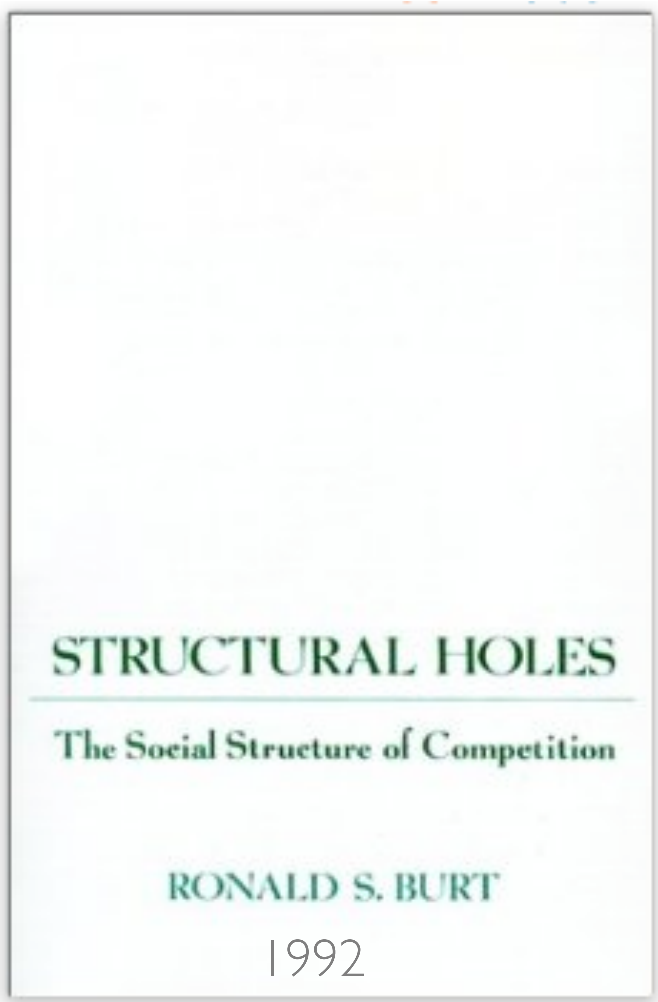
- Background
- Method
- Data
- Results
- Future Directions

CULTURAL

pollination, pollen, pollinators, flowers, flower, seed set, number flowers, pollinator, pollinated, nectar, stigmas, bees, anthers, outcrossing, floral, fruit set, pollen grains, breeding system, seed production, flowering, pollination ecology, selfing, ovules, number pollen, inbreeding depression, breeding systems, pollen transfer, flowers visited, flowering plants, stigma, pollen dispersal, inflorescence, inflorescences, nectar production, open flowers, flowers produced, inbreeding, bee, number seeds, individual flowers, floral display, flowers open, corolla, amount pollen, hermaphroditic, pollination biology, pollinator visitation, fruit seed, bumblebees, male function, flowers plants, floral traits, flowering season, visitation, experimental pollination, seeds produced, pollen production, flowers had, fruit production, pollinator behavior, pollinator visits, selfed, pollinations, pollen flow, outcrossing rate, fruit, visit flowers, pollen deposition, flower number, floral morphology, seeds, pollen limitation, outcrossing rates, female function, anthesis, seed, variation floral, plant reproductive, s c h, principles pollination, anther, variation pollen, grains, pollen deposited, flowers pollinated, flowering period, plants, reproductive biology, pollen donors, ipomopsis aggregata, pollen tubes, flowers plant, pollen load, flowers have, nectar pollen, plant populations, number ovules, floral biology, hummingbirds, flowering phenology, flower visitors, set fruit, pollen tube growth, visitors, handbook experimental, mating system, pollination success, pollen tube, visitation rates, visiting flowers, pollinator limitation, visited flowers, flower production, flower size, dioecy, hummingbird pollination, natural populations, journal botany, pollination systems, visited, flower fruit, seeds fruit, bumble bees, effects pollen, van der pijl, c h barrett, effect pollen, inflorescence size, pollinator activity, insect visitors, stamens, pollen removal, plant, b charlesworth, sex allocation, seed number, stigmatic, evolution dioecy, evolution floral, bagged, outcrossed, floral visitors, pollinator attraction, c e jones, fruits

STRUCTURAL

Structural Holes



1992

Ronald Stuart Burt
University of Chicago

"It is hardly possible to overrate the value...of placing human beings in contact with persons dissimilar to themselves…Such communication [is] one of the primary sources of progress." - Stuart Stuart Mill (1909)

Cultural Holes

2010

ANNUAL REVIEWS **Further**
Click here for quick links to Annual Reviews content online, including:
• Other articles in this volume
• Top cited articles
• Top downloaded articles
• Our comprehensive search

Cultural Holes: Beyond Relationality in Social Networks and Culture

Mark A. Pachucki[1] and Ronald L. Breiger[2]

[1]Department of Sociology, Harvard University, Cambridge, Massachusetts 02138; email: pachucki@fas.harvard.edu
[2]Department of Sociology, University of Arizona, Tucson, Arizona 85721; email: breiger@arizona.edu

Annu. Rev. Sociol. 2010. 36:205–24

First published online as a Review in Advance on April 20, 2010

The *Annual Review of Sociology* is online at soc.annualreviews.org

This article's doi: 10.1146/annurev.soc.012809.102615

Copyright © 2010 by Annual Reviews. All rights reserved

0360-0572/10/0811-0205$20.00

Key Words

interaction, meaning, structure, boundaries, local practices

Abstract

A burgeoning literature spanning sociologies of culture and social network methods has for the past several decades sought to explicate the relationships between culture and connectivity. A number of promising recent moves toward integration are worthy of review, comparison, critique, and synthesis. Network thinking provides powerful techniques for specifying cultural concepts ranging from narrative networks to classification systems, tastes, and cultural repertoires. At the same time, we see theoretical advances by sociologists of culture as providing a corrective to network analysis as it is often portrayed, as a mere collection of methods. Cultural thinking complements and sets a new agenda for moving beyond predominant forms of structural analysis that ignore action, agency, and intersubjective meaning. The notion of "cultural holes" that we use to organize our review points both to the cultural contingency of network structure and to the increasingly permeable boundary between studies of culture and research on social networks.

Common Culture

(shared meanings, tastes, interests)

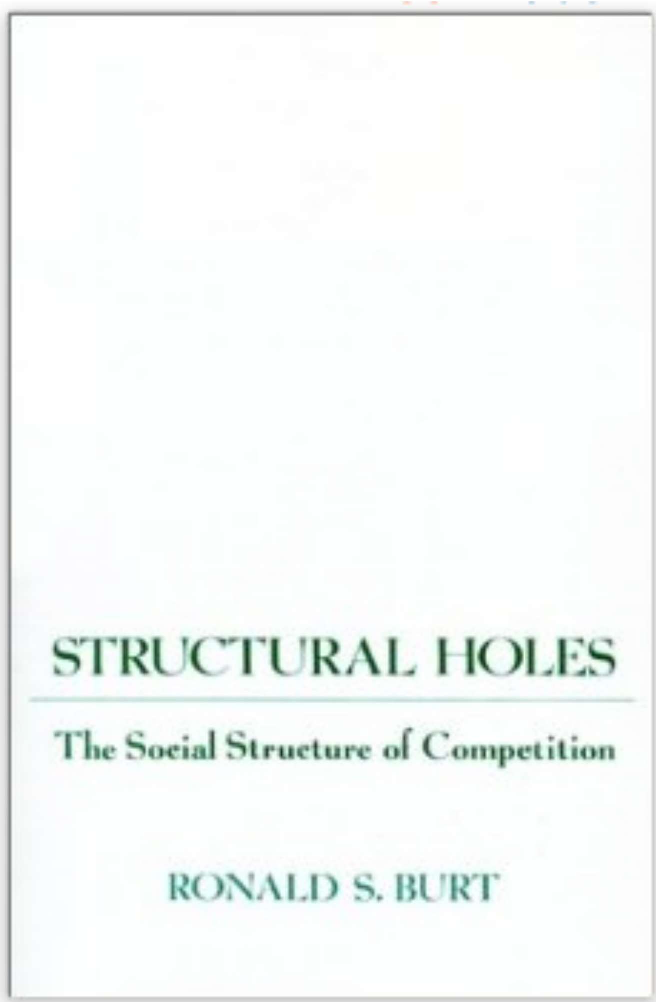
- enables ties between individuals and institutions
- when absent, 'holes' exist
- social actors may be structurally close but far away in 'matters of concern'

CULTURAL

pollination, pollen, pollinators, flowers, flower, seed set, number flowers, pollinator, pollinated, nectar, stigmas, bees, anthers, outcrossing, floral, fruit set, pollen grains, breeding system, seed production, flowering, pollination ecology, selfing, ovules, number pollen, inbreeding depression, breeding systems, pollen transfer, flowers visited, flowering plants, stigma, pollen dispersal, inflorescence, inflorescences, nectar production, open flowers, flowers produced, inbreeding, bee, number seeds, individual flowers, floral display, flowers open, corolla, amount pollen, hermaphroditic, pollination biology, pollinator visitation, fruit seed, bumblebees, male function, flowers plants, floral traits, flowering season, visitation, experimental pollination, seeds produced, pollen production, flowers had, fruit production, pollinator behavior, pollinator visits, selfed, pollinations, pollen flow, outcrossing rate, fruit, visit flowers, pollen deposition, flower number, floral morphology, seeds, pollen limitation, outcrossing rates, female function, anthesis, seed, variation floral, plant reproductive, s c h, principles pollination, anther, variation pollen, grains, pollen deposited, flowers pollinated, flowering period, plants, reproductive biology, pollen donors, ipomopsis aggregata, pollen tubes, flowers plant, pollen load, flowers have, nectar pollen, plant populations, number ovules, floral biology, hummingbirds, flowering phenology, flower visitors, set fruit, pollen tube growth, visitors, handbook experimental, mating system, pollination success, pollen tube, visitation rates, visiting flowers, pollinator limitation, visited flowers, flower production, flower size, dioecy, hummingbird pollination, natural populations, journal botany, pollination systems, visited, flower fruit, seeds fruit, bumble bees, effects pollen, van der pijl, c h barrett, effect pollen, inflorescence size, pollinator activity, insect visitors, stamens, pollen removal, plant, b charlesworth, sex allocation, seed number, stigmatic, evolution dioecy, evolution floral, bagged, outcrossed, floral visitors, pollinator attraction, c e jones, fruits

STRUCTURAL

How do we model the communicative burden imposed by cultural holes?



ANNUAL REVIEWS **Further**

Click here for quick links to Annual Reviews content online, including:

- Other articles in this volume
- Top cited articles
- Top downloaded articles
- Our comprehensive search

Cultural Holes: Beyond Relationality in Social Networks and Culture

Mark A. Pachucki[1] and Ronald L. Breiger[2]

[1]Department of Sociology, Harvard University, Cambridge, Massachusetts 02138; email: pachucki@fas.harvard.edu

[2]Department of Sociology, University of Arizona, Tucson, Arizona 85721; email: breiger@arizona.edu

Annu. Rev. Sociol. 2010. 36:205–24

First published online as a Review in Advance on April 20, 2010

The *Annual Review of Sociology* is online at soc.annualreviews.org

This article's doi: 10.1146/annurev.soc.012809.102615

Copyright © 2010 by Annual Reviews. All rights reserved

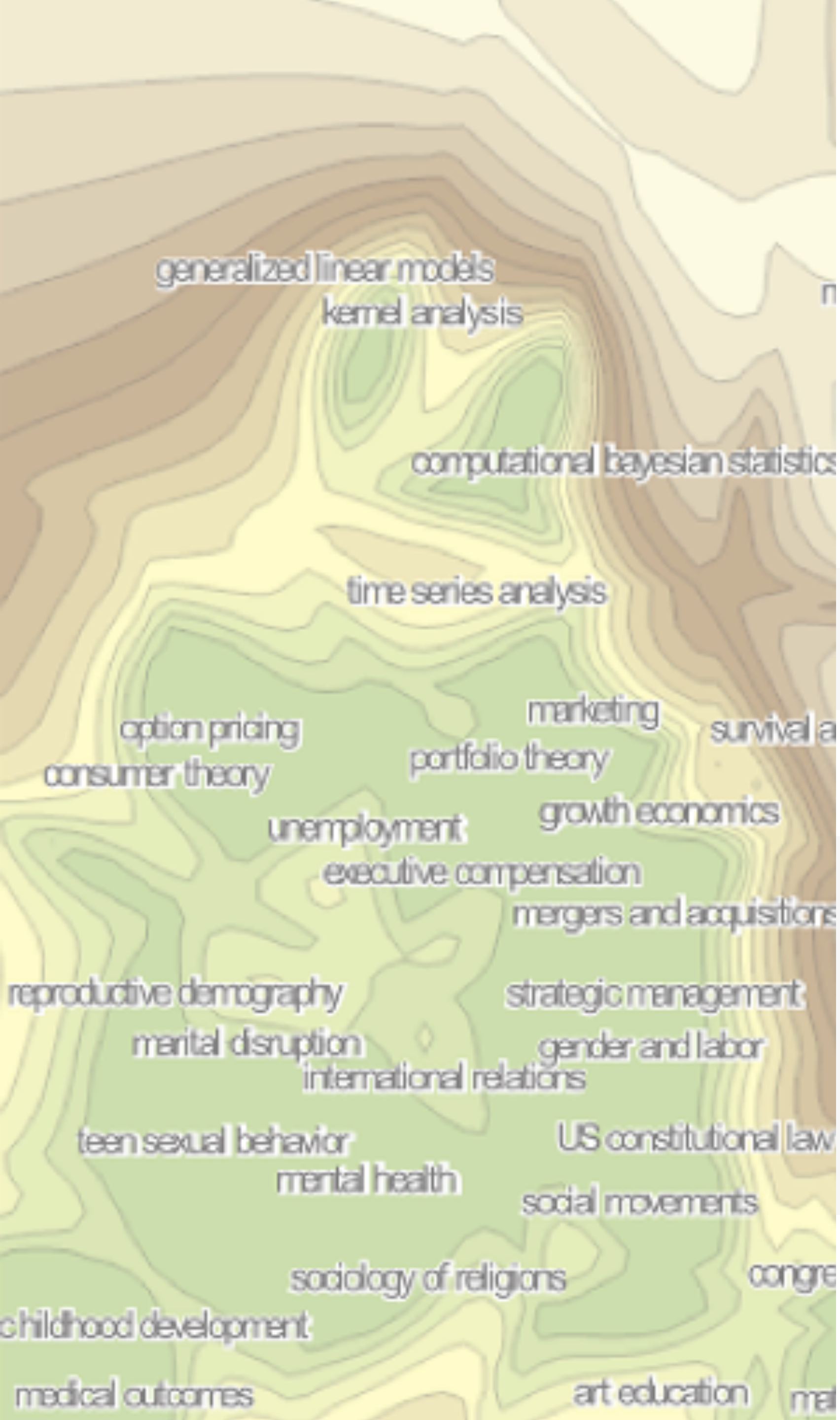
0360-0572/10/0811-0205$20.00

Key Words

interaction, meaning, structure, boundaries, local practices

Abstract

A burgeoning literature spanning sociologies of culture and social network methods has for the past several decades sought to explicate the relationships between culture and connectivity. A number of promising recent moves toward integration are worthy of review, comparison, critique, and synthesis. Network thinking provides powerful techniques for specifying cultural concepts ranging from narrative networks to classification systems, tastes, and cultural repertoires. At the same time, we see theoretical advances by sociologists of culture as providing a corrective to network analysis as it is often portrayed, as a mere collection of methods. Cultural thinking complements and sets a new agenda for moving beyond predominant forms of structural analysis that ignore action, agency, and intersubjective meaning. The notion of "cultural holes" that we use to organize our review points both to the cultural contingency of network structure and to the increasingly permeable boundary between studies of culture and research on social networks.

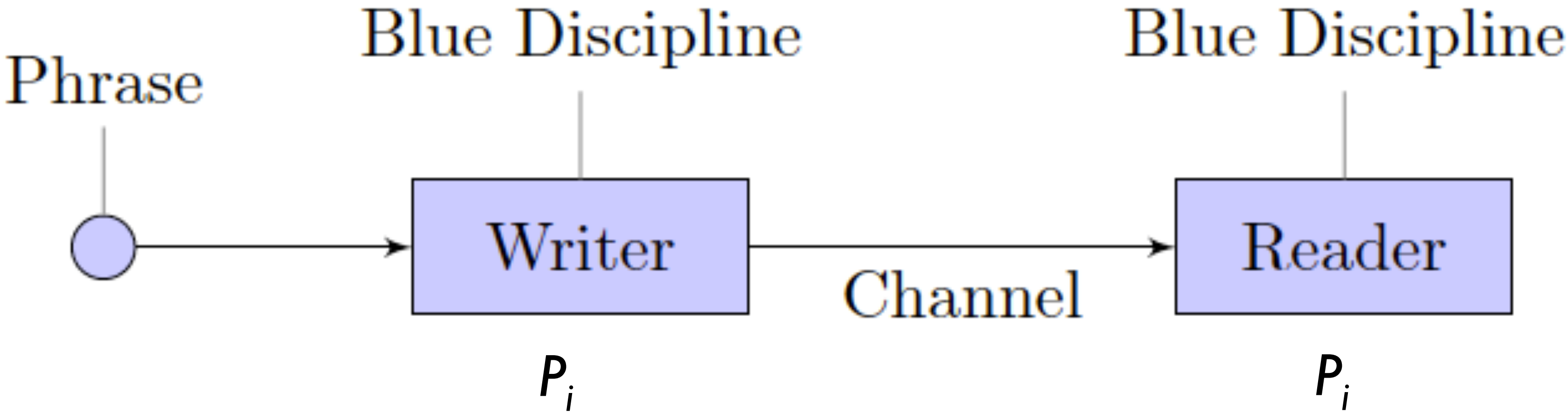


Outline

- Background
- **Method**
- Data
- Results
- Future Directions

Methods

- Model of optimal communication
- Penalty for communicating across language sets
- Operationalize building blocks of model



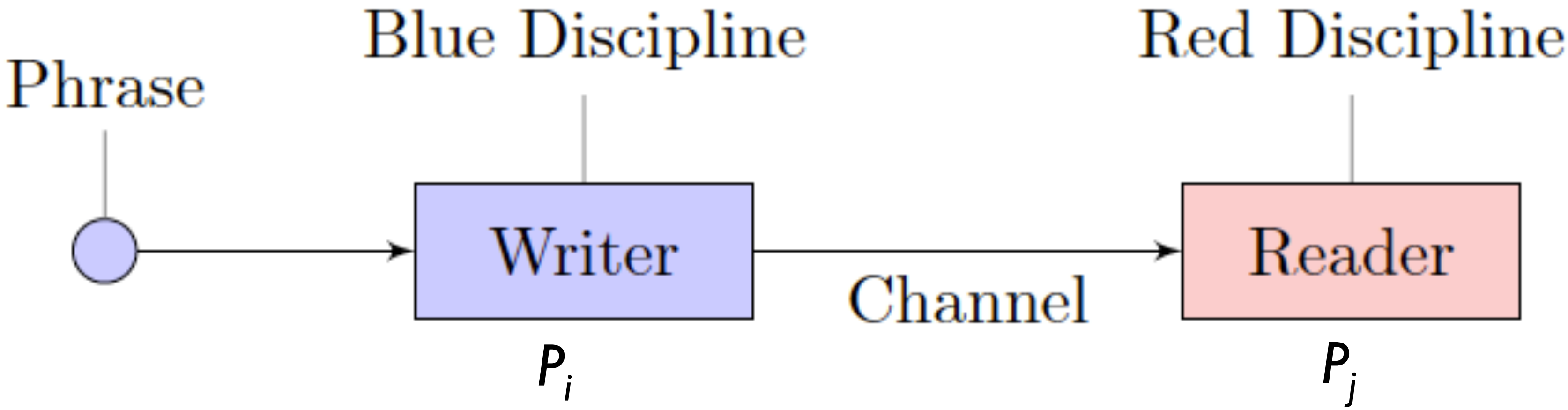
$\boldsymbol{X}$ ~ space of all phrases

P_i ~ probability distribution over χ_i with values $x \in \boldsymbol{X}$

- writer chooses phrases with probability $p_i(x)$
- optimal codeword has length $-\log_2 p_i(x)$

expected message length: $H(X_i) = -\sum_{x \in \mathcal{X}} p_i(x) \log_2 p_i(x)$

assumption: language of each scientific field is *optimized* based on frequency of phrases



cross entropy

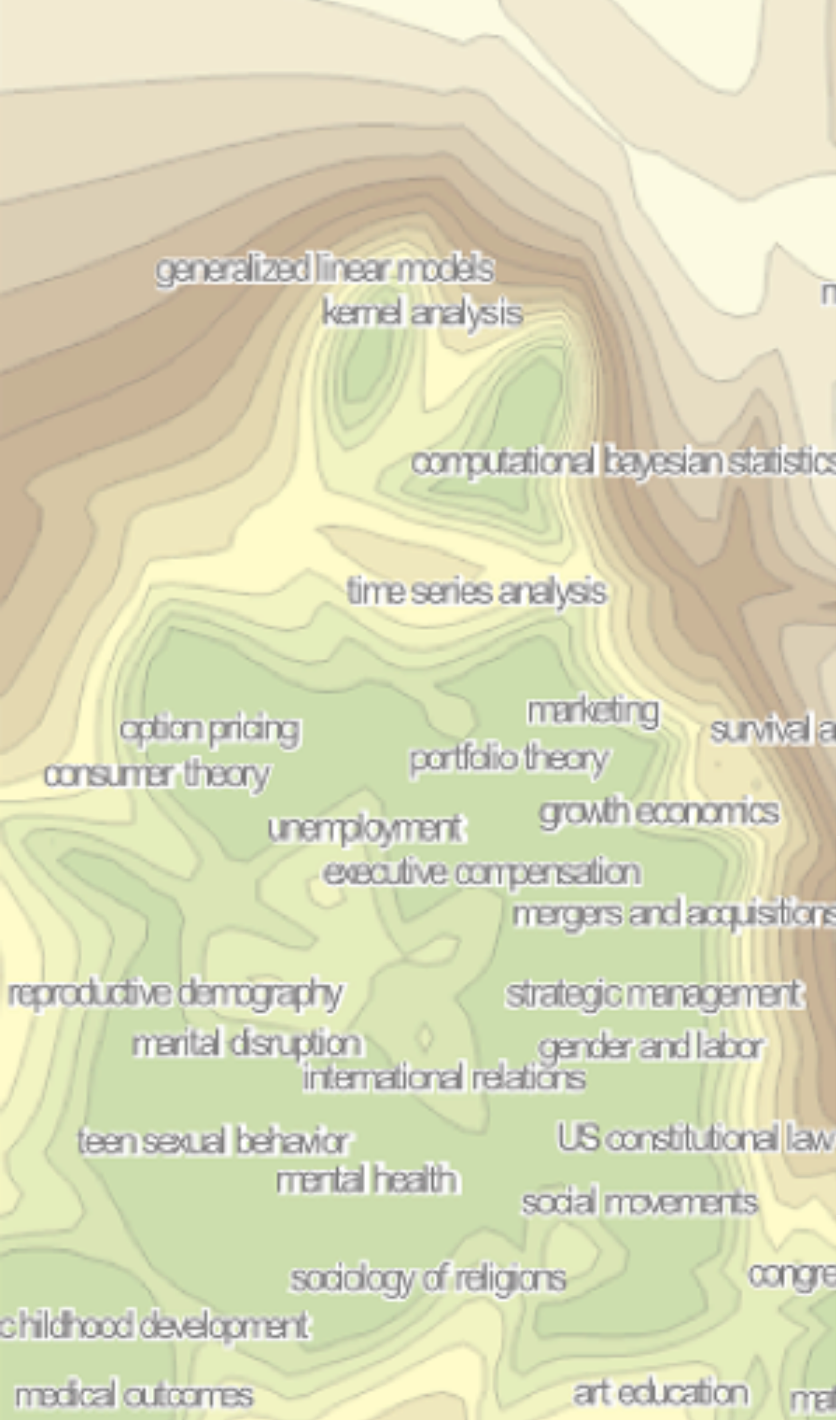
expected message length: $Q(p_i||p_j) = -\sum_{x \in \mathcal{X}} p_i(x) \log_2 p_j(x)$

efficiency of communication

$$E_{ij} = \frac{H(X_i)}{Q(p_i||p_j)} = \frac{-\sum_{x \in \mathcal{X}} p_i(x) \log_2 p_i(x)}{-\sum_{x \in \mathcal{X}} p_i(x) \log_2 p_j(x)}$$

$$C_{ij} = 1 - E_{ij}$$

cultural hole



Outline

- Background
- Method
- **Data**
- Results
- Future Directions

345 years

8,000,000 full text articles

500,000 unique authors

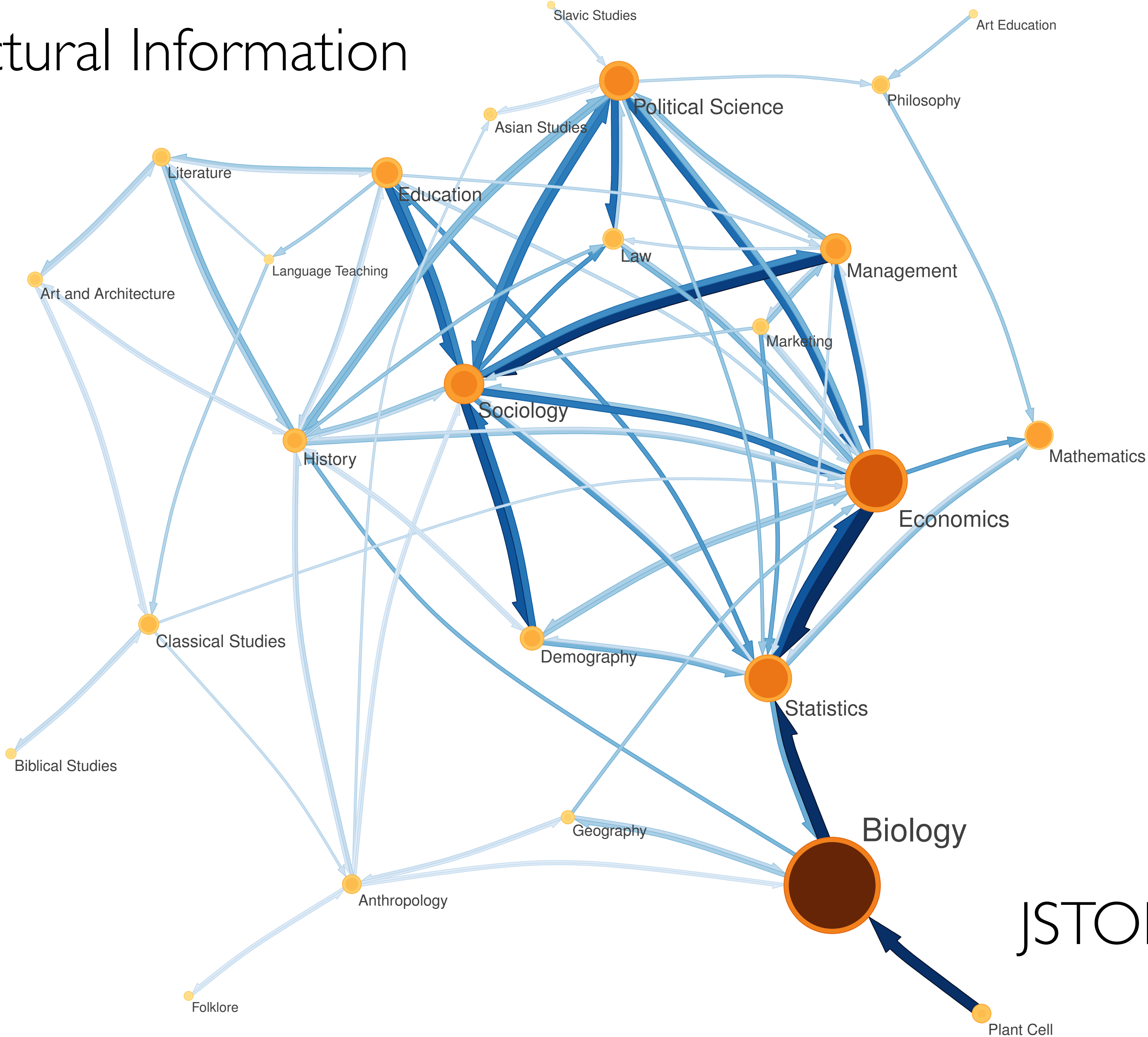
1. Identify the disciplinary boundaries using citation patterns. Select the 60 largest research areas for study.

The map equation

$$L(\mathsf{M}) = q_{\curvearrowright} H(\mathcal{Q}) + \sum_{i=1}^{m} p_{\circlearrowright}^{i} H(\mathcal{P}^{i})$$

Rosvall and Bergstrom (2008) *PNAS*

Structural Information

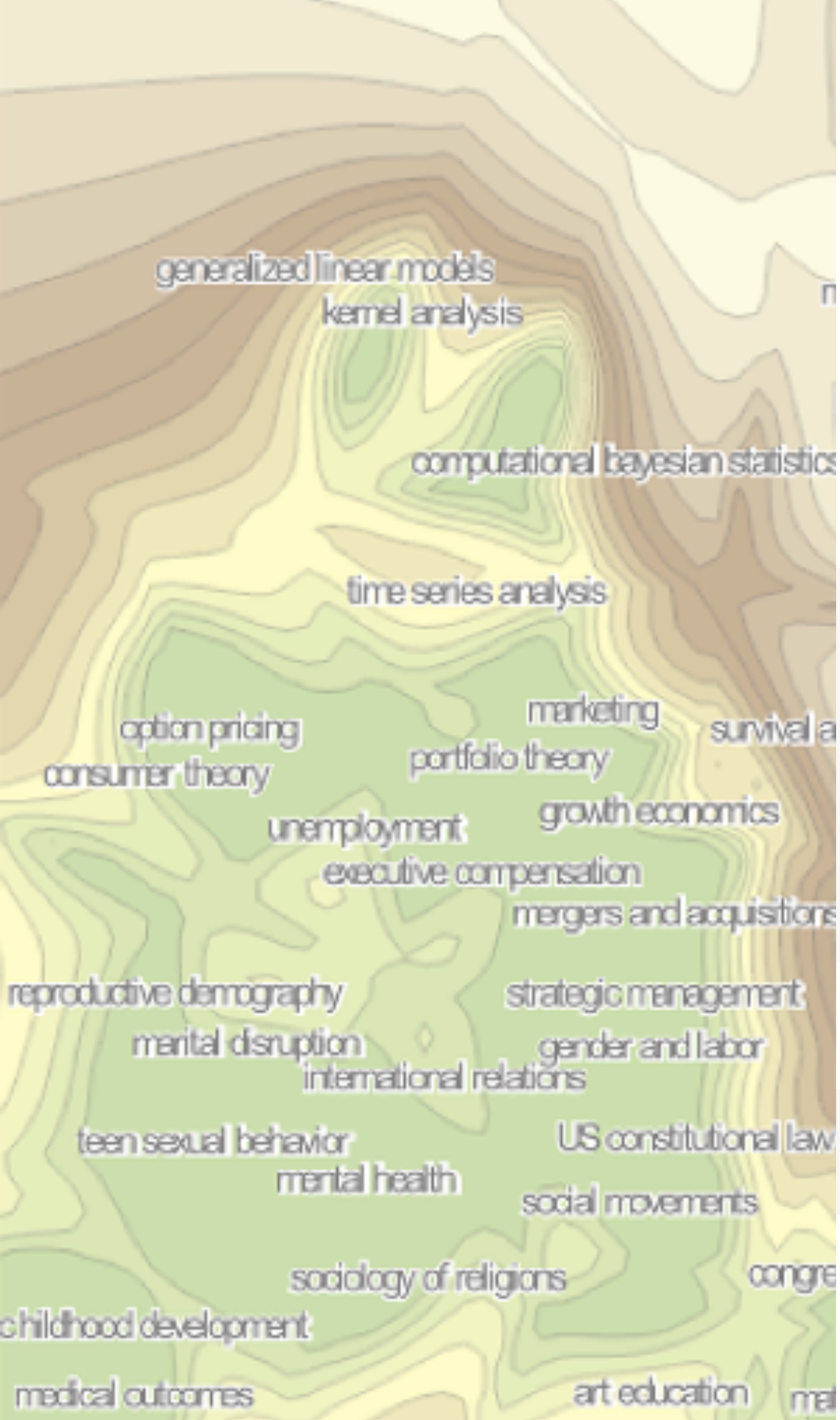


JSTOR

Cultural Information

"The phrase frequency distribution for each scholarly field was assembled using the empirical frequency of each triplet of consecutive words (trigram)."

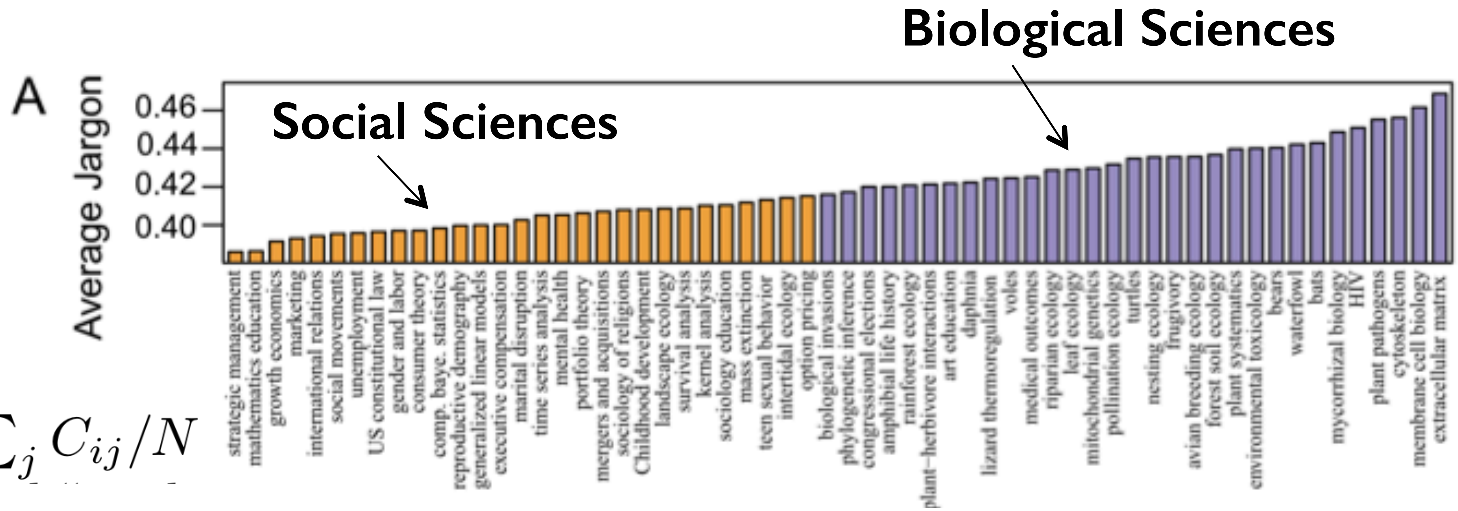
trigrams: *the phrase frequency*, *phrase frequency distribution*…



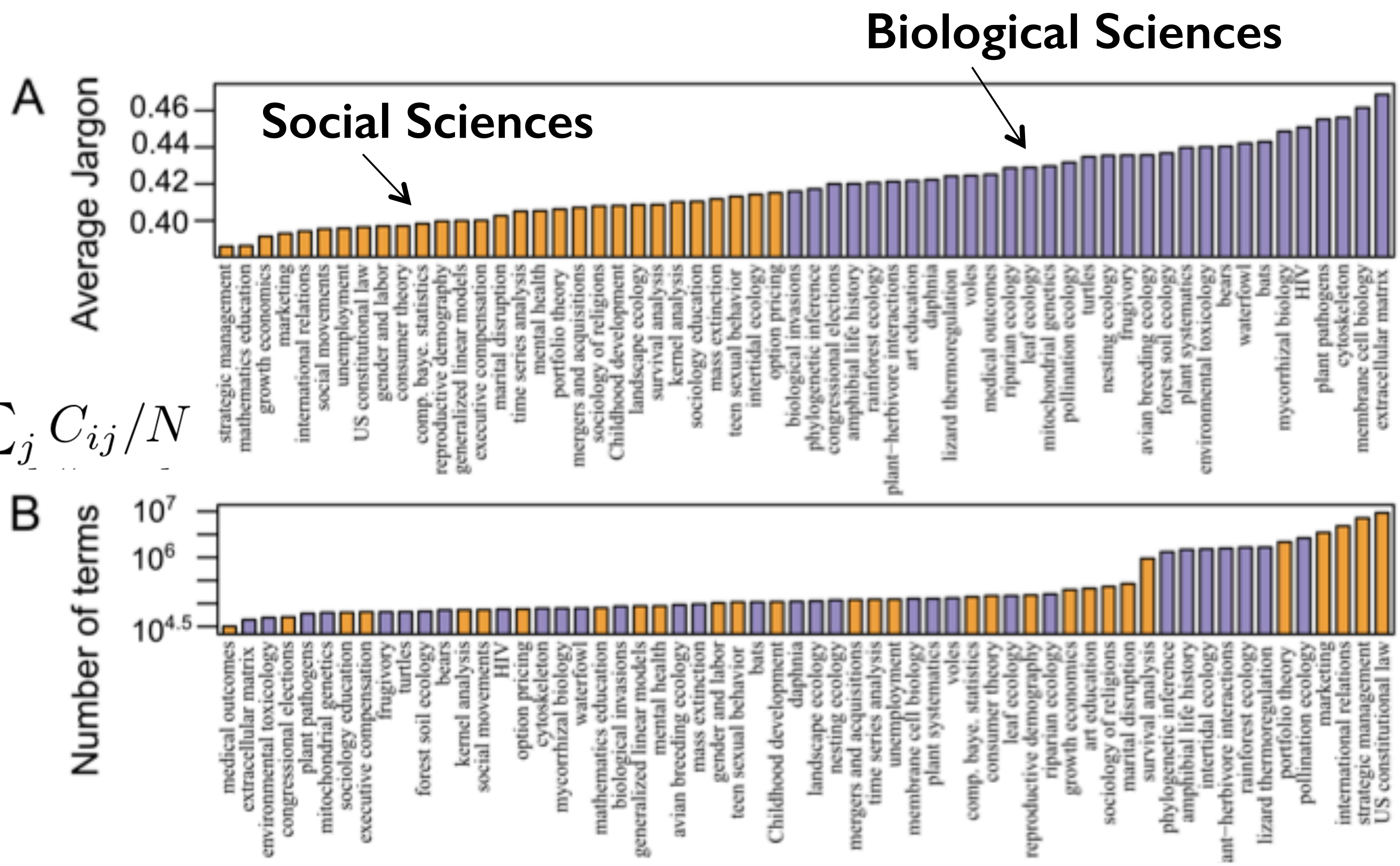
Outline

- Background
- Method
- Data
- **Results**
- Future Directions

$$C_i = \sum_j C_{ij}/N$$

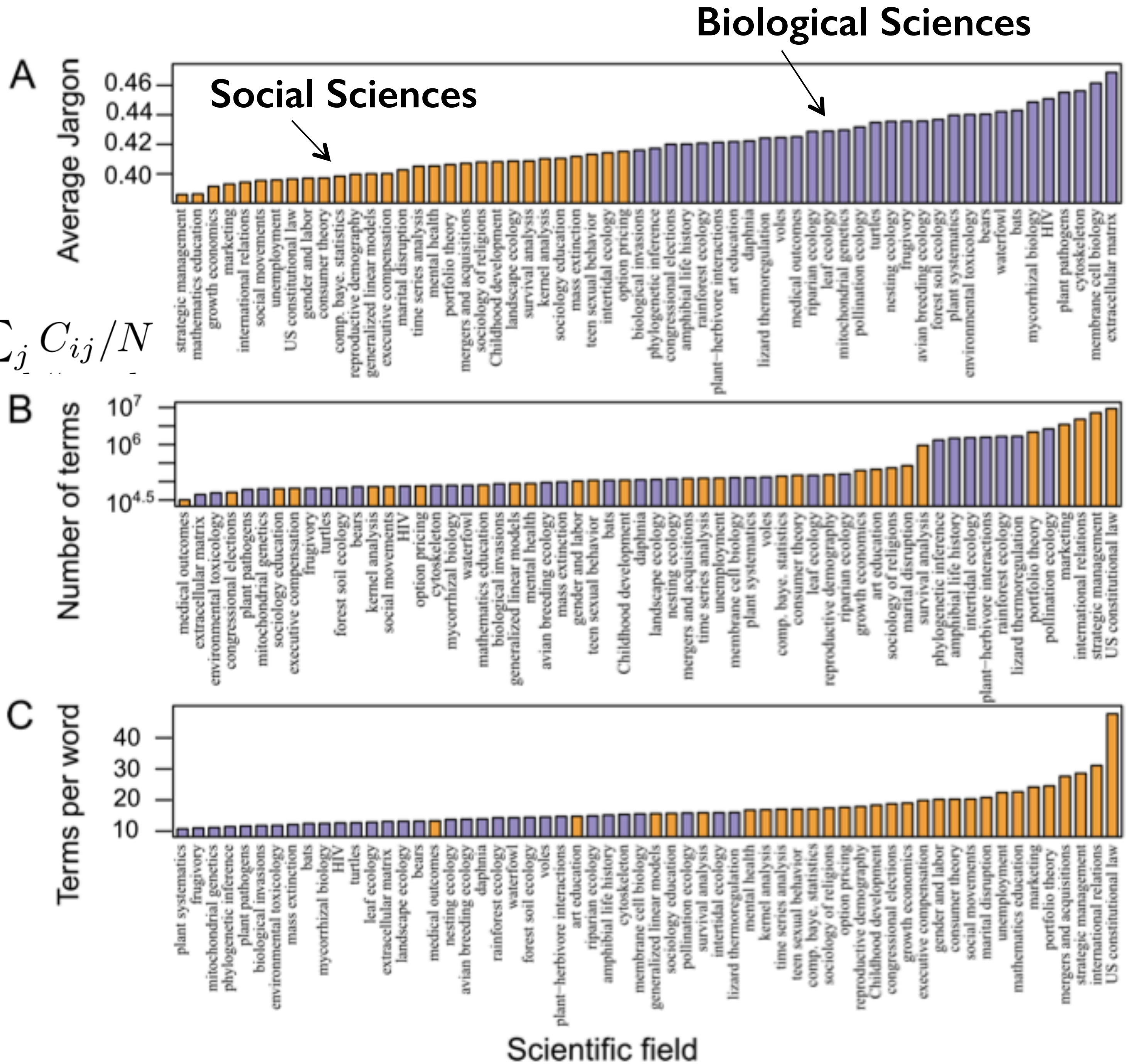


scientific field



$$C_i = \sum_j C_{ij}/N$$

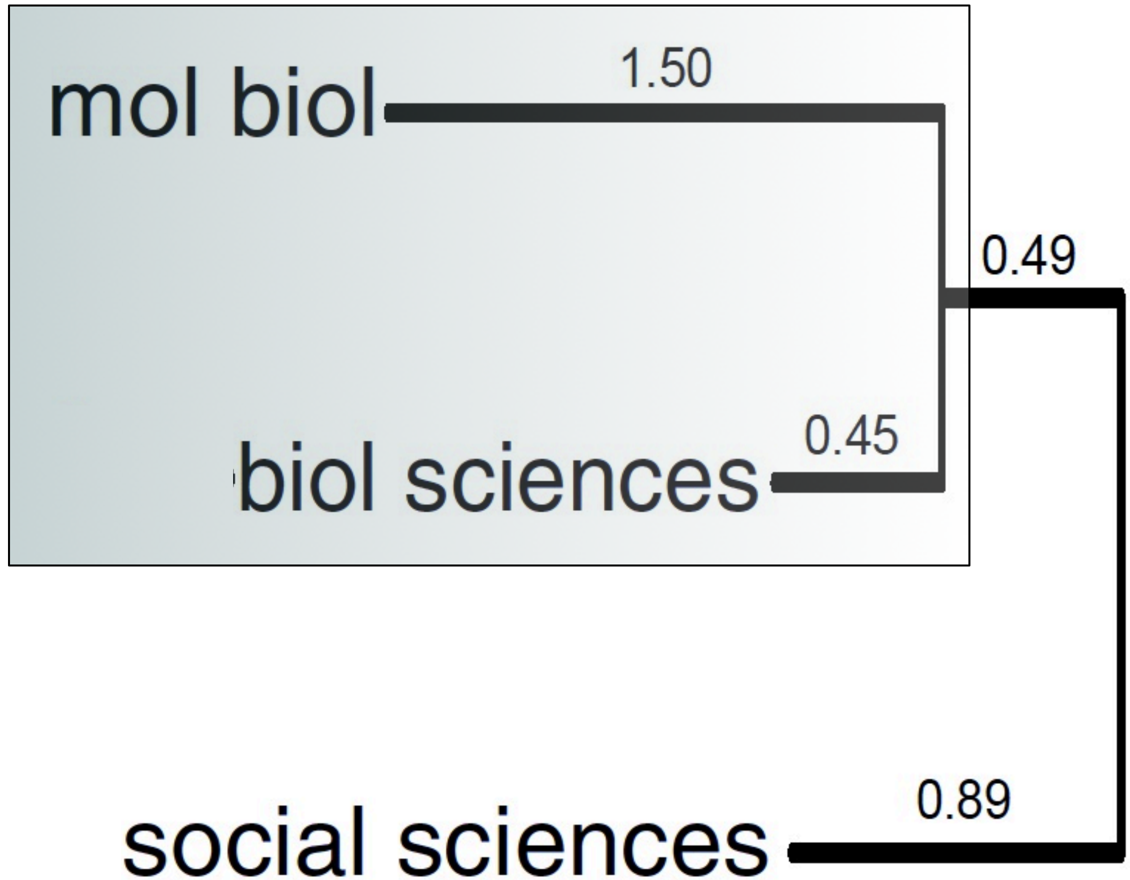
scientific field


Biological Sciences
Social Sciences
A
Average Jargon
0.46
0.44
0.42
0.40
strategic management
mathematics education
growth economics
marketing
international relations
social movements
unemployment
US constitutional law
gender and labor
consumer theory
comp. baye. statistics
reproductive demography
generalized linear models
executive compensation
marital disruption
time series analysis
mental health
portfolio theory
mergers and acquisitions
sociology of religions
Childhood development
landscape ecology
survival analysis
kernel analysis
sociology education
mass extinction
teen sexual behavior
intertidal ecology
option pricing
biological invasions
phylogenetic inference
congressional elections
amphibial life history
rainforest ecology
plant−herbivore interactions
art education
daphnia
lizard thermoregulation
voles
medical outcomes
riparian ecology
leaf ecology
mitochondrial genetics
pollination ecology
turtles
nesting ecology
frugivory
avian breeding ecology
forest soil ecology
plant systematics
environmental toxicology
bears
waterfowl
bats
mycorrhizal biology
HIV
plant pathogens
cytoskeleton
membrane cell biology
extracellular matrix
B
Number of terms
10^7
10^6
10^4.5
medical outcomes
extracellular matrix
environmental toxicology
congressional elections
plant pathogens
mitochondrial genetics
sociology education
executive compensation
frugivory
turtles
forest soil ecology
bears
kernel analysis
social movements
HIV
option pricing
cytoskeleton
mycorrhizal biology
waterfowl
mathematics education
biological invasions
generalized linear models
mental health
avian breeding ecology
mass extinction
gender and labor
teen sexual behavior
bats
Childhood development
daphnia
landscape ecology
nesting ecology
mergers and acquisitions
time series analysis
unemployment
membrane cell biology
plant systematics
voles
comp. baye. statistics
consumer theory
leaf ecology
reproductive demography
riparian ecology
growth economics
art education
sociology of religions
marital disruption
survival analysis
phylogenetic inference
amphibial life history
intertidal ecology
plant−herbivore interactions
rainforest ecology
lizard thermoregulation
portfolio theory
pollination ecology
marketing
international relations
strategic management
US constitutional law
C
Terms per word
40
30
20
10
plant systematics
frugivory
mitochondrial genetics
phylogenetic inference
plant pathogens
biological invasions
environmental toxicology
mass extinction
bats
mycorrhizal biology
HIV
turtles
leaf ecology
extracellular matrix
landscape ecology
bears
medical outcomes
nesting ecology
avian breeding ecology
daphnia
rainforest ecology
waterfowl
forest soil ecology
voles
plant−herbivore interactions
art education
riparian ecology
amphibial life history
cytoskeleton
membrane cell biology
generalized linear models
sociology education
pollination ecology
survival analysis
intertidal ecology
lizard thermoregulation
mental health
kernel analysis
time series analysis
teen sexual behavior
comp. baye. statistics
sociology of religions
option pricing
reproductive demography
Childhood development
congressional elections
growth economics
executive compensation
gender and labor
consumer theory
social movements
marital disruption
unemployment
mathematics education
marketing
portfolio theory
mergers and acquisitions
strategic management
international relations
US constitutional law
Scientific field

$$C_i = \sum_j C_{ij}/N$$

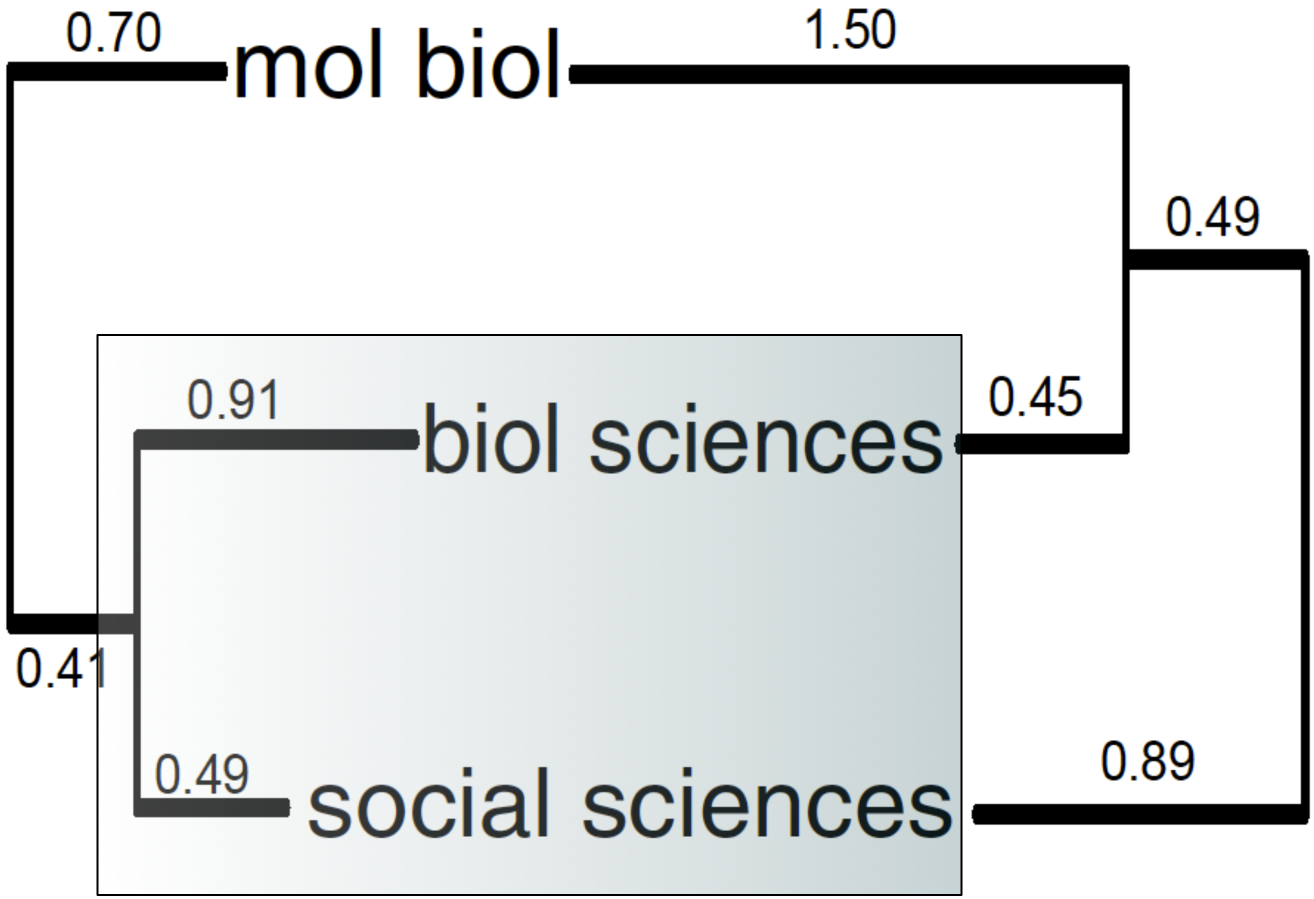
3. Cluster according to citation distance or communication cost using UPGMA.

Citation clustering

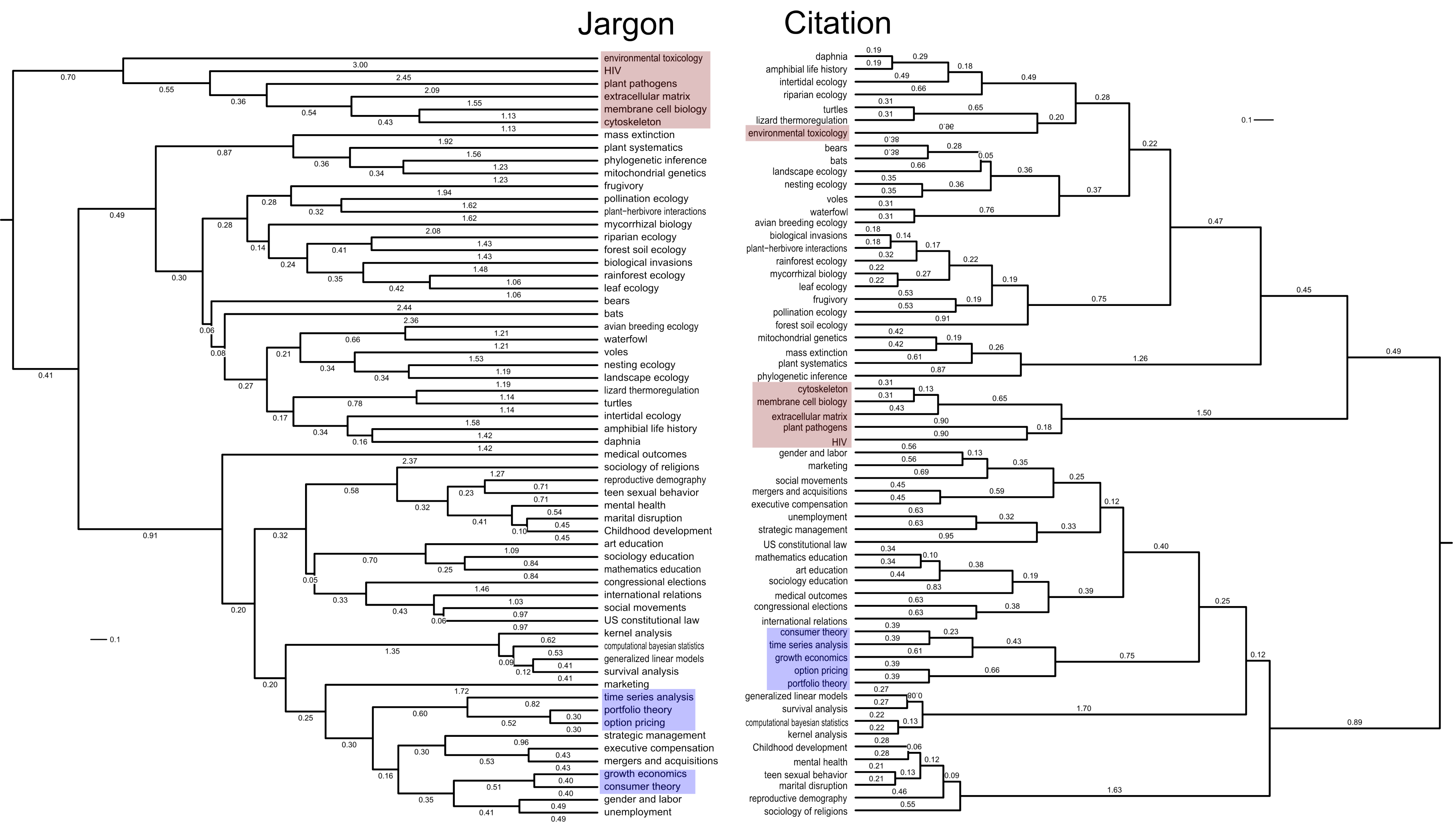


Method: hierarchical clustering (UPGMA) on the average shortest citation path between fields

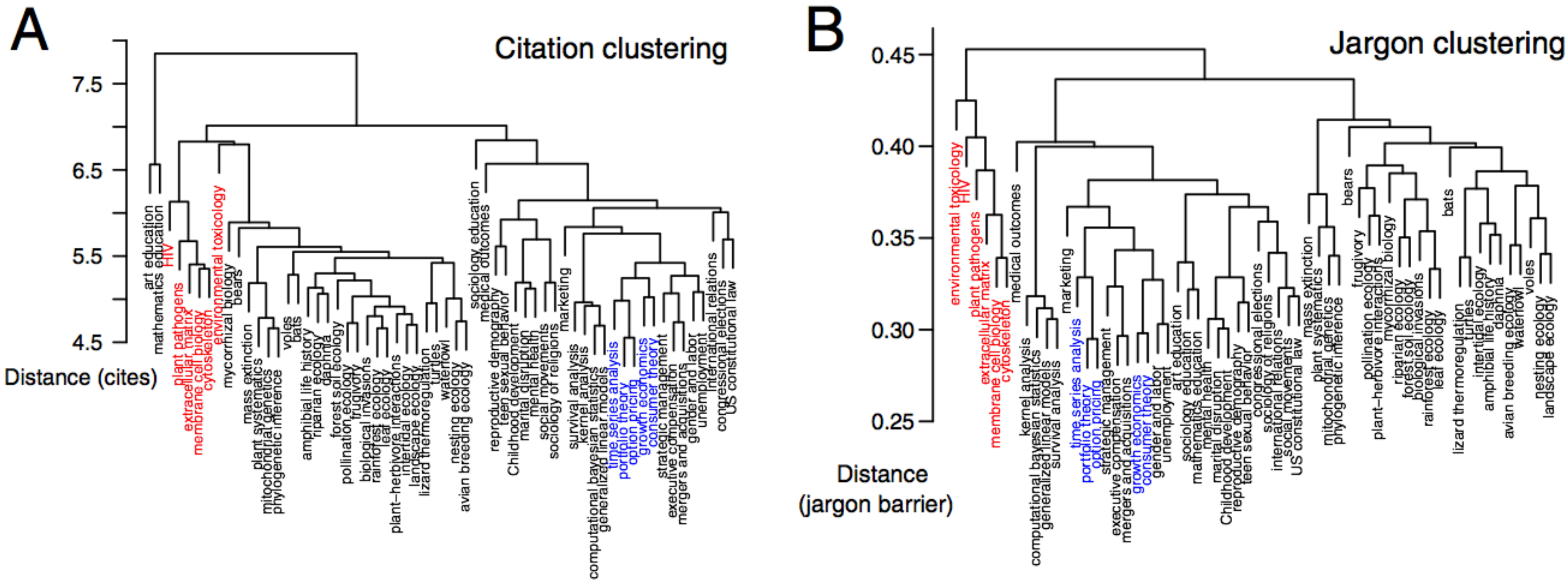
Jargon Citation

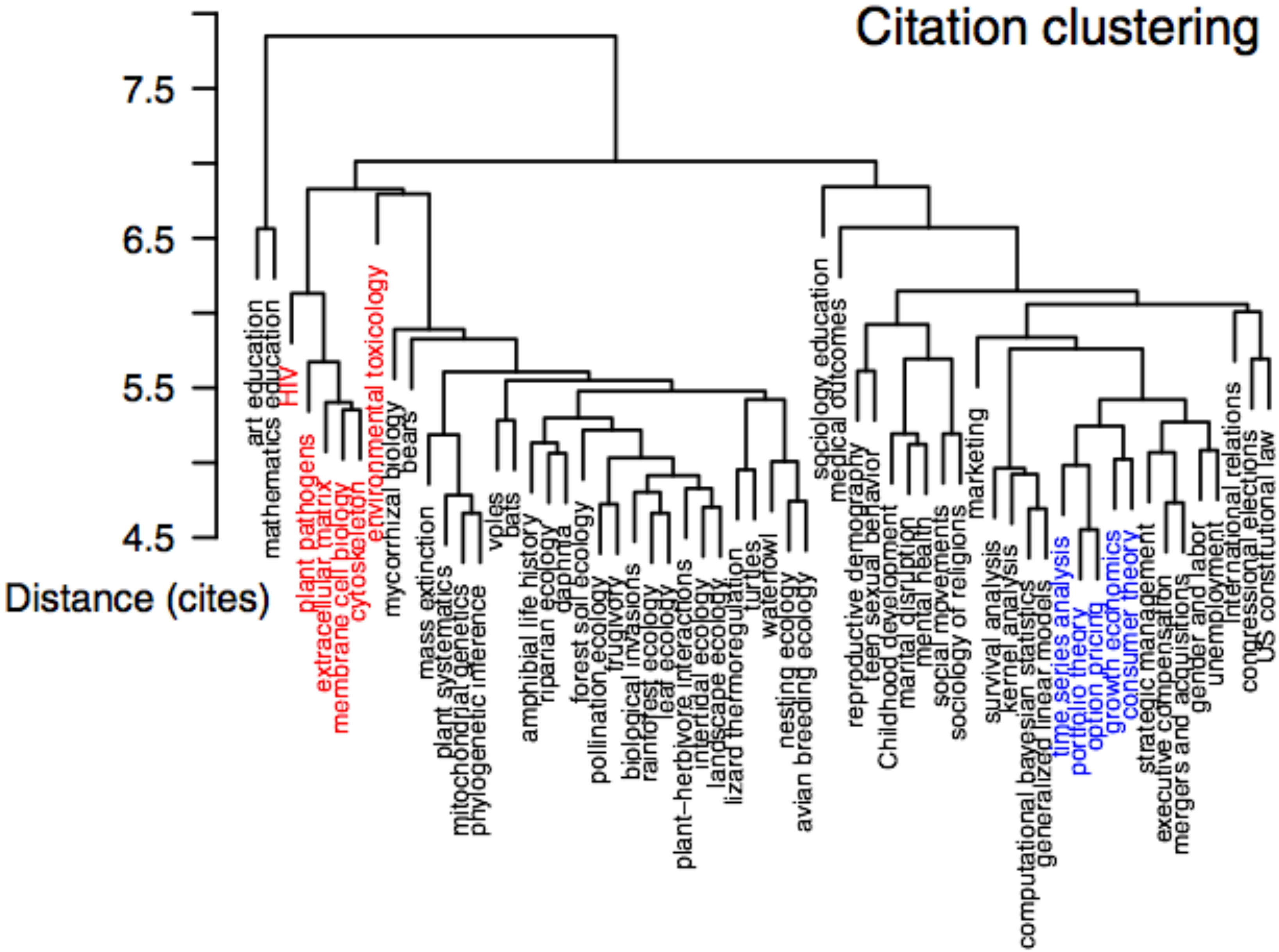


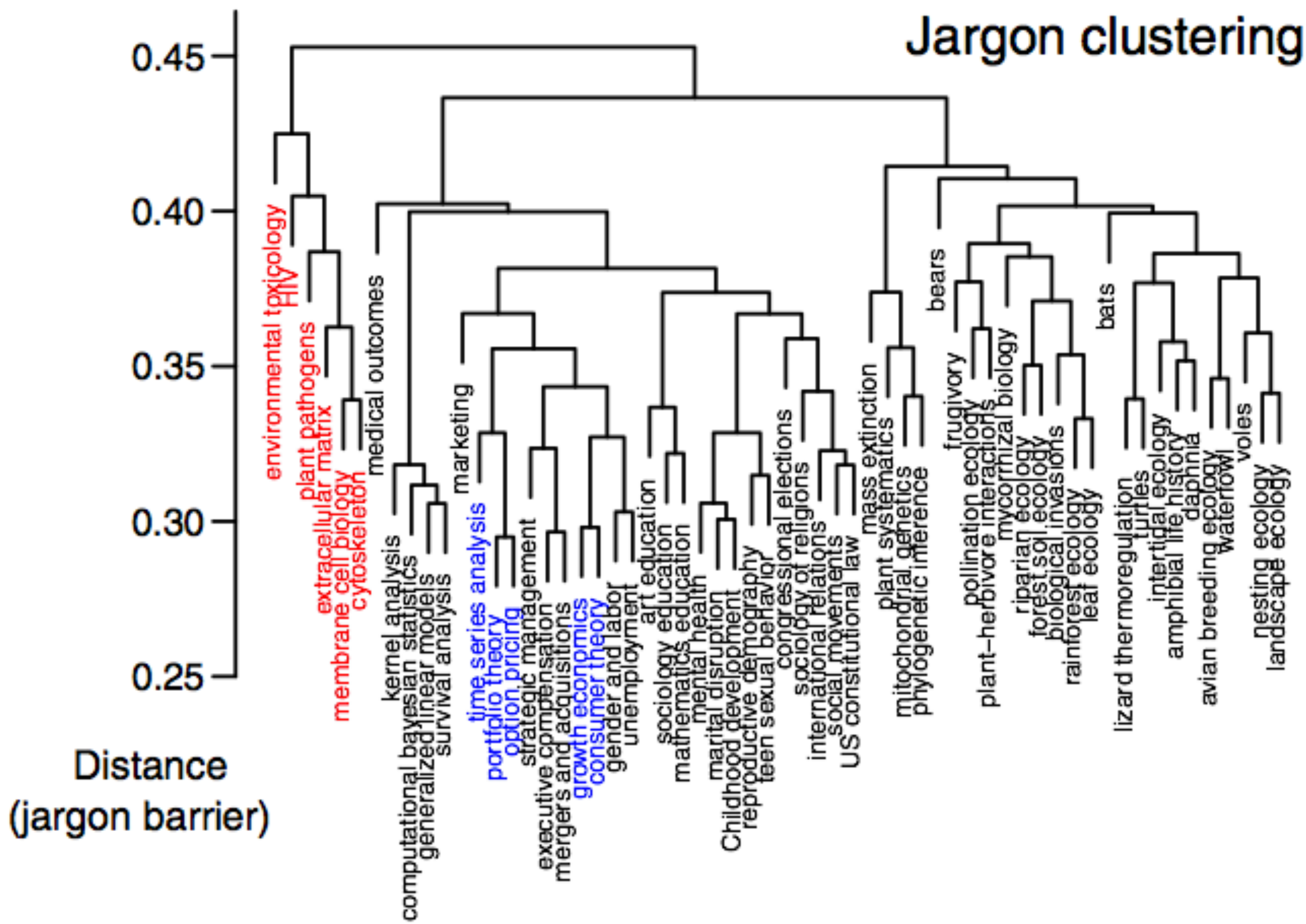
$$\tilde{C}_{ij} = (C_{ij} + C_{ji})/2$$



Jargon
Citation
environmental toxicology
HIV
plant pathogens
extracellular matrix
membrane cell biology
cytoskeleton
mass extinction
plant systematics
phylogenetic inference
mitochondrial genetics
frugivory
pollination ecology
plant-herbivore interactions
mycorrhizal biology
riparian ecology
forest soil ecology
biological invasions
rainforest ecology
leaf ecology
bears
bats
avian breeding ecology
waterfowl
voles
nesting ecology
landscape ecology
lizard thermoregulation
turtles
intertidal ecology
amphibial life history
daphnia
medical outcomes
sociology of religions
reproductive demography
teen sexual behavior
mental health
marital disruption
Childhood development
art education
sociology education
mathematics education
congressional elections
international relations
social movements
US constitutional law
kernel analysis
computational bayesian statistics
generalized linear models
survival analysis
marketing
time series analysis
portfolio theory
option pricing
strategic management
executive compensation
mergers and acquisitions
growth economics
consumer theory
gender and labor
unemployment
0.1

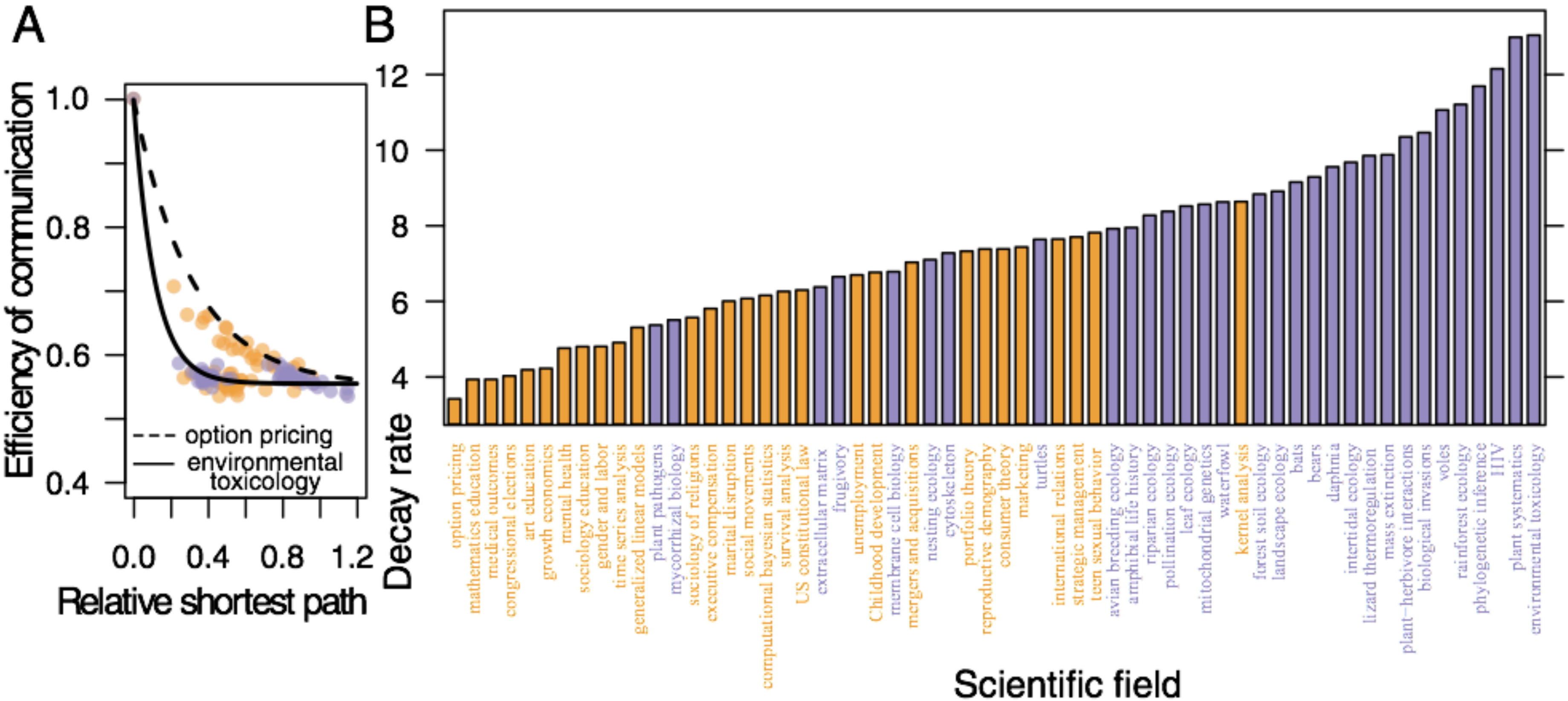

A
Citation clustering
7.5
6.5
5.5
4.5
Distance (cites)
art education
mathematics education
HIV
plant pathogens
extracellular matrix
membrane cell biology
cytoskeleton
environmental toxicology
mycorrhizal biology
bears
mass extinction
plant systematics
mitochondrial genetics
phylogenetic inference
voles
bats
amphibial life history
riparian ecology
daphnia
forest soil ecology
pollination ecology
frugivory
biological invasions
rainforest ecology
leaf ecology
plant−herbivore interactions
intertidal ecology
landscape ecology
lizard thermoregulation
turtles
waterfowl
nesting ecology
avian breeding ecology
sociology education
medical outcomes
reproductive demography
teen sexual behavior
Childhood development
marital disruption
mental health
social movements
sociology of religions
marketing
survival analysis
kernel analysis
computational bayesian statistics
generalized linear models
time series analysis
portfolio theory
option pricing
growth economics
consumer theory
strategic management
executive compensation
mergers and acquisitions
gender and labor
unemployment
international relations
congressional elections
US constitutional law
B
Jargon clustering
0.45
0.40
0.35
0.30
0.25
Distance
(jargon barrier)
environmental toxicology
HIV
plant pathogens
extracellular matrix
membrane cell biology
cytoskeleton
medical outcomes
kernel analysis
computational bayesian statistics
generalized linear models
survival analysis
marketing
time series analysis
portfolio theory
option pricing
strategic management
executive compensation
mergers and acquisitions
growth economics
consumer theory
gender and labor
unemployment
art education
sociology education
mathematics education
mental health
marital disruption
Childhood development
reproductive demography
teen sexual behavior
congressional elections
sociology of religions
international relations
social movements
US constitutional law
mass extinction
plant systematics
mitochondrial genetics
phylogenetic inference
bears
frugivory
pollination ecology
plant−herbivore interactions
mycorrhizal biology
riparian ecology
forest soil ecology
biological invasions
rainforest ecology
leaf ecology
bats
lizard thermoregulation
turtles
intertidal ecology
amphibial life history
daphnia
avian breeding ecology
waterfowl
voles
nesting ecology
landscape ecology


Citation clustering
7.5
6.5
5.5
4.5
Distance (cites)
art education
mathematics education
HIV
plant pathogens
extracellular matrix
membrane cell biology
cytoskeleton
environmental toxicology
mycorrhizal biology
bears
mass extinction
plant systematics
mitochondrial genetics
phylogenetic inference
voles
bats
amphibial life history
riparian ecology
daphnia
forest soil ecology
pollination ecology
frugivory
biological invasions
rainforest ecology
leaf ecology
plant−herbivore interactions
intertidal ecology
landscape ecology
lizard thermoregulation
turtles
waterfowl
nesting ecology
avian breeding ecology
sociology education
medical outcomes
reproductive demography
teen sexual behavior
Childhood development
marital disruption
mental health
social movements
sociology of religions
marketing
survival analysis
kernel analysis
computational bayesian statistics
generalized linear models
time series analysis
portfolio theory
option pricing
growth economics
consumer theory
strategic management
executive compensation
mergers and acquisitions
gender and labor
unemployment
international relations
congressional elections
US constitutional law

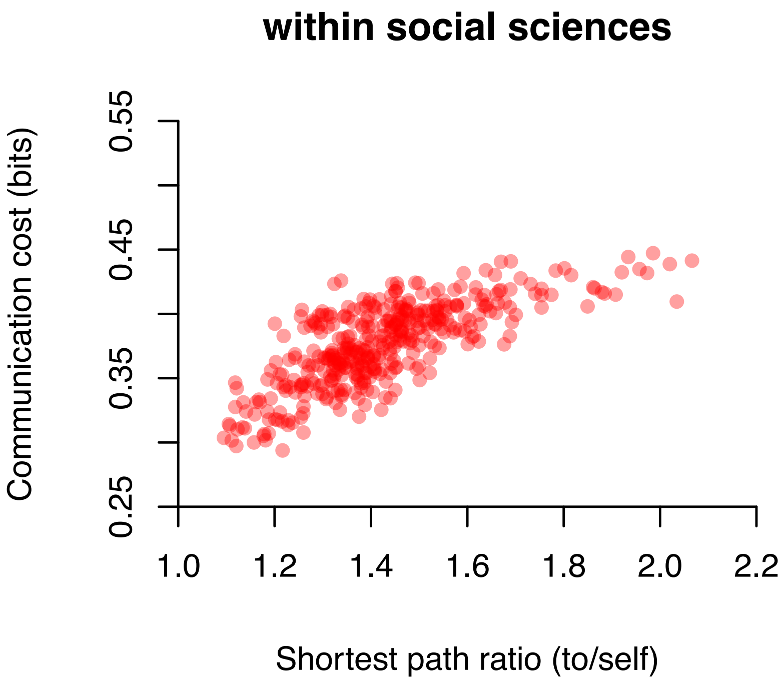


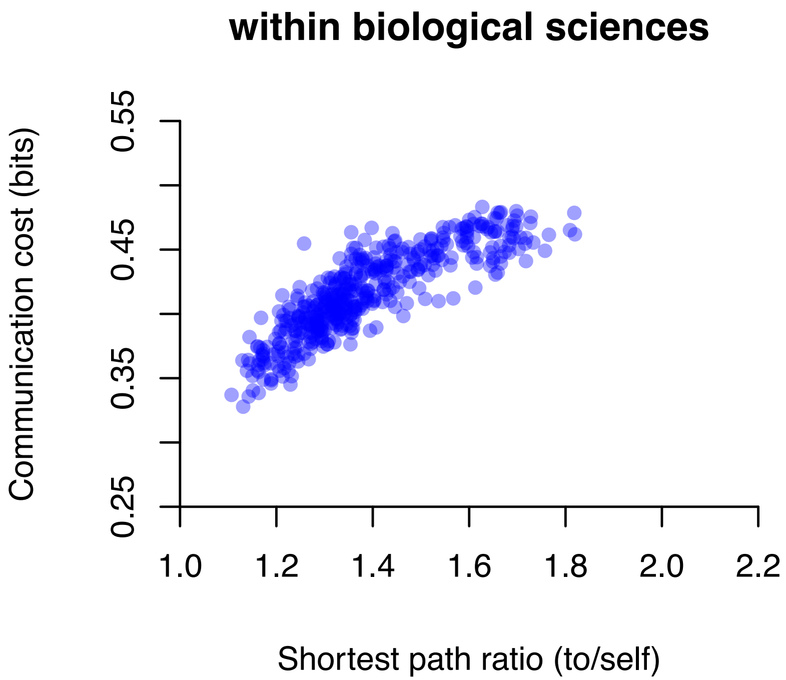
Jargon clustering
0.45
0.40
0.35
0.30
0.25
Distance
(jargon barrier)
environmental toxicology
HIV
plant pathogens
extracellular matrix
membrane cell biology
cytoskeleton
medical outcomes
kernel analysis
computational bayesian statistics
generalized linear models
survival analysis
marketing
time series analysis
portfolio theory
option pricing
strategic management
executive compensation
mergers and acquisitions
growth economics
consumer theory
gender and labor
unemployment
art education
sociology education
mathematics education
mental health
marital disruption
Childhood development
reproductive demography
teen sexual behavior
congressional elections
sociology of religions
international relations
social movements
US constitutional law
mass extinction
plant systematics
mitochondrial genetics
phylogenetic inference
bears
frugivory
pollination ecology
plant–herbivore interactions
mycorrhizal biology
riparian ecology
forest soil ecology
biological invasions
rainforest ecology
leaf ecology
bats
lizard thermoregulation
turtles
intertidal ecology
amphibian life history
daphnia
avian breeding ecology
waterfowl
voles
nesting ecology
landscape ecology

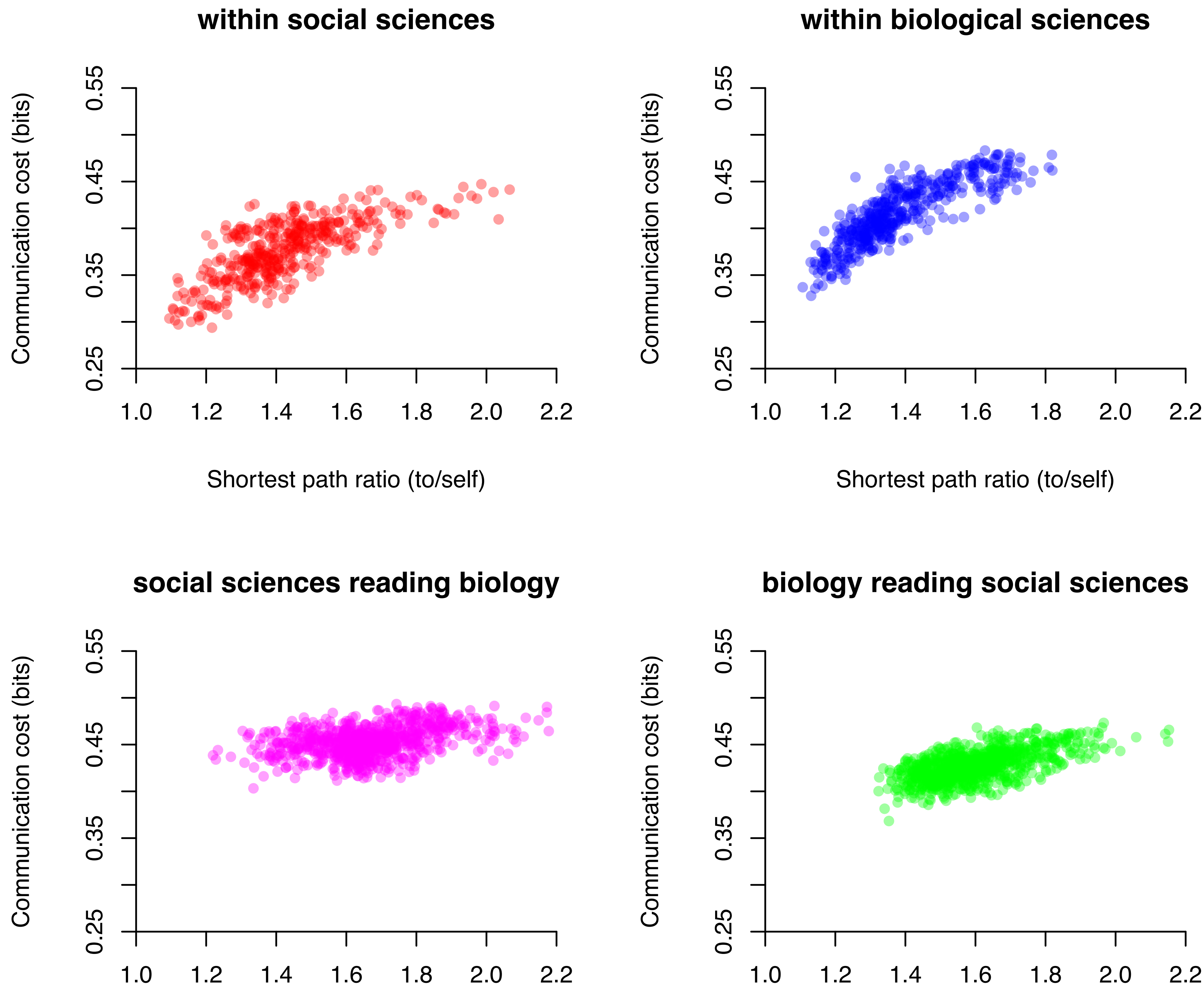
Decay in communication efficiency



$$E_{ij} = 1 - \beta(1 - e^{-\gamma d_{ij}})$$


within social sciences
Communication cost (bits)
0.25
0.35
0.45
0.55
1.0
1.2
1.4
1.6
1.8
2.0
2.2
Shortest path ratio (to/self)


within biological sciences
Communication cost (bits)
0.25
0.35
0.45
0.55
1.0
1.2
1.4
1.6
1.8
2.0
2.2
Shortest path ratio (to/self)

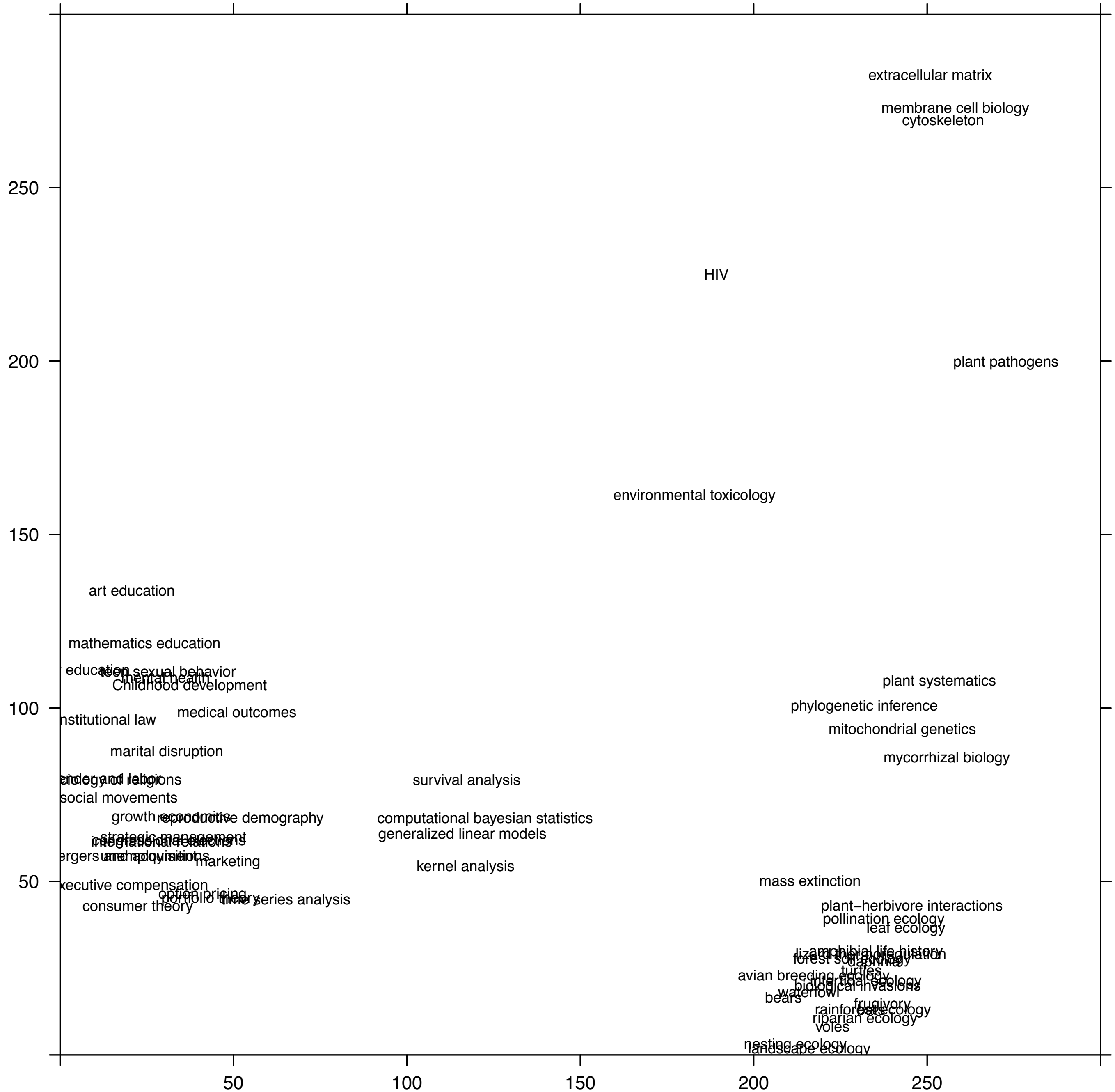


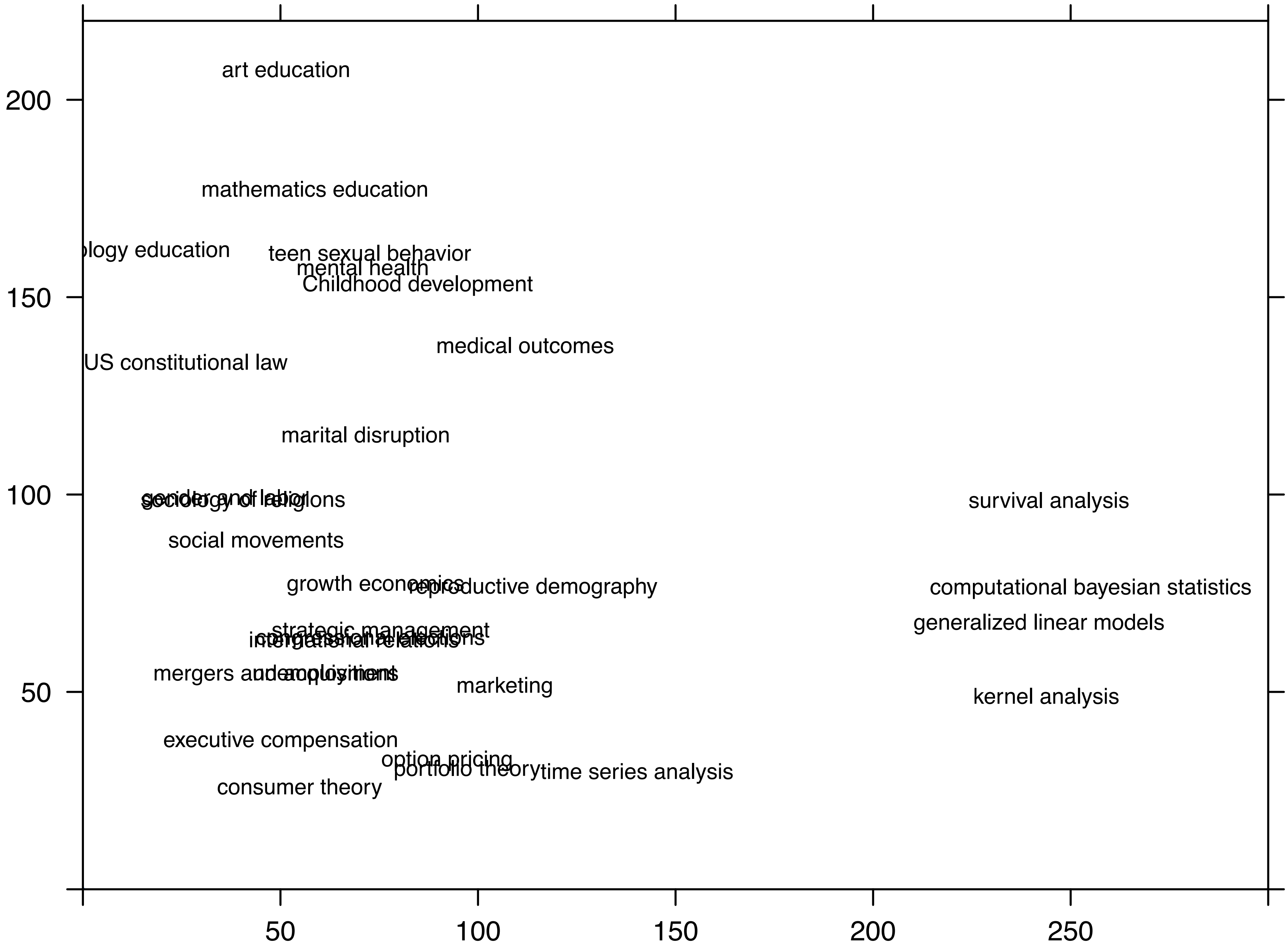
within social sciences
within biological sciences
social sciences reading biology
biology reading social sciences
Communication cost (bits)
Shortest path ratio (to/self)
0.25
0.35
0.45
0.55
1.0
1.2
1.4
1.6
1.8
2.0
2.2

Mapping ‘cultural holes’

Step 1: Compute x-y position of each discipline based on mean citation distance between disciplines.

- principle coordinate analysis (**PCoA**), dissimilarity matrix uses average shortest path


extracellular matrix
membrane cell biology
cytoskeleton
HIV
plant pathogens
environmental toxicology
art education
mathematics education
education
teen sexual behavior
mental health
Childhood development
medical outcomes
nstitutional law
marital disruption
survival analysis
social movements
reproductive demography
computational bayesian statistics
generalized linear models
marketing
kernel analysis
executive compensation
consumer theory
time series analysis
plant systematics
phylogenetic inference
mitochondrial genetics
mycorrhizal biology
mass extinction
plant−herbivore interactions
pollination ecology
leaf ecology
amphibian life history
turtles
avian breeding ecology
biological invasions
waterfowl
bears
frugivory
riparian ecology
voles
nesting ecology
landscape ecology
250
200
150
100
50
50
100
150
200
250



art education
mathematics education
ology education
teen sexual behavior
mental health
Childhood development
medical outcomes
US constitutional law
marital disruption
gender and labor
sociology of religions
survival analysis
social movements
growth economics
reproductive demography
computational bayesian statistics
generalized linear models
strategic management
congressional elections
international relations
mergers and acquisitions
marketing
kernel analysis
executive compensation
option pricing
portfolio theory
time series analysis
consumer theory
200
150
100
50
50
100
150
200
250

Mapping 'cultural holes'

Step 1: Compute x-y position of each discipline based on mean citation distance between disciplines.

- principle coordinate analysis (**PCoA**), dissimilarity matrix uses average shortest path

Step 2: Map jargon barriers on to the z-axis

$$\longrightarrow \; w^i_{P_{xy}} = \left(\sqrt{(x - F^i_x)^2 + (y - F^i_y)^2}\right)^{-\kappa}$$

$$\longrightarrow \; P_{xy} = \sum_{i \neq \max(\overrightarrow{w_{P_{xy}}})} \; \sum_{j \neq \max(\overrightarrow{w_{P_{xy}}}), j \neq i} w^i_{P_{xy}} w^j_{P_{xy}} \tilde{C}_{ij}$$

Map height

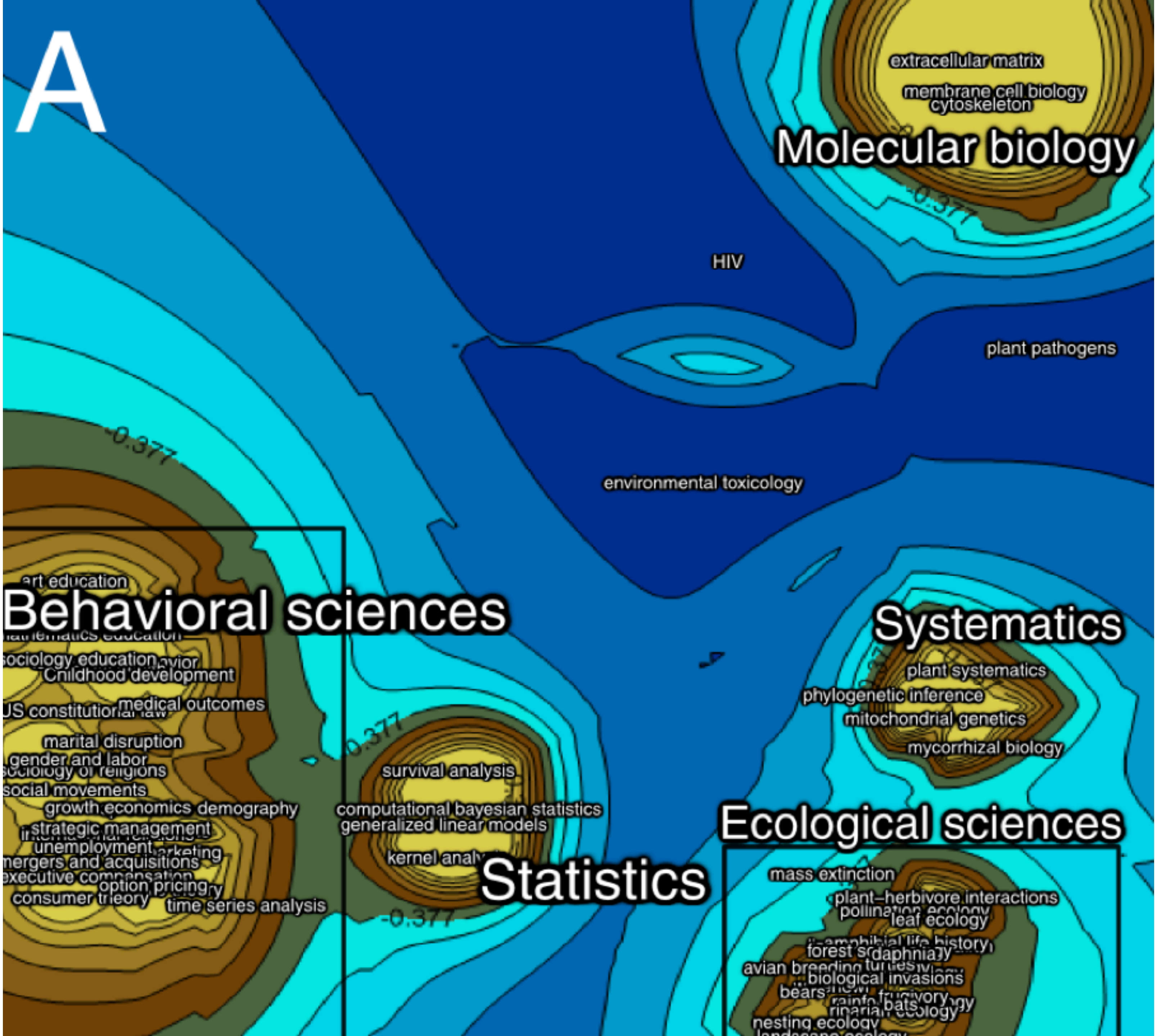
Weighting function

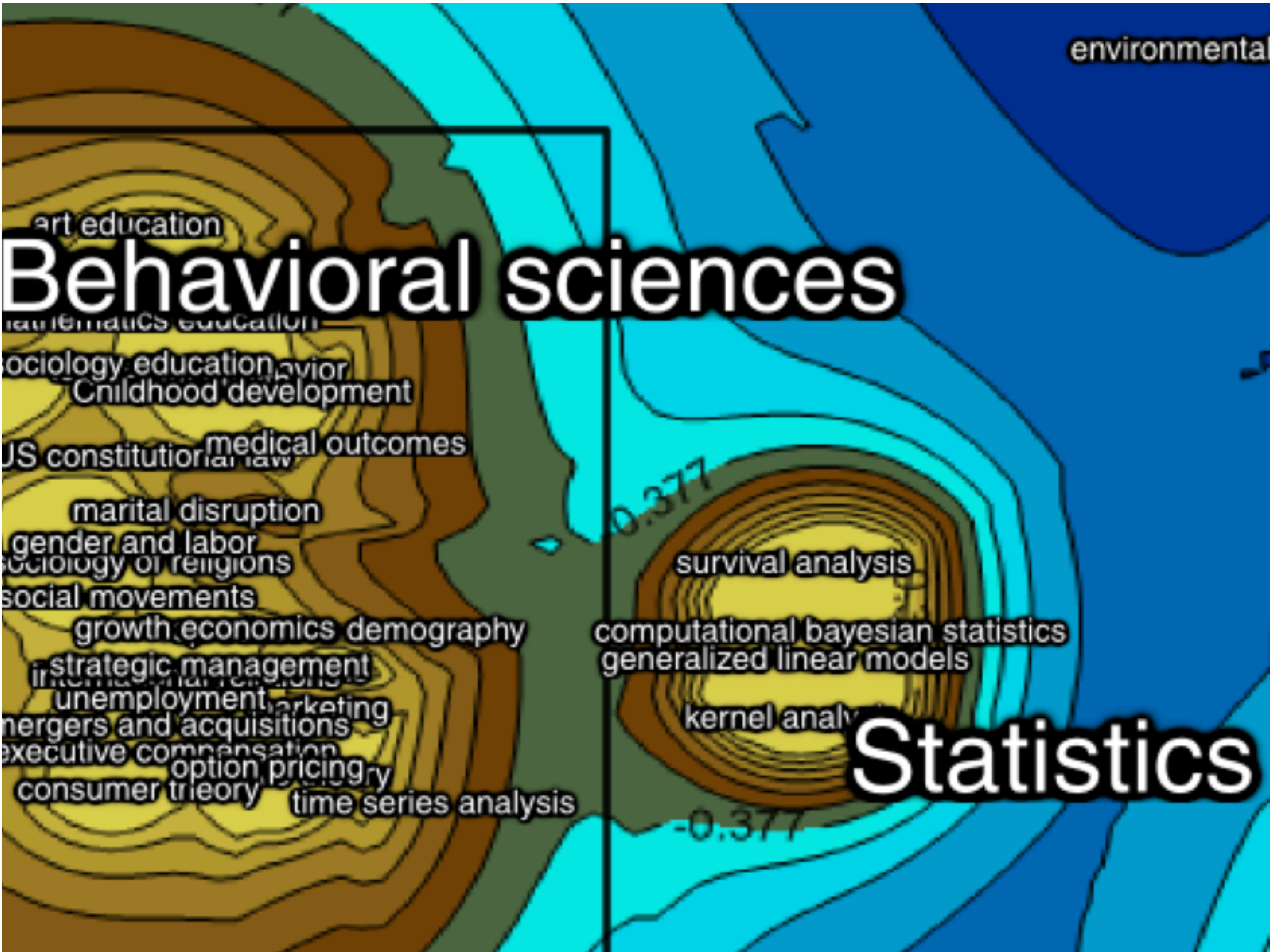
Jargon distance

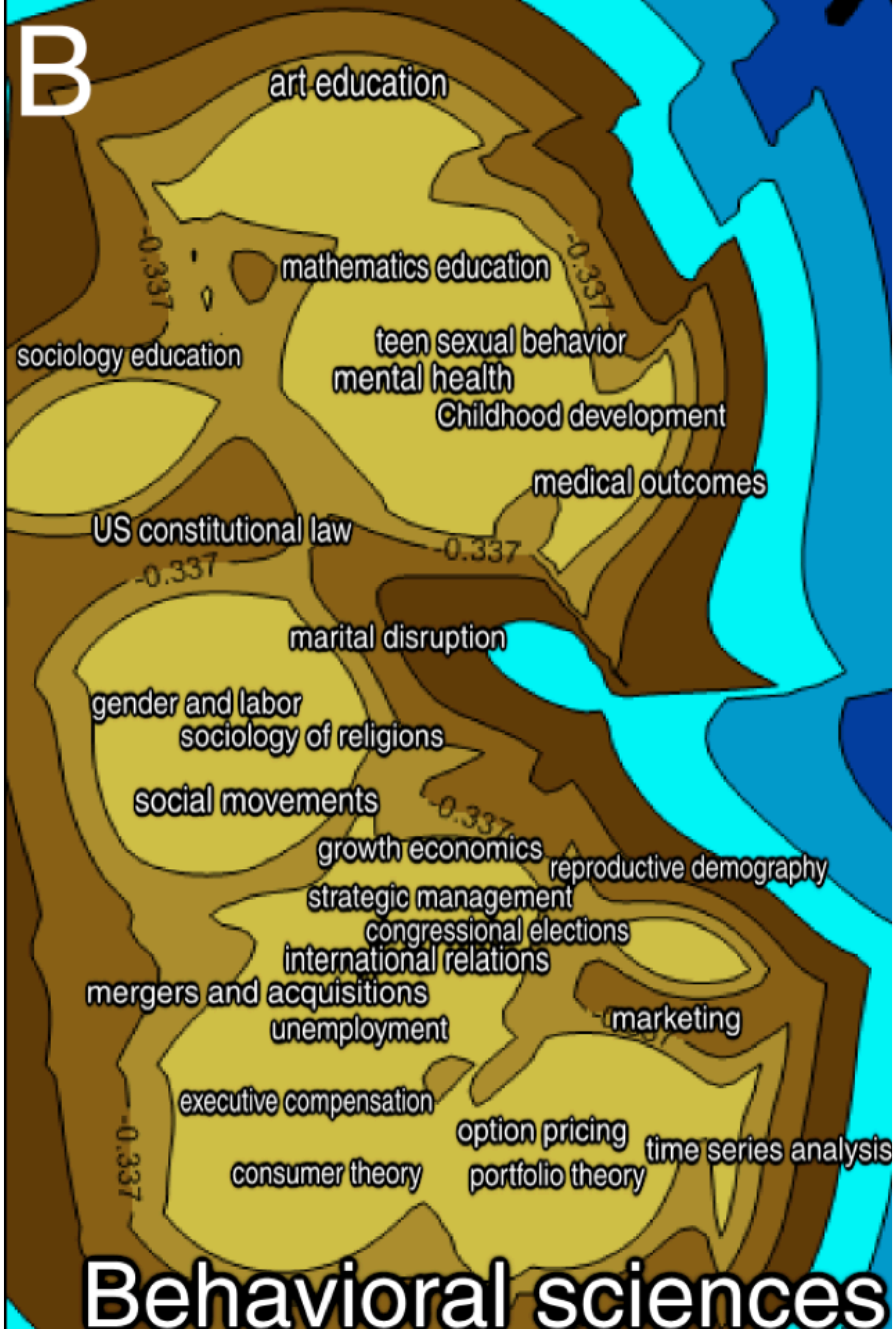
$$P_{xy} = \sum_{i \neq max(\overrightarrow{w_{P_{xy}}})} \sum_{j \neq max(\overrightarrow{w_{P_{xy}}}), j \neq i} \overrightarrow{w^i_{P_{xy}}} \overrightarrow{w^j_{P_{xy}}} J_{ij}$$

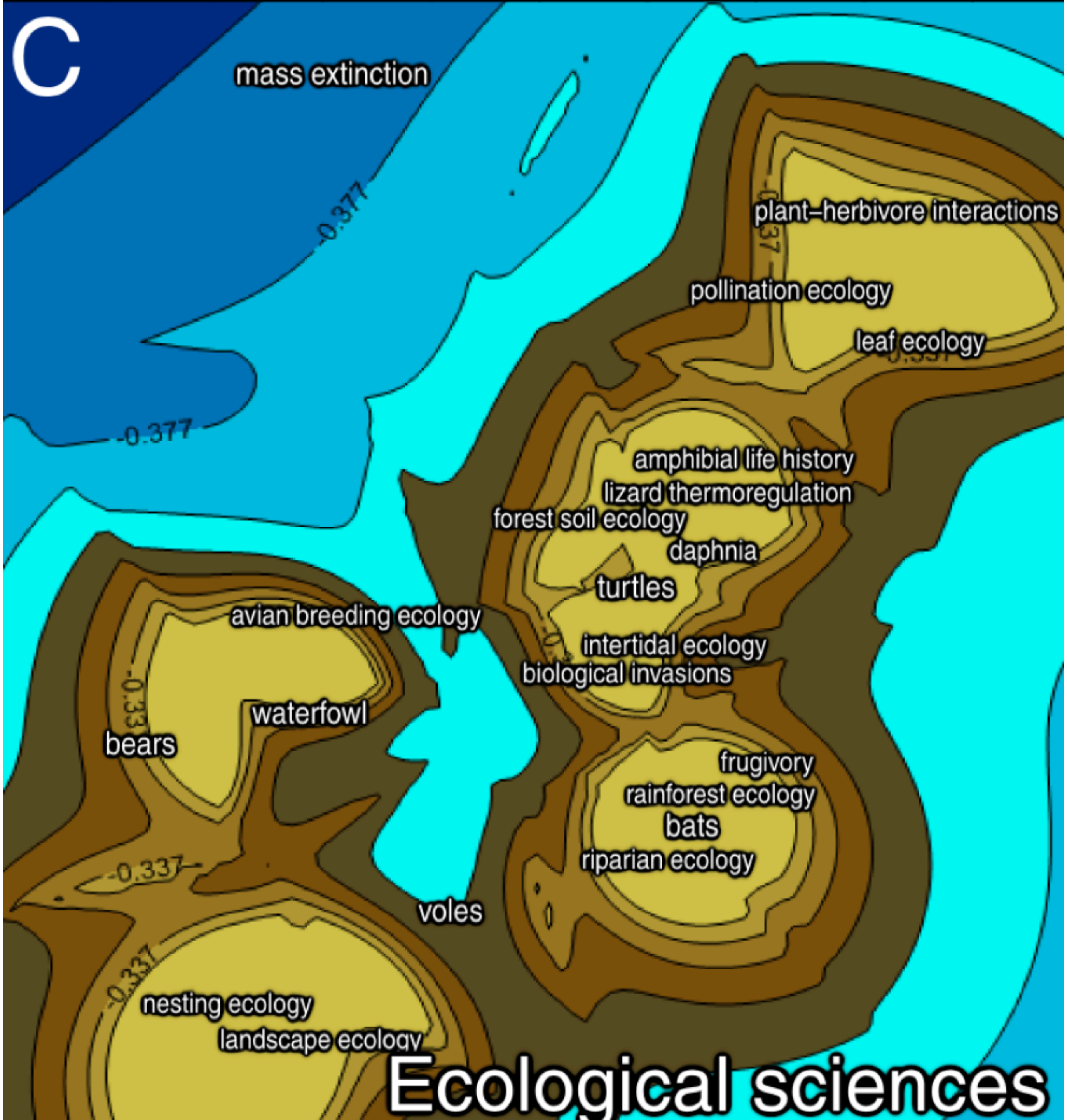
where

$$\overrightarrow{w_{P_{xy}}} = (\sqrt{(x - F^i_x)^2 + (y - F^i_y)^2})^{-\alpha}$$

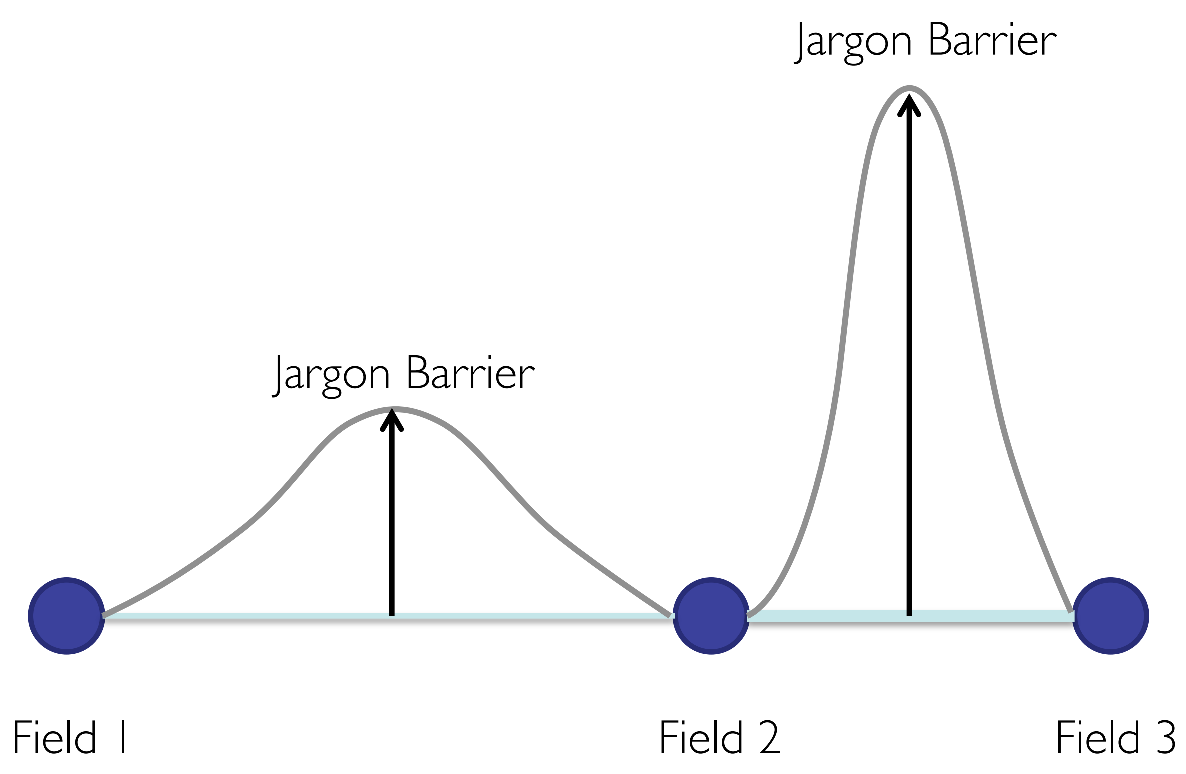

A
Molecular biology
extracellular matrix
membrane cell biology
cytoskeleton
-0.377
HIV
plant pathogens
environmental toxicology
Behavioral sciences
art education
sociology education
Childhood development
medical outcomes
marital disruption
gender and labor
social movements
growth economics demography
strategic management
unemployment
mergers and acquisitions
option pricing
consumer theory
time series analysis
Statistics
survival analysis
computational bayesian statistics
generalized linear models
Systematics
plant systematics
phylogenetic inference
mitochondrial genetics
mycorrhizal biology
Ecological sciences
mass extinction
plant–herbivore interactions
leaf ecology
daphnia
biological invasions
bears
nesting ecology


environmental
art education
Behavioral sciences
sociology education
Childhood development
medical outcomes
marital disruption
gender and labor
social movements
growth economics demography
strategic management
unemployment
mergers and acquisitions
option pricing
consumer theory
time series analysis
-0.377
survival analysis
computational bayesian statistics
generalized linear models
Statistics
-0.377


B
art education
mathematics education
-0.337
-0.337
teen sexual behavior
sociology education
mental health
Childhood development
medical outcomes
US constitutional law
-0.337
-0.337
marital disruption
gender and labor
sociology of religions
social movements
-0.337
growth economics
reproductive demography
strategic management
congressional elections
international relations
mergers and acquisitions
marketing
unemployment
executive compensation
-0.337
option pricing
time series analysis
consumer theory
portfolio theory
Behavioral sciences


C
mass extinction
-0.377
-0.377
plant–herbivore interactions
pollination ecology
leaf ecology
amphibial life history
lizard thermoregulation
forest soil ecology
daphnia
turtles
avian breeding ecology
intertidal ecology
biological invasions
waterfowl
bears
frugivory
rainforest ecology
bats
riparian ecology
-0.337
voles
-0.337
nesting ecology
landscape ecology
Ecological sciences

Mapping Jargon in Science



paper screening

arXiv.org

information retrieval

Summary

- Modeled communication barriers for spanning ‘cultural holes’
- Found that structural and cultural information don’t perfectly align
- Communicative efficiency decays with citation distance specific to fields
- Ecological sciences balkanized by jargon more so than social sciences

Future Directions

- Continue to refine the model and our assumptions of the model
- Generalize model to other forms of cultural communication
- Apply to other disciplines and other non-scientific corpora
- Look at the evolution of cultural and structural holes
- Make the code available so that others can build these kinds of maps

Acknowledgements

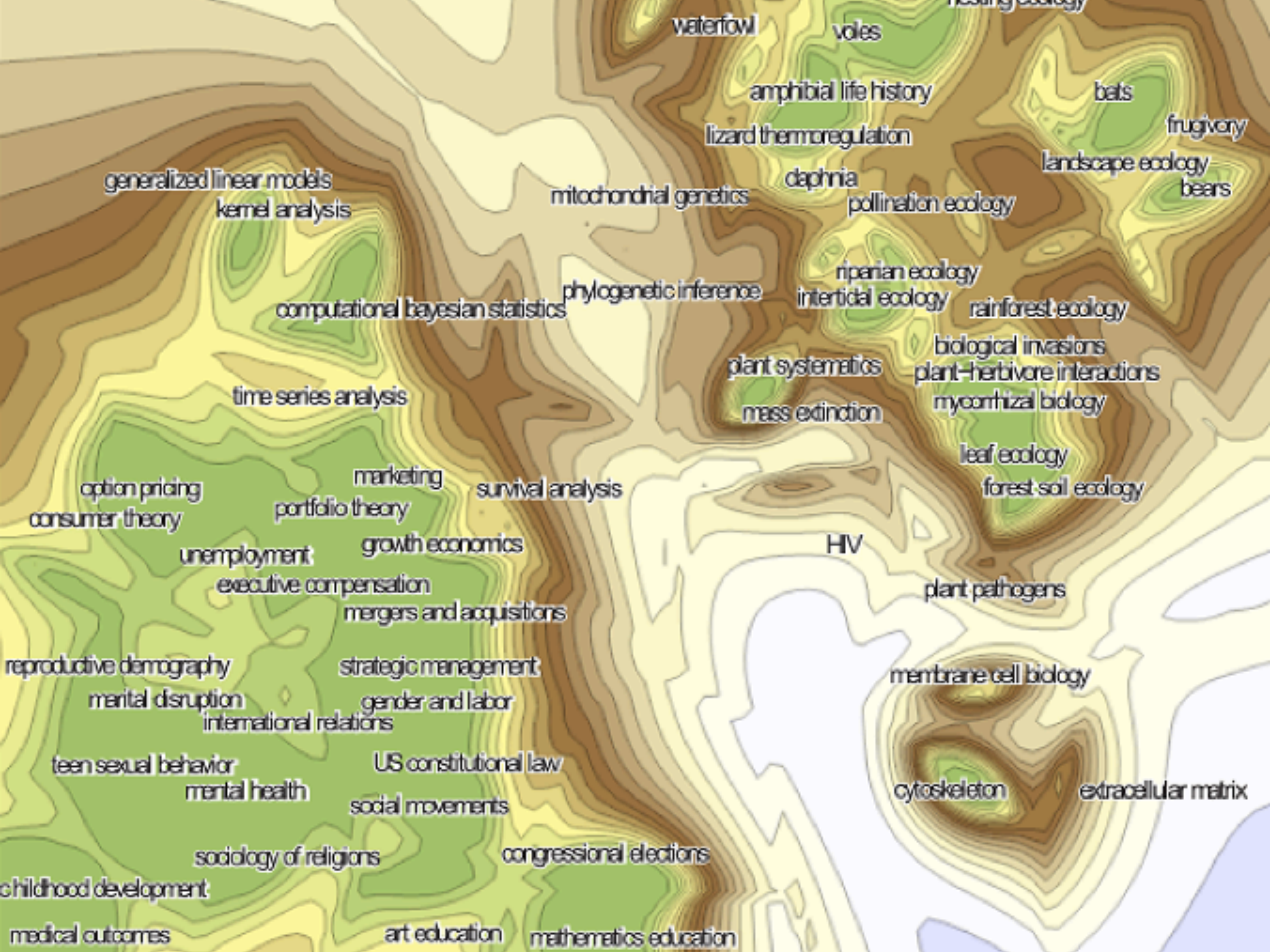
Jacob Foster, Department of Sociology, UCLA

Daril Vilhena, Amazon Data Science Division

Carl Bergstrom, Department of Biology, University of Washington

James Evans, Department of Sociology, University of Chicago

Martin Rosvall, Department of Physics, Umea University



waterfowl
voles
amphibial life history
bats
frugivory
lizard thermoregulation
landscape ecology
bears
daphnia
generalized linear models
mitochondrial genetics
pollination ecology
kernel analysis
riparian ecology
phylogenetic inference
intertidal ecology
rainforest ecology
computational bayesian statistics
biological invasions
plant systematics
plant–herbivore interactions
time series analysis
mycorrhizal biology
mass extinction
leaf ecology
marketing
option pricing
survival analysis
forest soil ecology
consumer theory
portfolio theory
growth economics
HIV
unemployment
executive compensation
plant pathogens
mergers and acquisitions
reproductive demography
strategic management
membrane cell biology
marital disruption
gender and labor
international relations
teen sexual behavior
US constitutional law
mental health
cytoskeleton
extracellular matrix
social movements
sociology of religions
congressional elections
childhood development
medical outcomes
art education
mathematics education