Exploring futures studies scholarly literature

T. Gauthier, J. Subtil, and P.-A. Fonta

To reveal the **social structures** that bond researchers and practitioners together
as well as
the specialized language and **subject matters** featured in the scholarly literature

Materials and methods

Scholarly journals

17 peer-reviewed journals identified by the World Futures Studies Federation:

- ❏ 10 indexed by Web of Science
- ❏ 5 not indexed by Web of Science
- ❏ 2 indexed by Web of Science but not included in the analysis (> 15-year old publications)

10 indexed scholarly journals

Foresight

Futures

Info

International Journal of Foresight and Innovation Policy

Long Range Planning

10 indexed scholarly journals

Technological Forecasting and Social Change

The International Journal of Forecasting

The Journal of Forecasting

Time & Society

NanoEthics

5 scholarly journals not indexed

European Journal of Futures Research

Journal of Evolution and Technology

Journal of Futures Studies

Policy Futures in Education

World Future Review

2 indexed scholarly journals not included

Futuribles (1975-99)

World Futures (1981-84)

Publications crawl

Using open-source, Python-based Scrapy framework

~ 37GB of HTML crawled

Data processing

Using large-scale data processing engine Apache Spark

Steps:

1. Scraping fields (incl. title, authors, abstract, etc) from HTML
2. Entity (incl. publications, authors, institutions, etc) creation
3. Entity disambiguation
4. CSV export

Entity volume

16,958 futures studies publications

14,621 authors

206,488 publications cited by futures studies publications

Data analysis

Using Tableau Software to:

1. Compute metrics
2. Create data visualizations

Social network analysis

Using open source Gephi software to create networks

Nodes: futures studies publications' authors

Edges: futures studies publications co-authorship

Edge weight: number of coauthored futures studies publications

Social network analysis

Betweenness centrality

Number of shortest paths from all nodes to all others that pass through that node

Modularity (community detection)

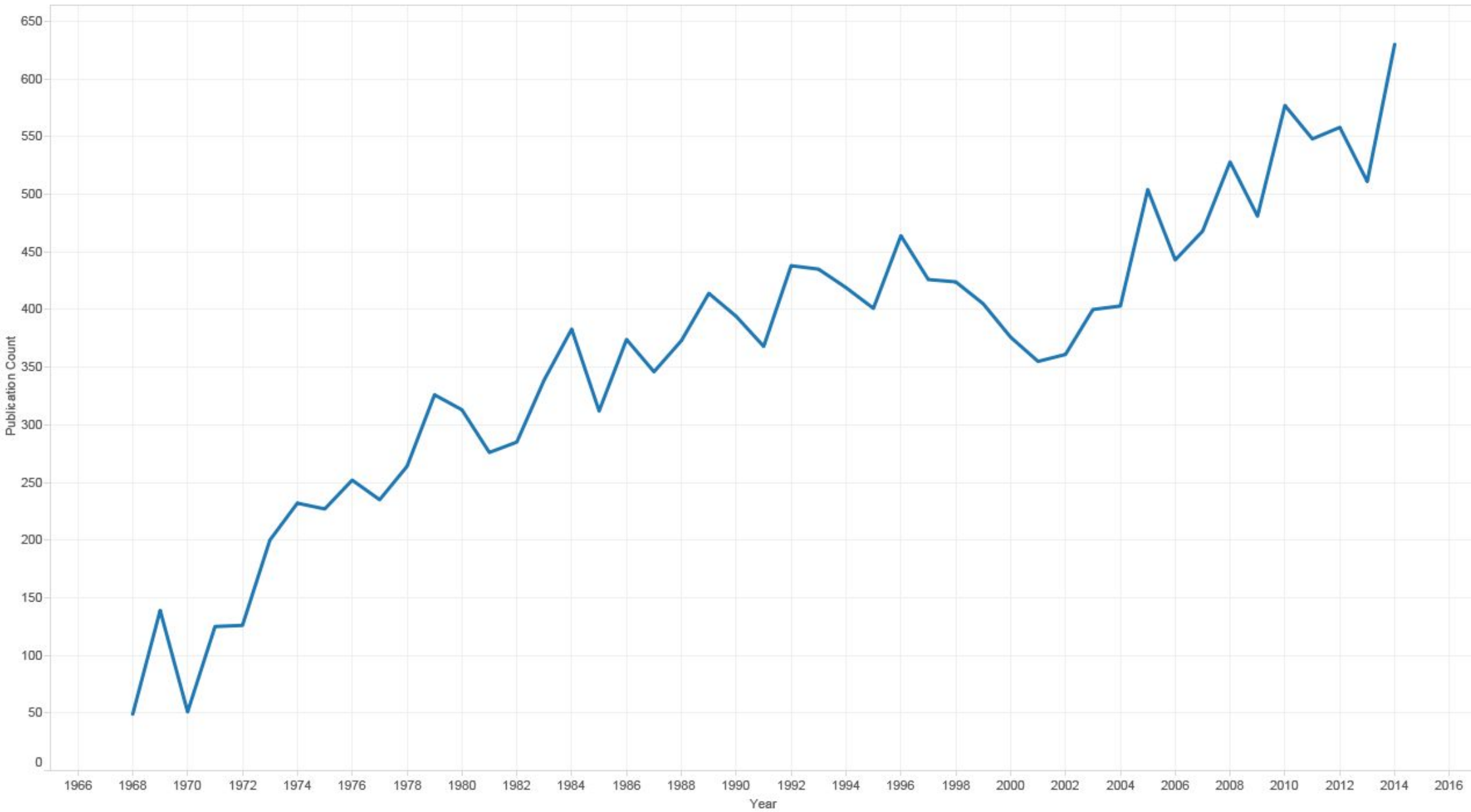Limitations: nondeterministic, community membership may vary slightly

Giant component

Keep the biggest component of the graph
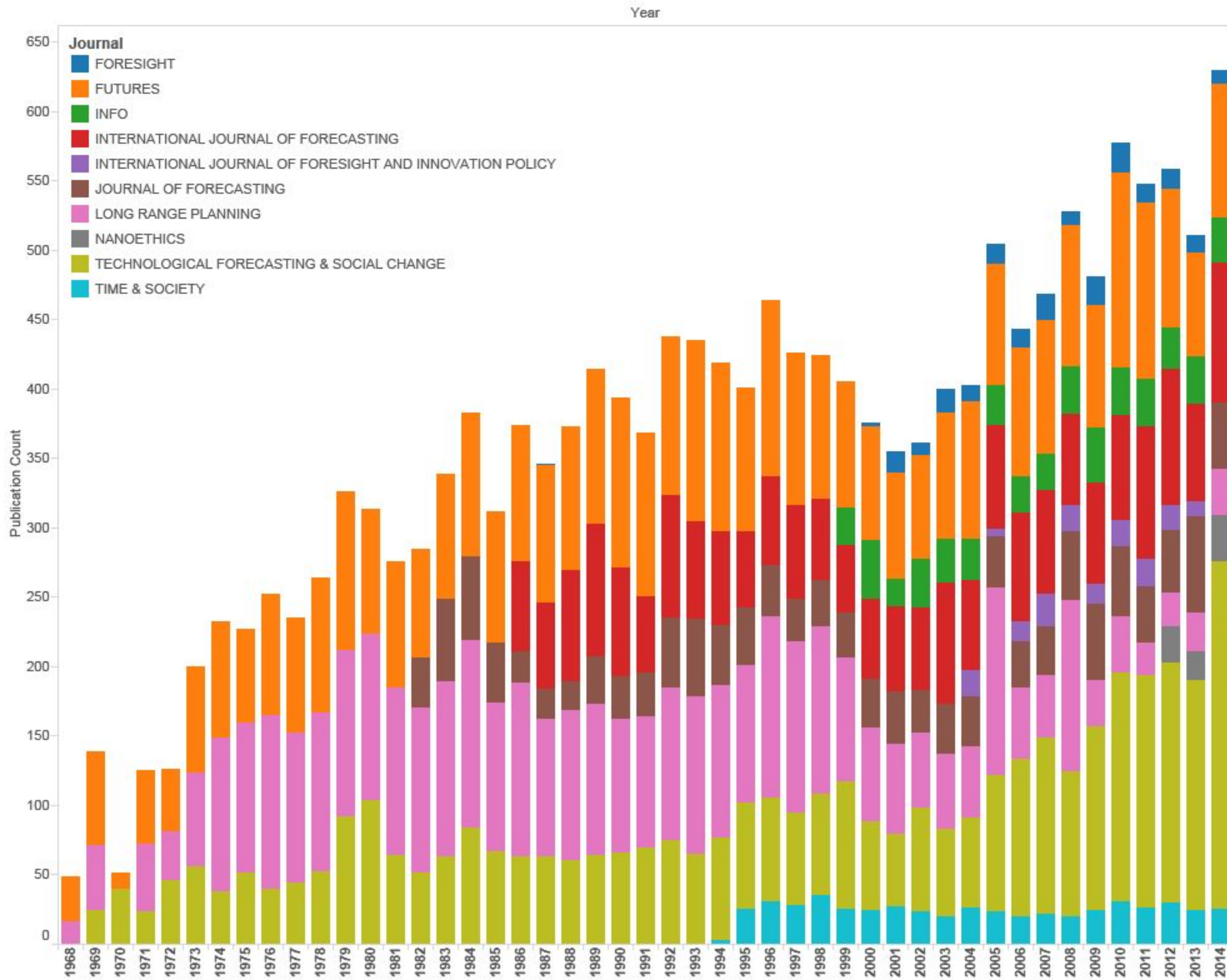
Filter out unconnected authors and small communities

Results

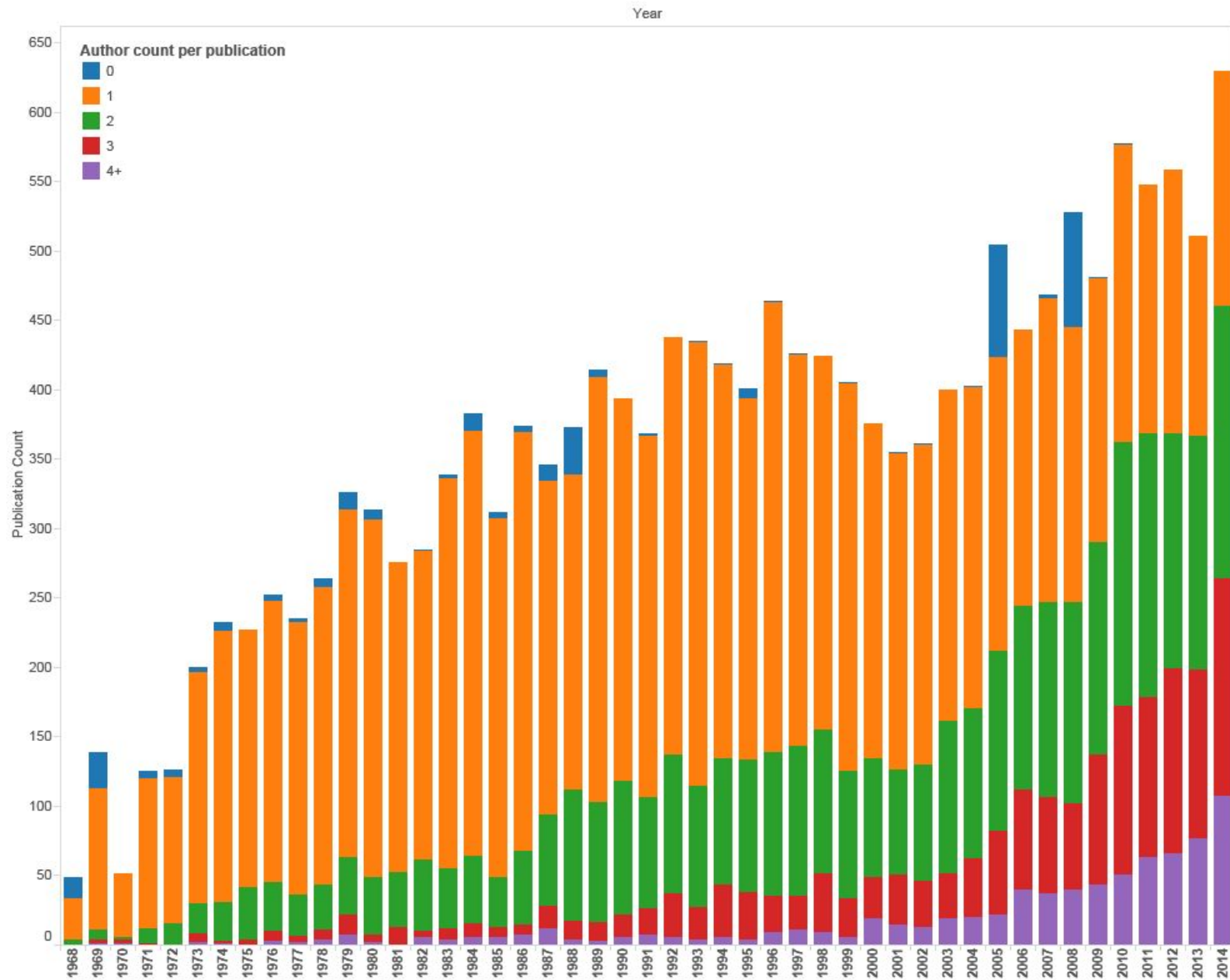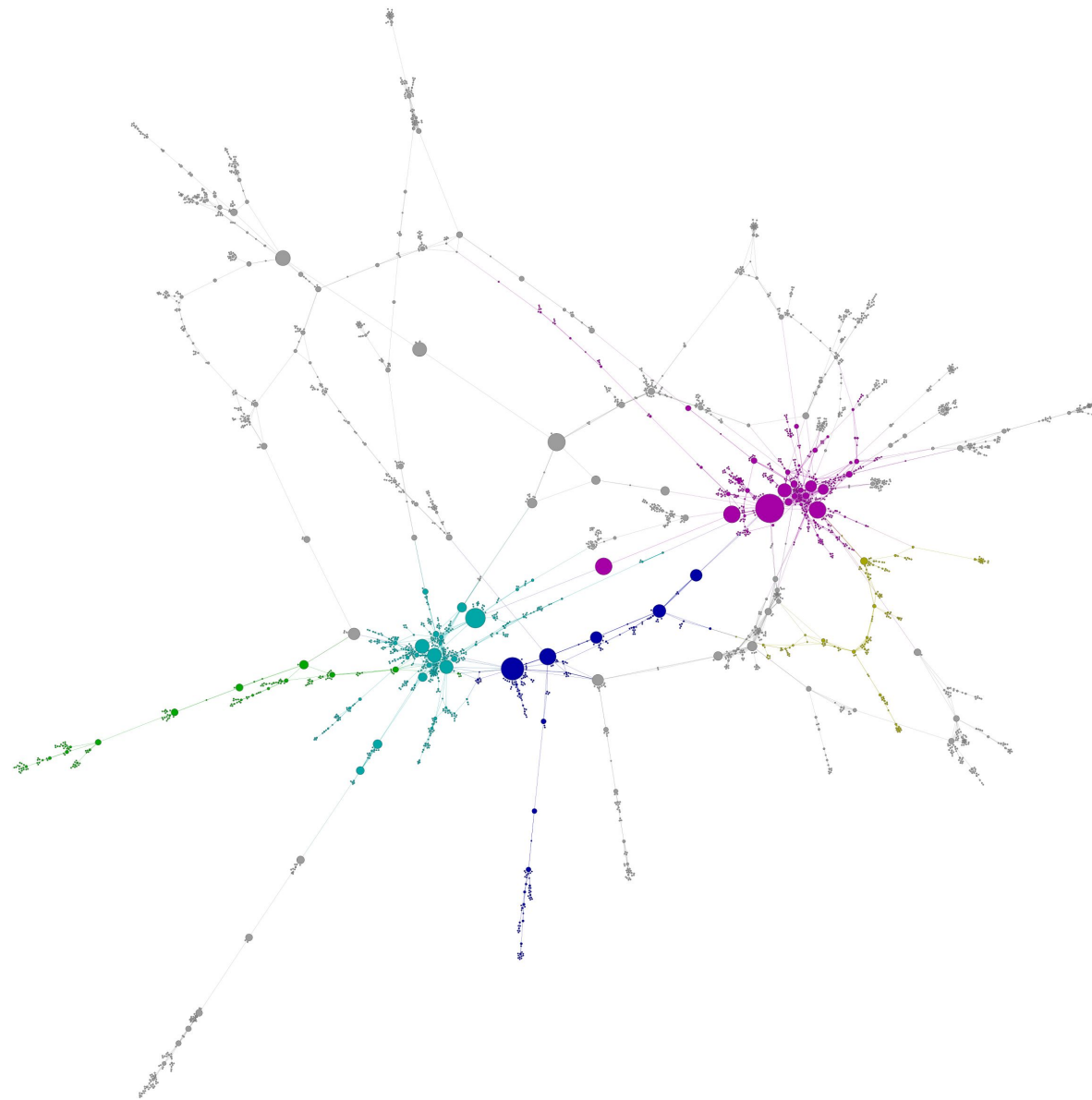# Futures publications (1968-2014)

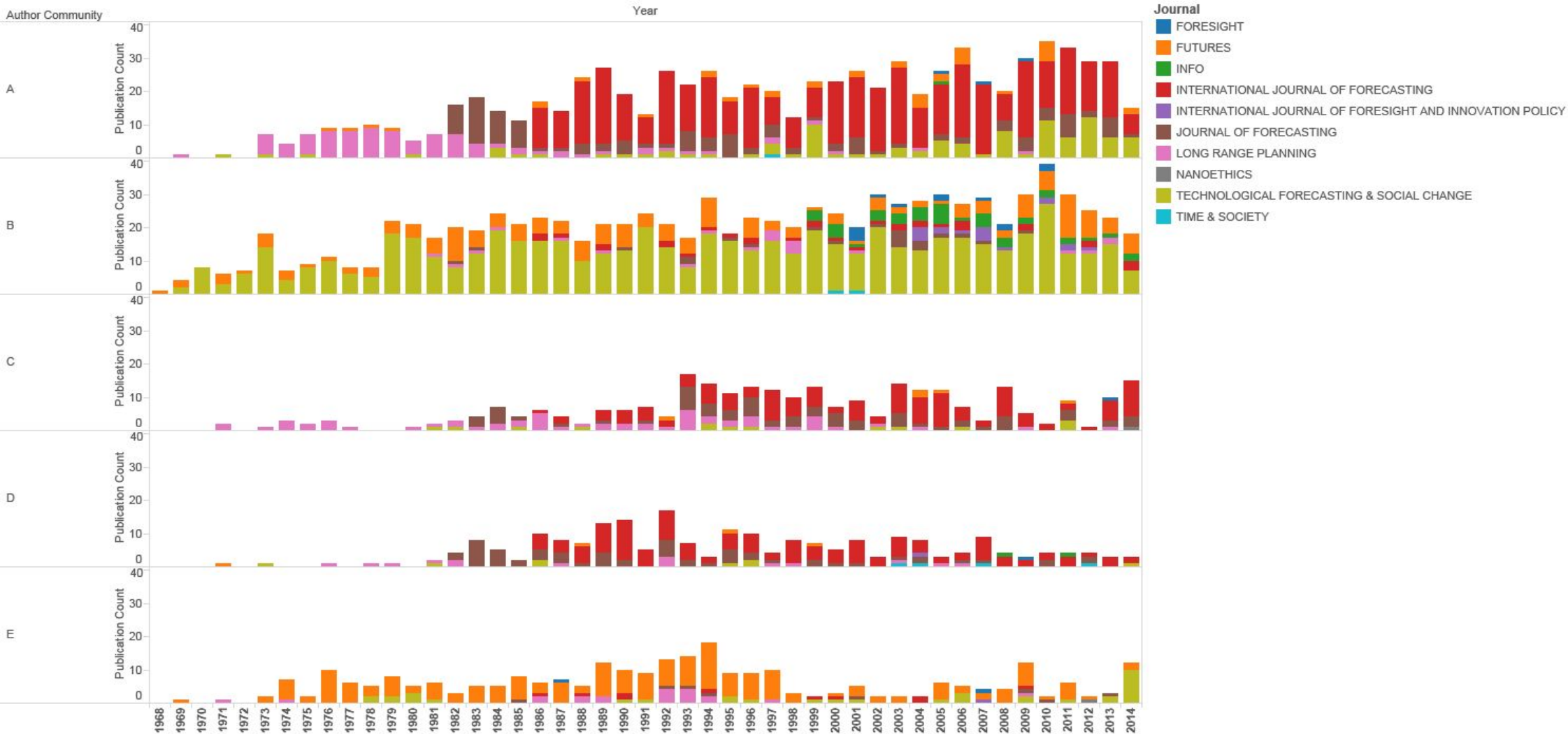# Futures publications per journal (1968-2014)

Co-authorship patterns

Author networks and communities

Publication patterns by author community

Author keyword analysis per community

Community A

forecasting

exponential smoothing

judgmental forecasting

time series

judgment

combining forecasts

m-competition

prediction intervals

forecast accuracy

state space models

accuracy

judgement

markov chain monte carlo

model selection

delphi

forecast evaluation

forecasting practice

scenario planning

seasonality
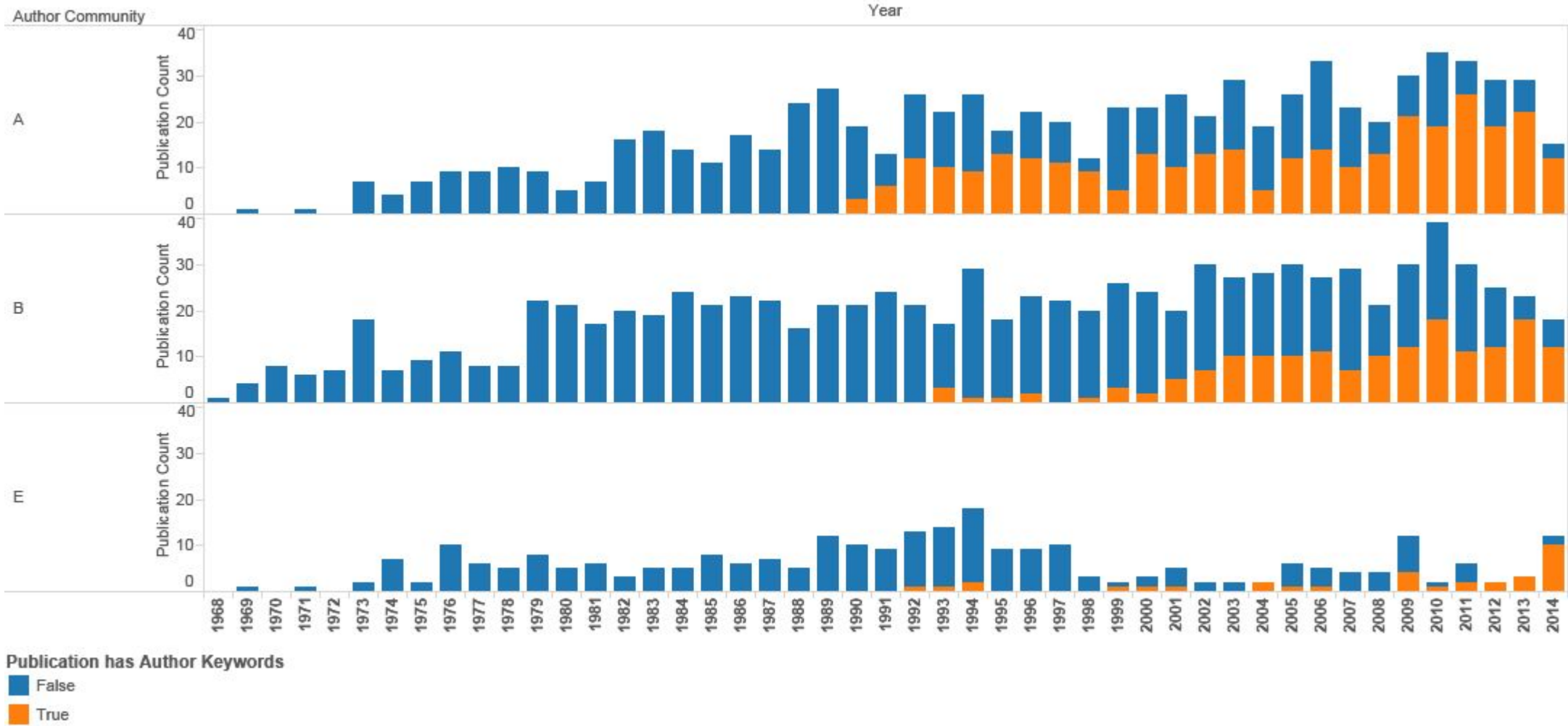
% of Total Publication Count

2.236%   12.141%

Community A author keywords

Methods

Exponential smoothing (8%)

Time series/time-series forecasting (10%)

Judgmental forecasting (17%)

Scenarios (2%)

Community A author keywords over time

Methods

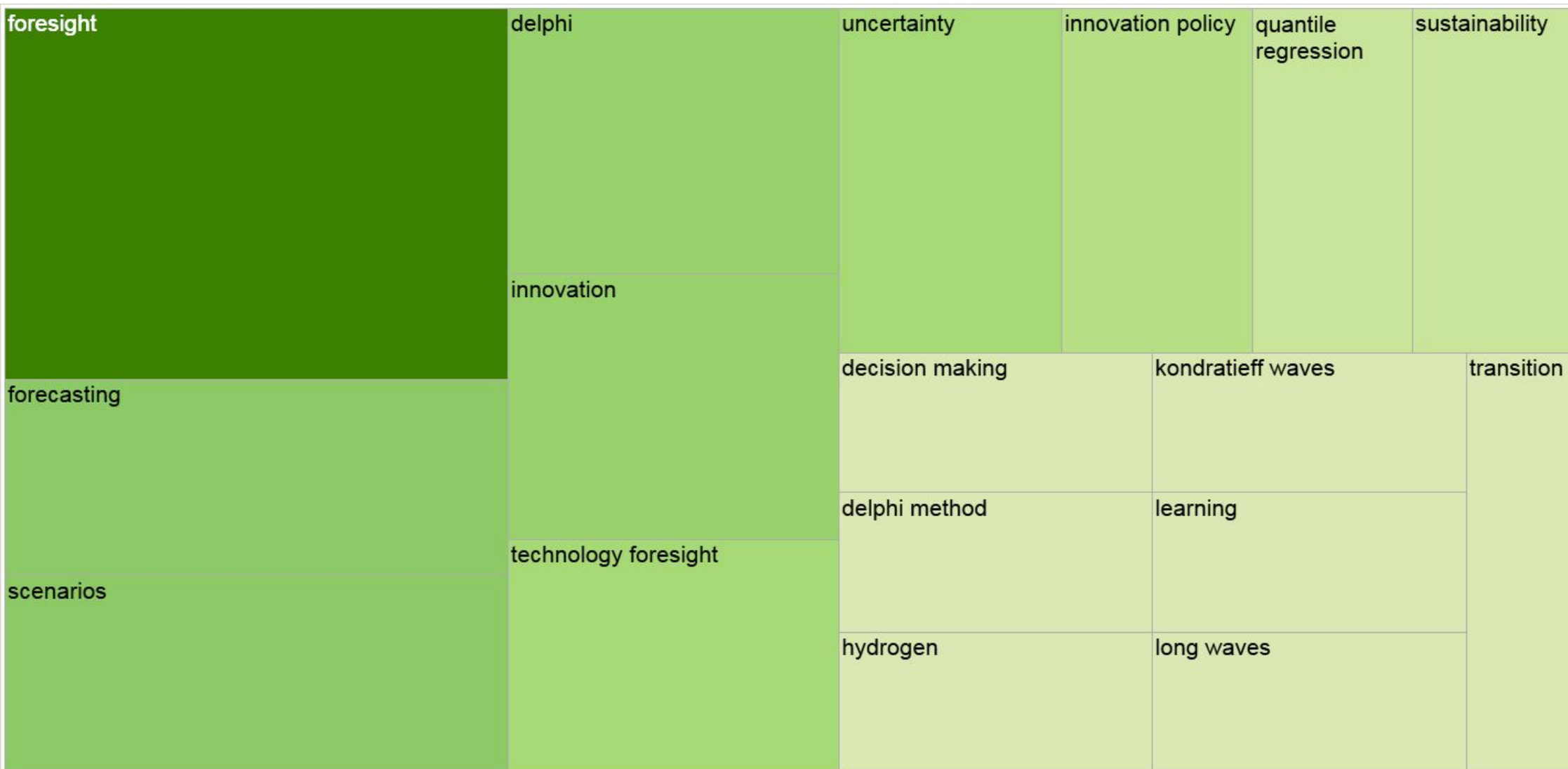Exponential smoothing (1990-94: 10%, 1995-99: 0%, 2000-04: 13%, 2005-09: 9%, 2010-14 8%)

Time series/time-series forecasting (1990-94: 15%, 1995-99: 24%, 2000-04: 13%, **2005-09: 0%, 2010-14 4%**)

Judgmental forecasting (1990-94: 17%, 1995-99: 36%, 2000-04: 16%, 2005-09: 10%, 2010-14 10%)

Scenarios (1990-94: 0%, 1995-99: 0%, 2000-04: 0%, 2005-09: 0%, **2010-14 4%**)

Community B

foresight

delphi

innovation

technology foresight

forecasting

scenarios

uncertainty

innovation policy

quantile regression

sustainability

decision making

kondratieff waves

transition

delphi method

learning

hydrogen

long waves

% of Total Publication Count

2.410%

10.241%

Community B author keywords

Methods

Scenarios/scenario planning (7%)

Delphi (7%)

Exponential smoothing (1%)

Topics

Innovation (23%)

Sustainability (7%)
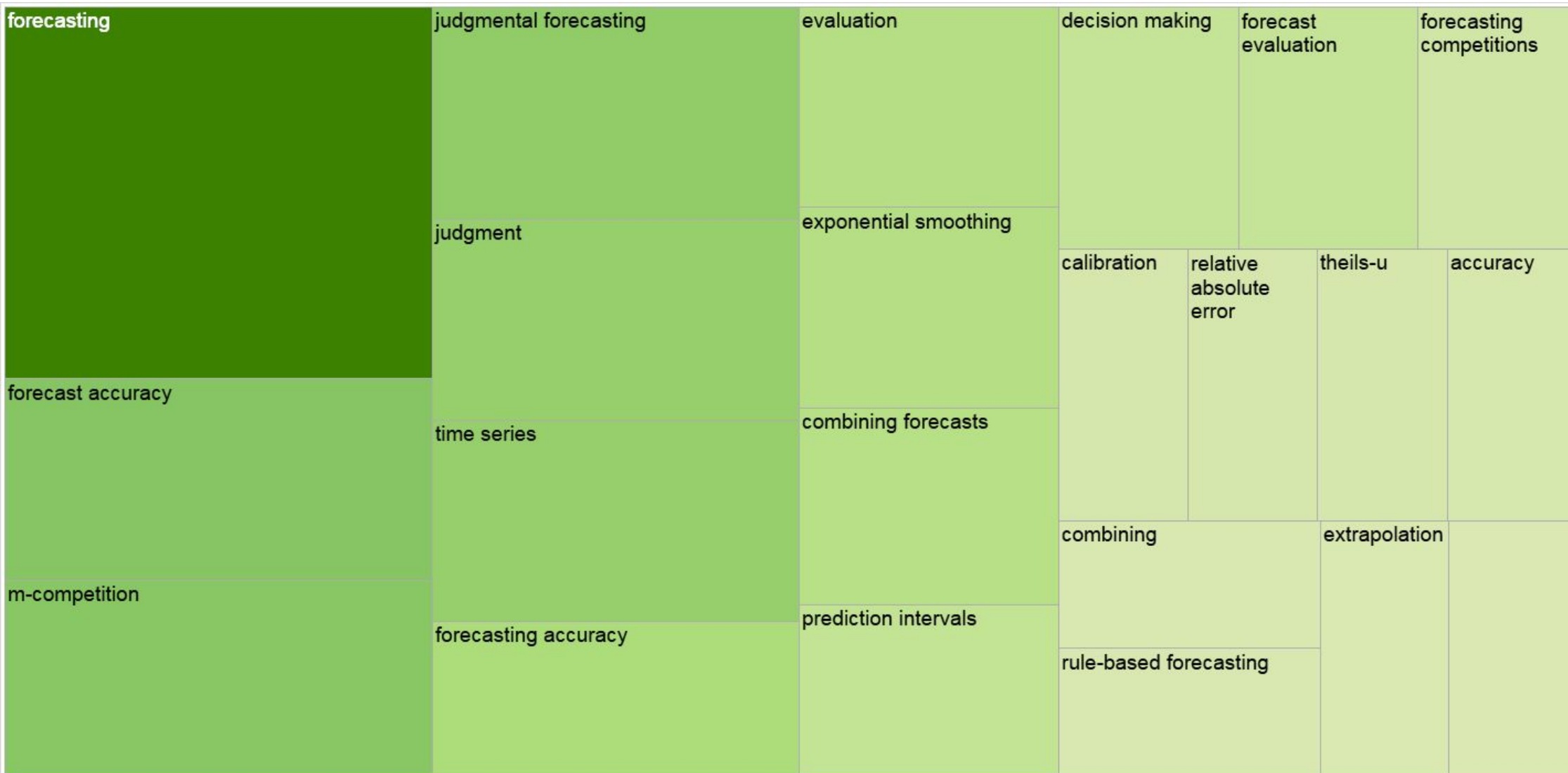
Community B author keywords over time

Methods

Scenarios/scenario planning (1990-94: 0%, 1995-99: 0%, 2000-04: 0%, 2005-09: 4%, 2010-14 12.7%)

Delphi (1990-94: 0%, 1995-99: 0%, 2000-04: 0%, **2005-09: 6%, 2010-14: 7%**)

Topics

Innovation (1990-94: 0%, 1995-99: 0%, **2000-04: 17.7%, 2005-09: 8%, 2010-14: 9.9%**)
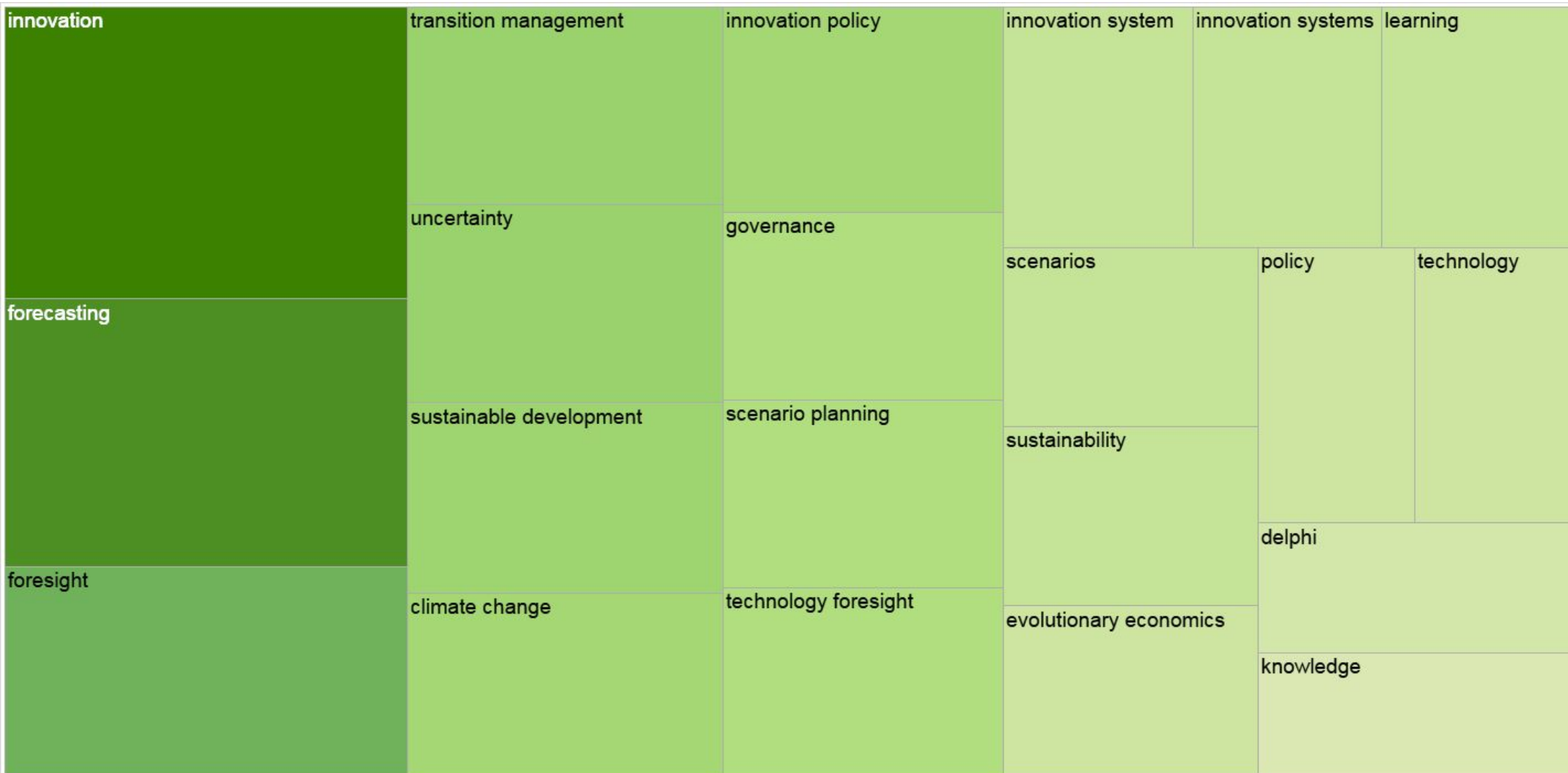
Community A references

forecasting

judgmental forecasting

judgment

time series

forecasting accuracy

forecast accuracy

m-competition

evaluation

exponential smoothing

combining forecasts

prediction intervals

decision making

forecast evaluation

forecasting competitions

calibration

relative absolute error

theils-u

accuracy

combining

extrapolation

rule-based forecasting

% of Total Main Publication Count
10.24%   51.21%

Community A references' author keywords

Methods

M-competition (27%)

time series/time-series forecasting (up to 68%)

Community B references

% of Total Main Publication Count
7.44% — 22.79%

Community B references' author keywords

Methods

Uncertainty (12%)

Topics

Sustainability/sustainable development (20%)

Transition management (12%)

Climate change (11%)

Challenges and limitations

Data crawl and analysis

Availability of journals/publications/author keywords

"Intent" of <u>author</u> keywords (intent: why selected these vs. those keywords? keywords terminology may evolve over time? susceptible to fashion?)

Large crawls → Web of Science session expires → need to handle resets

Crawled data requires in-depth preparation to get to analysis-ready entities

Disambiguation

Author disambiguation

Uncovered cases: true homonyms, misspelled names, disambiguation on ambiguous authors e.g. J. Smith

Publication disambiguation

Several publications with same title:

- Some are definitely distinct material (e.g. journal article vs. book chapter)
- Some look like re-editions
- Some might be duplicates

Potential next steps

Potential next steps

Institution-level analysis

Text mining analysis (title, abstract, full-text when available)

Interviews/narratives/data combination

Futures studies ontology/ontologies

Thank you

thomas.gauthier@hesge.ch