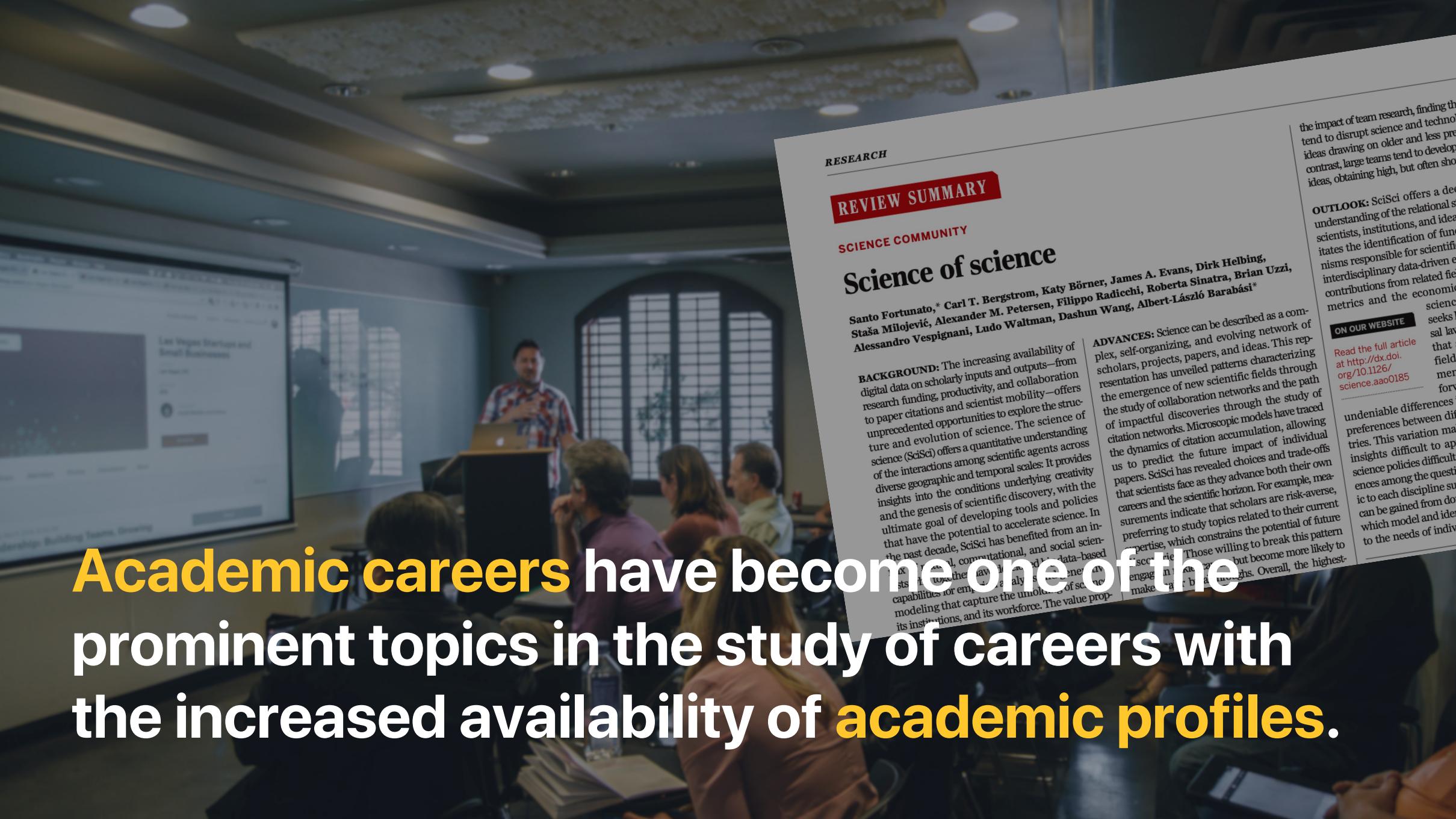


How to achieve career success is a long-standing research question that has been studied in various social science disciplines.



A photograph of a man giving a presentation to an audience in a lecture hall. He is standing at a podium, gesturing with his hands as he speaks. The audience is seated in rows of chairs, facing him. In the foreground, a large screen displays a presentation slide with text and graphics.

Academic careers have become one of the prominent topics in the study of careers with the increased availability of academic profiles.

RESEARCH

REVIEW SUMMARY

SCIENCE COMMUNITY

Science of science

Santo Fortunato,* Carl T. Bergstrom, Katy Börner, James A. Evans, Dirk Helbing, Staša Milojević, Alexander M. Petersen, Filippo Radicchi, Roberta Sinatra, Brian Uzzi, Alessandro Vespignani, Ludo Waltman, Dashun Wang, Albert-László Barabási*

BACKGROUND: The increasing availability of digital data on scholarly inputs and outputs—from research funding, productivity, and collaboration to paper citations and scientist mobility—offers unprecedented opportunities to explore the structure and evolution of science. The science of science (SciSci) offers a quantitative understanding of the interactions among scientific agents across diverse geographic and temporal scales; It provides insights into the conditions underlying creativity and the genesis of scientific discovery, with the ultimate goal of developing tools and policies that have the potential to accelerate science. In the past decade, SciSci has benefited from an interdisciplinary, computational, and social scientific approach to the analysis and modeling of scientific profiles. Those willing to break this pattern of engagement in science can make more contributions to its institutions, and its workforce. The value of the impact of team research, finding ideas drawing on older and less pre-contrast, large teams tend to develop ideas, obtaining high, but often short-term success.

ADVANCES: Science can be described as a complex, self-organizing, and evolving network of scholars, projects, papers, and ideas. This representation has unveiled patterns characterizing the emergence of new scientific fields through the study of collaboration networks and the path of impactful discoveries through the study of citation networks. Microscopic models have traced the dynamics of citation accumulation, allowing us to predict the future impact of individual papers. SciSci has revealed choices and trade-offs that scientists face as they advance both their own careers and the scientific horizon. For example, measurements indicate that scholars are risk-averse, preferring to study topics related to their current expertise, which constrains the potential of future research. Overall, the highest-impact research tends to disrupt science and technology, drawing on older and less pre-existing ideas, obtaining high, but often short-term success.

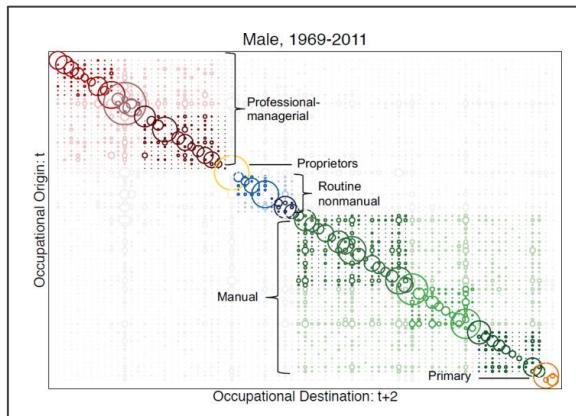
OUTLOOK: SciSci offers a deeper understanding of the relational scientists, institutions, and ideas that facilitates the identification of mechanisms responsible for scientific contributions from related fields and the economic

ON OUR WEBSITE
Read the full article
[at <http://dx.doi.org/10.1126/science.aoa0185>](http://dx.doi.org/10.1126/science.aoa0185)

undeniable differences in scientific preferences between disciplines. This variation makes it difficult to apply science policies that are appropriate to each discipline such as those that can be gained from doing research which model and identify the needs of individual scientists.

Analyzing Academic Careers

New analytical demands in today's boundaryless career world



B. F. Jarvis and X. Song, 2017

Point-to-Point Transition

Fails to Capture Long-term Impact



Table 3 Comparing school success

	S high grades, L not	Both high grades	Both no high grades	L high grades, S not
Secondary education	3	8	2	4
Master's degree	4	1	10	2
Ph.D.	2	0	7	3

S = Talent who stayed; L = Talent who left.

Table 4 Family conditions

	Talents who stayed			Talents who left		
	Male	Female	All	Male	Female	All
Have children	10	3	13	8	8	16
Supported by partner	8	5	13	2	1	3
S yes, L no			Both yes	Both no		
Have children	1	10	0	0	2	2
Supported by partner	9	3	1	1	0	0

B. V. Balen et al., 2012

Table 4 Log-Linear Model Fit Statistics

Model	df	χ^2	p-value	BIC	Δ
<i>Men (N = 56,200)</i>					
<i>Baseline Models</i>					
0a: Origin + Destination + Period	27,972	244,675	<.01	-61,245	69.0
1a: 0a + Origin \times Period + Destination \times Period	27,380	239,997	<.01	-59,449	68.5
2a: 1a + Parameter-ABCD	27,292	27,431	.28	-271,052	18.2
<i>Period-Varying Topological Mobility Models</i>					
3a: 2a + Period \times Parameter-A	26,992	24,874	>.99	-270,328	14.2
4a: 2a + Period \times Parameter-AB	26,956	24,723	>.99	-270,086	14.1
5a: 2a + Period \times Parameter-ABC	26,944	24,661	>.99	-270,017	14.1
6a: 2a + Period \times Parameter-ABCD	26,940	24,648	>.99	-269,986	14.1
7a: 2a + Period \times Parameter-A'B'C'D'	27,276	25,659	>.99	-272,650	16.0
<i>Women (N = 47,180)</i>					
<i>Baseline Models</i>					
0b: Origin + Destination + Period	27,972	174,593	<.01	-126,434	66.6
1b: 0b + Origin \times Period + Destination \times Period	27,380	166,191	<.01	-128,465	65.7
2b: 1b + Parameter-ABCD	27,292	19,635	>.99	-274,074	15.2
<i>Period-Varying Topological Mobility Models</i>					
3b: 2b + Period \times Parameter-A	26,992	18,287	>.99	-272,193	12.3
4b: 2b + Period \times Parameter-AB	26,956	18,154	>.99	-271,939	12.2
5b: 2b + Period \times Parameter-ABC	26,944	18,102	>.99	-271,862	12.1
6b: 2b + Period \times Parameter-ABCD	26,940	18,077	>.99	-271,844	12.1
7b: 2b + Period \times Parameter-A'B'C'D'	27,276	18,879	>.99	-274,657	13.9

B. F. Jarvis and X. Song, 2017

Point-to-Point Transition

Fails to Capture Long-term Impact



Impacts of Individual Factors

Lack of Social Factor Analysis



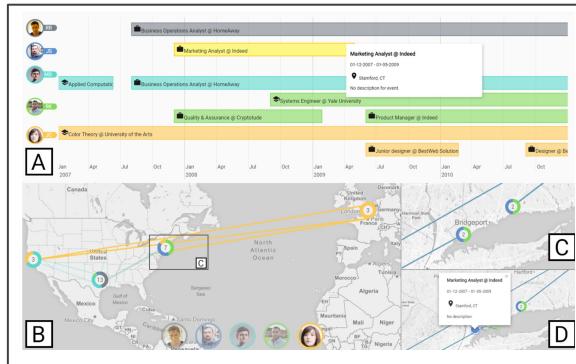
Statistical and Hypothesis-Driven

Require Computationally Efficient Ways

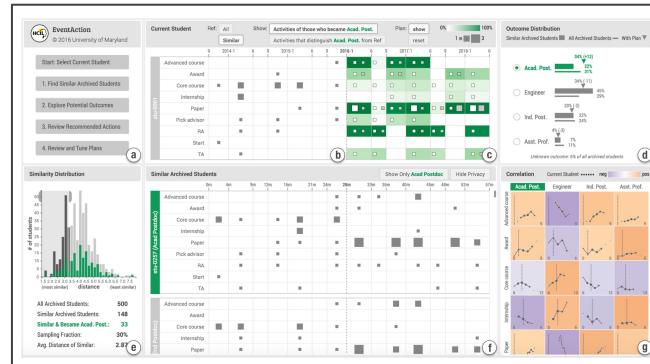


Analyzing Academic Careers

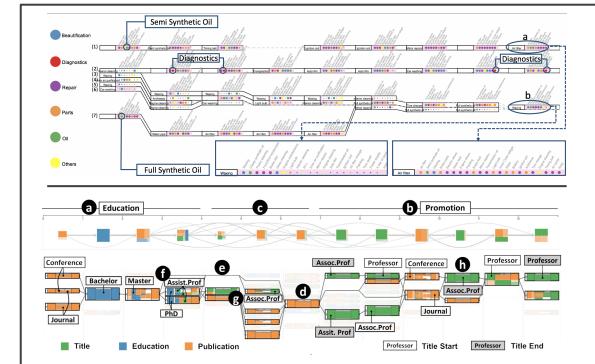
New analytical demands in today's boundaryless career world



V.Filipov et al., 2019



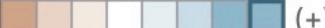
F.Du et al., 2016



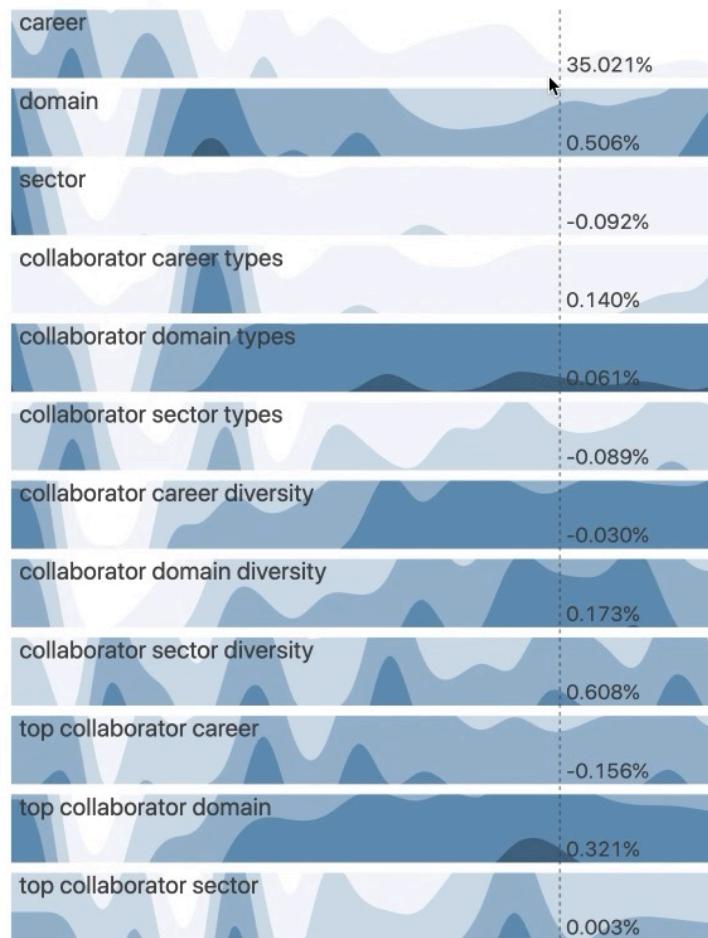
S. Guo et al., 2017; S. Guo et al., 2018

Only Focus on Sequential Properties
Lack of Individual and Social Factor Analysis

Factor View

p-value (0)  (1)coef (-)  (+)

Window

Factors Independent Scale

Cluster View

Person View

Career

Senior	Mediate	Junior		
Q4	Q3	Q2	Q1	Q0

Domain Vis Others Sec

Challenges

Factor distillation and dynamic
Impacts analysis

Intuitive visual representations
of the effects of factors

Effective exploration of
academic profiles and the multi-
factor effects

Challenges

Factor distillation and dynamic Impacts analysis

Intuitive visual representations
of the effects of factors

Effective exploration of
academic profiles and multi-
factor effects



Operationalizing **social factors** (e.g., collaborations) aggravates
the complexity of the problem given their network nature

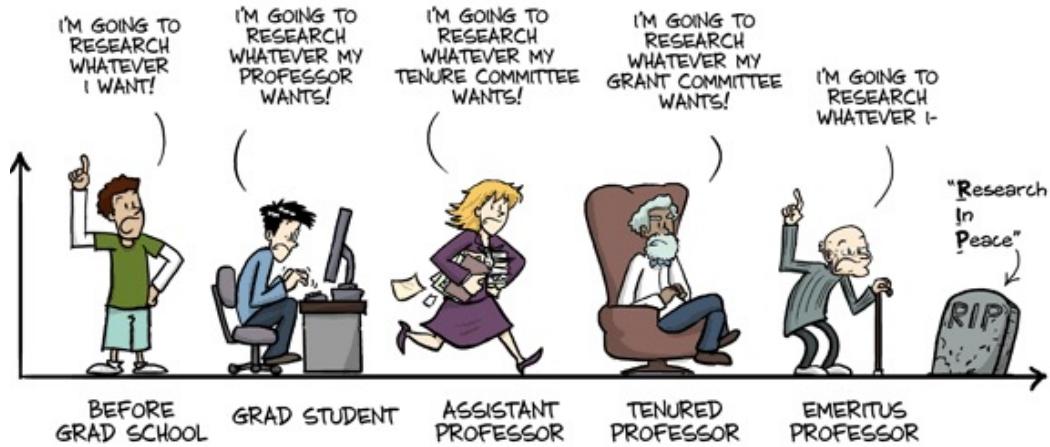
Challenges

Factor distillation and dynamic Impacts analysis

Intuitive visual representations of the effects of factors

Effective exploration of academic profiles and multi-factor effects

THE EVOLUTION OF INTELLECTUAL FREEDOM



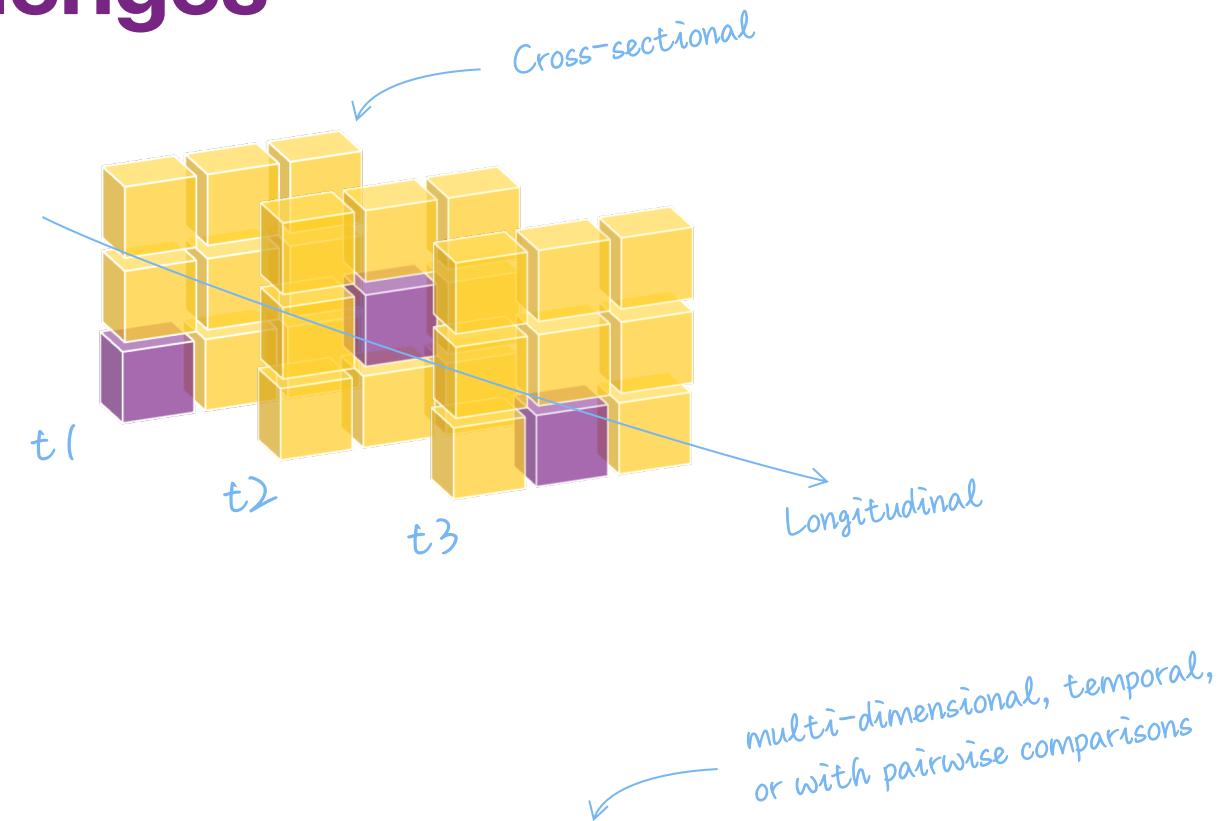
Capturing the effects of factors requires **a dynamic multivariate analytical framework** given the long period of research fields

Challenges

Factor distillation and dynamic
Impacts analysis

Intuitive visual representations of the effects of factors

Effective exploration of academic
profiles and multi-factor effects



Visual representations for complex data structures require supporting
both **cross-sectional** and **longitudinal** studies of factors

Challenges

Factor distillation and dynamic
Impacts analysis

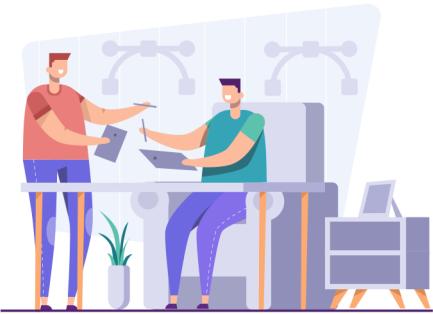
Intuitive visual representations
of the effects of factors

**Effective exploration of academic
profiles and multi-factor effects**

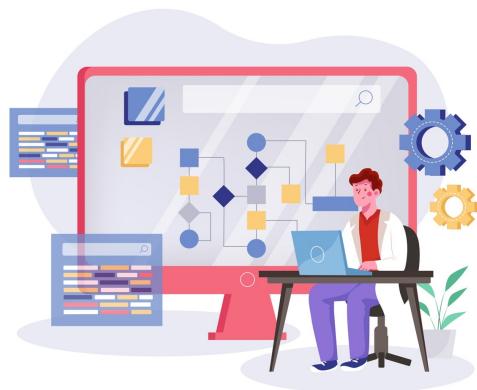


Combining **academic profiles** with the **impacts and patterns** to
comprehensively understand the results is necessary

Contributions



**Problem Domain
Characterization**



**Multi-factor Effects
Analytical Framework**



**ACSeeker Implementation;
Real World Data Demonstration**

Background and Concepts

- **Career Success**
 - Or called **Career Performance**, is the outcomes (e.g., title ranks and incomes) of one's working experiences. In academic careers, **citations** of a scholar's research outputs are commonly used to measure career success



Background and Concepts

- **Career Factors**

- Potential factors (i.e., variates) that can affect one's career success, includes both individual and social factors.
- In academic careers,
 - **Individual Factors:** job title ranks (e.g., junior and senior), sectors (e.g., academia, industry, and government agencies), and research domains
 - **Social Factors:** collaborators' individual factors as social factors

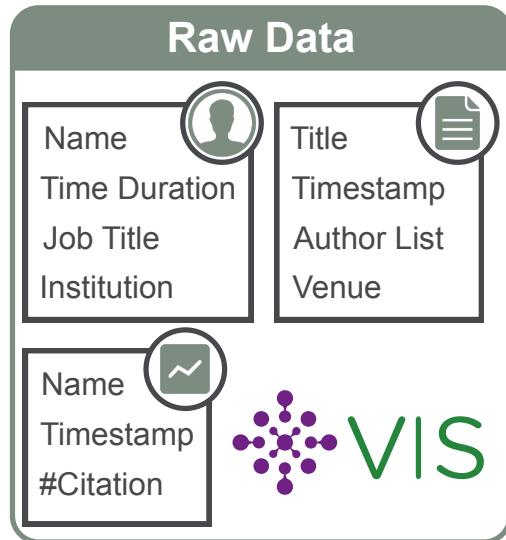


Background and Concepts

- **Factor Categories**
 - Different groups that are identified based on the values within a factor according to user-specified definitions



Data Sources



1100+ VIS Researchers

- **Career Data**

- Researchers' websites, their institutional webpages, and LinkedIn

- **Bibliographic Data**

- Aminer [1]

- **Citation Data**

- Google Scholar

Job title, year, organization, ...

Paper title, venue, year, coauthors, ...

Citations by year



[1] J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su. ArnetMiner: Extraction and mining of academic social networks. In Proceedings of ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 990–998, 2008.

Task Analysis



Inter-factor-level

- **T1** How does a **factor** influence career success over time?
- **T2** How do multiple factors differ in their impacts on career success?



Intra-factor-level

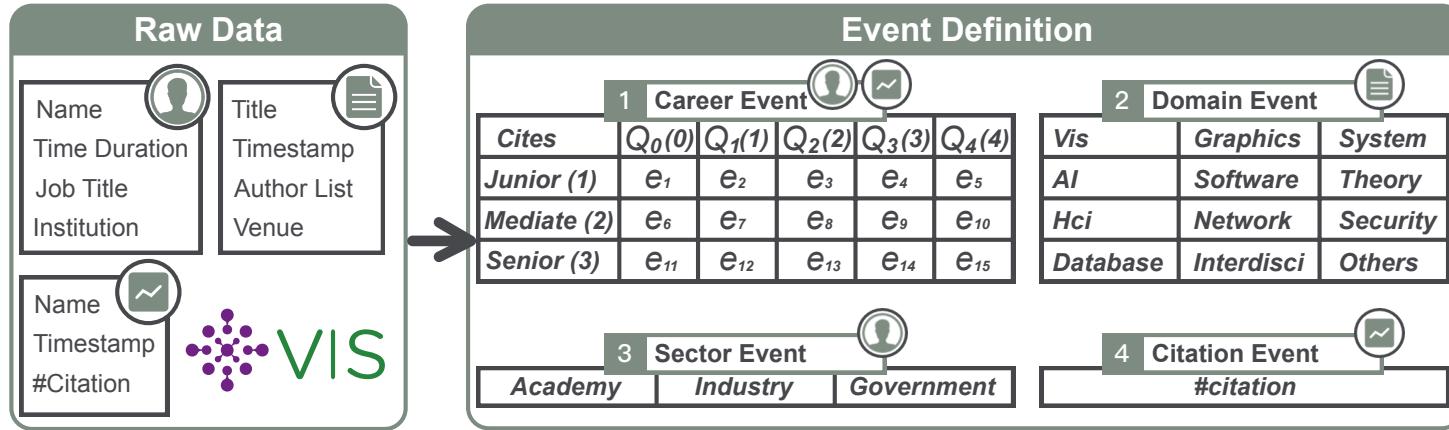
- **T3** How does a **category** within a factor change over time to affect career success?
- **T4** How do the categories within a factor differ from affecting career success?



Individual-level

- **T5** How does a researcher's career path change over time?
- **T6** How do the different factors of a researcher change over time?

Data Preprocessing

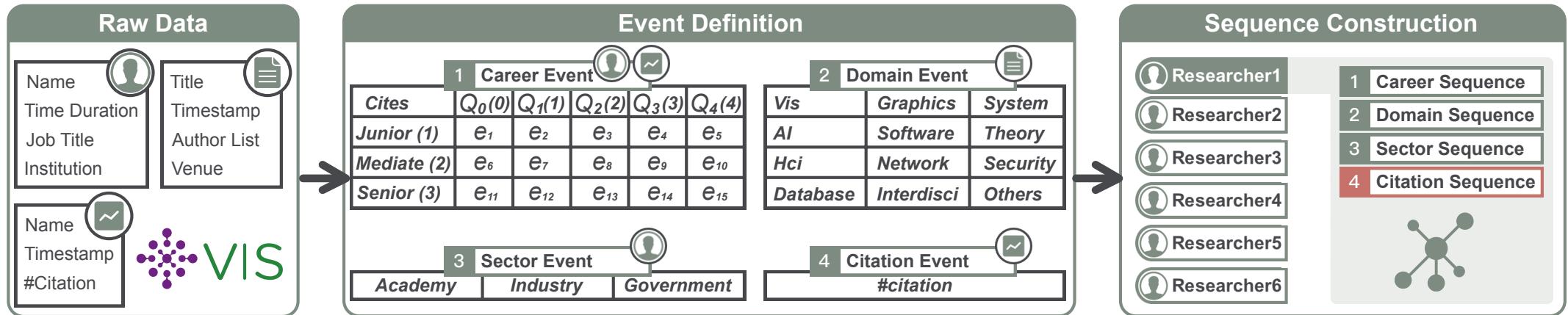


- **Event Definition:** based on career lengths/titles, citations, paper venues, and sectors [1]

Academia, Industry, or Government

[1] For more details, please read our supplementary materials on wangyifang.top.

Data Preprocessing



- **Sequence Construction**: based on four event types

- Career Sequence
- Domain Sequence
- Sector Sequence
- Citation Sequence

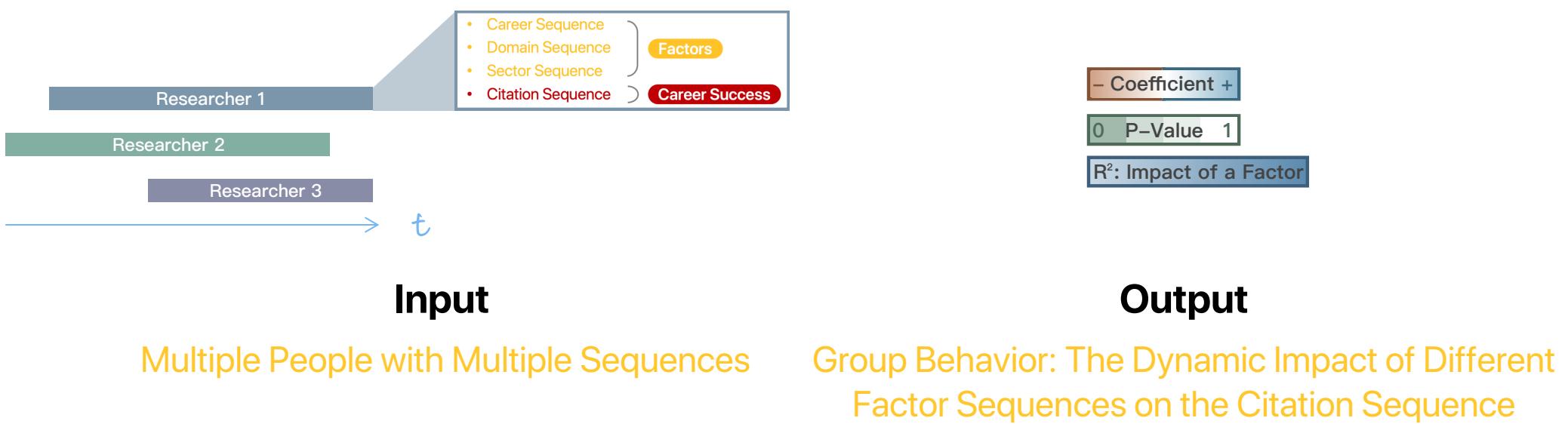
} Factors
Career Success

- **Dynamic Network Construction**: record all **collaborators** of each researcher by year. We use their individual factor sequences as **social factors**



Data Model

Sequence History Analysis (SHA) → Multi-factor Impact Analysis (MIA)



Data Model

Sequence History Analysis (SHA)

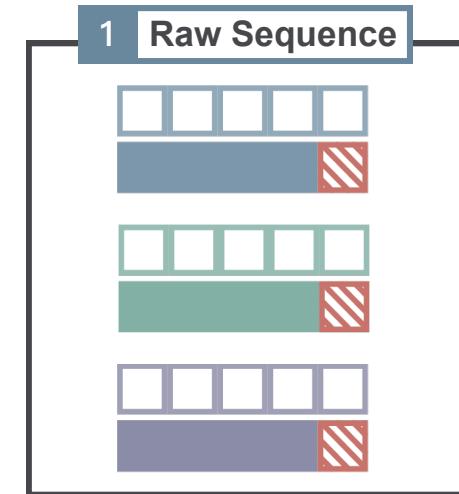


E.g., both parents (without siblings),
one parent and sibling(s), ...

- **Example:** how the **co-residence trajectories** affect teenagers' **timing of the departure** from the parental home [1]

E.g., departure at age 14

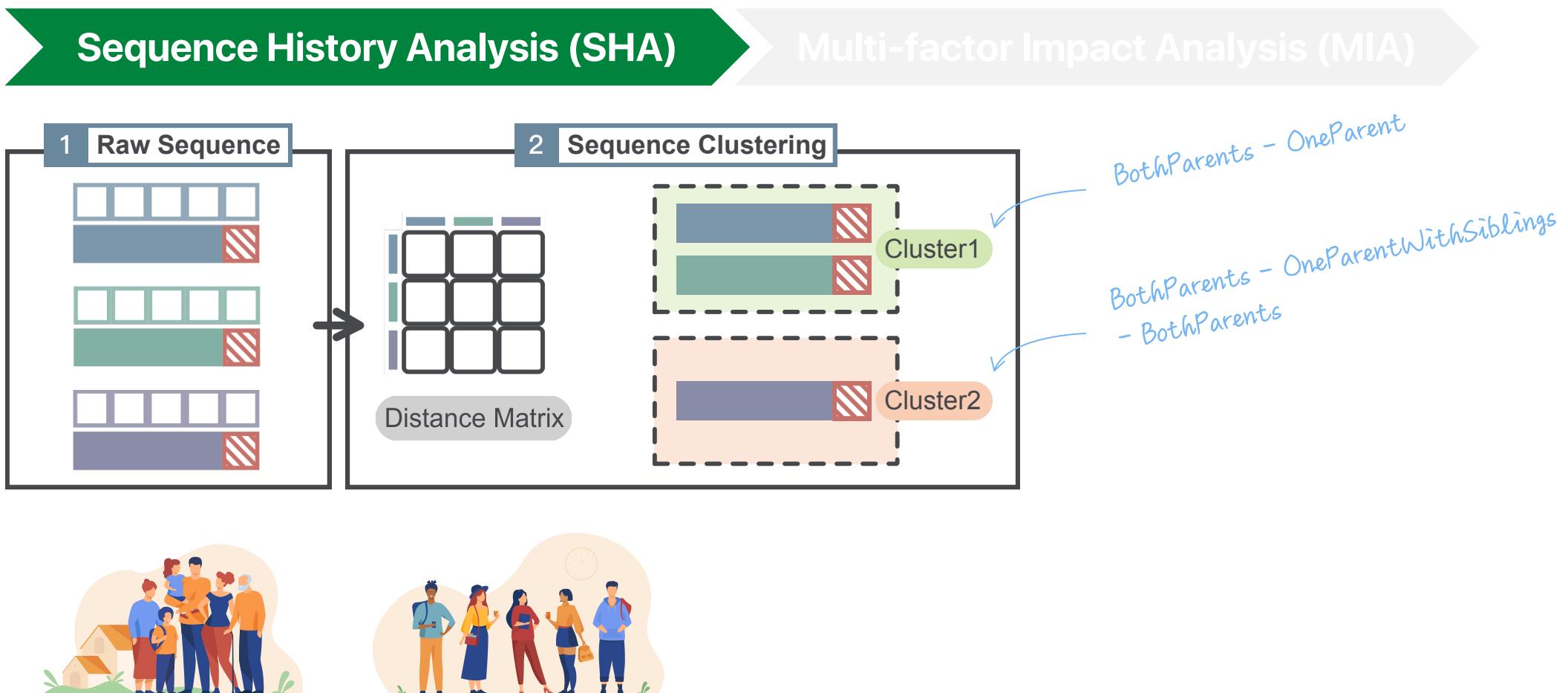
Multi-factor Impact Analysis (MIA)



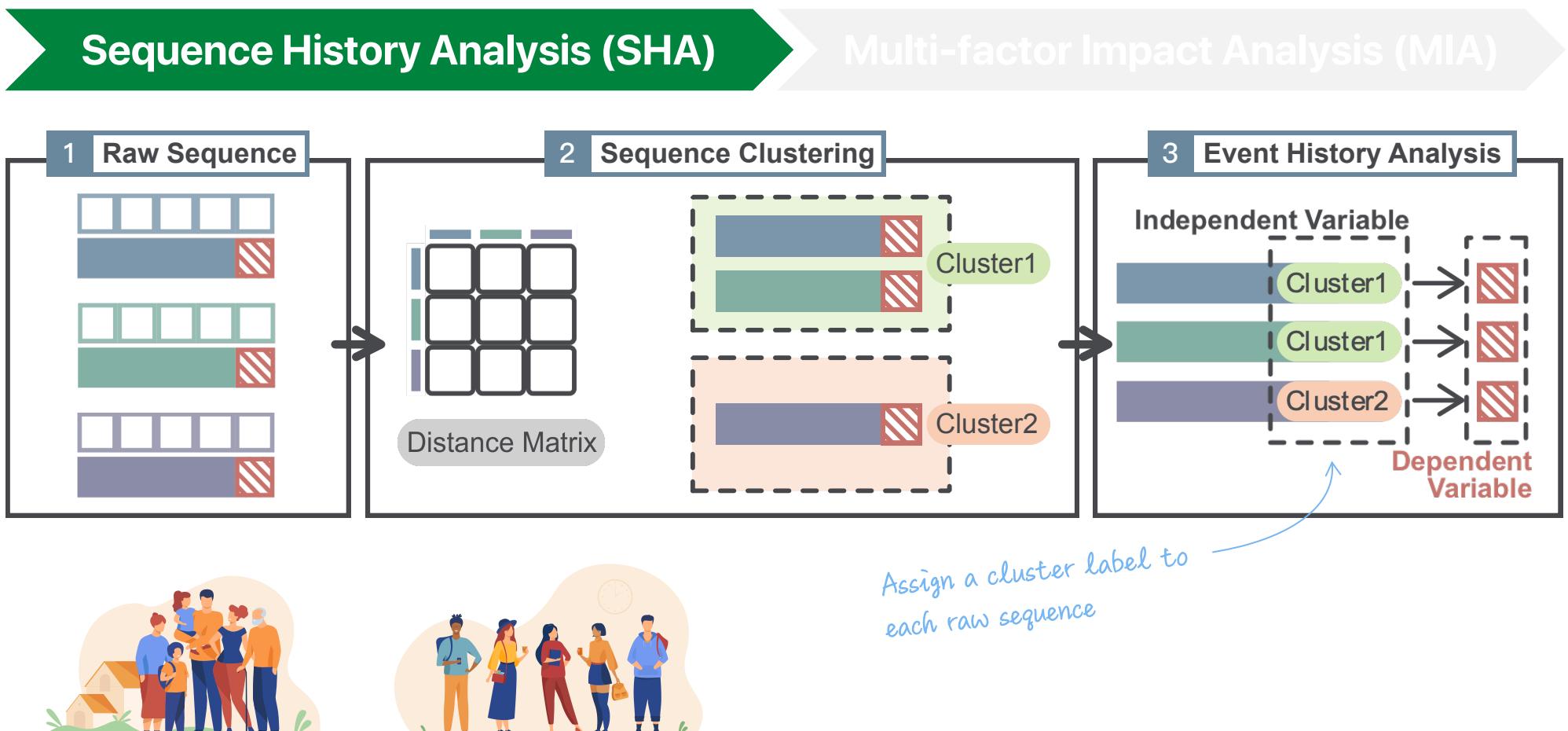
BothParents - OneParent	Departure at 15
BP - OP	Departure at 17
BP - OPS - BP	Departure at 14

[1] F. Rossignon, M. Studer, J.-A. Gauthier, and J.-M. Le Goff. Sequence history analysis (SHA): Estimating the effect of past trajectories on an upcoming event. In Sequence Analysis and Related Approaches, pp. 83–100. 2018.

Data Model

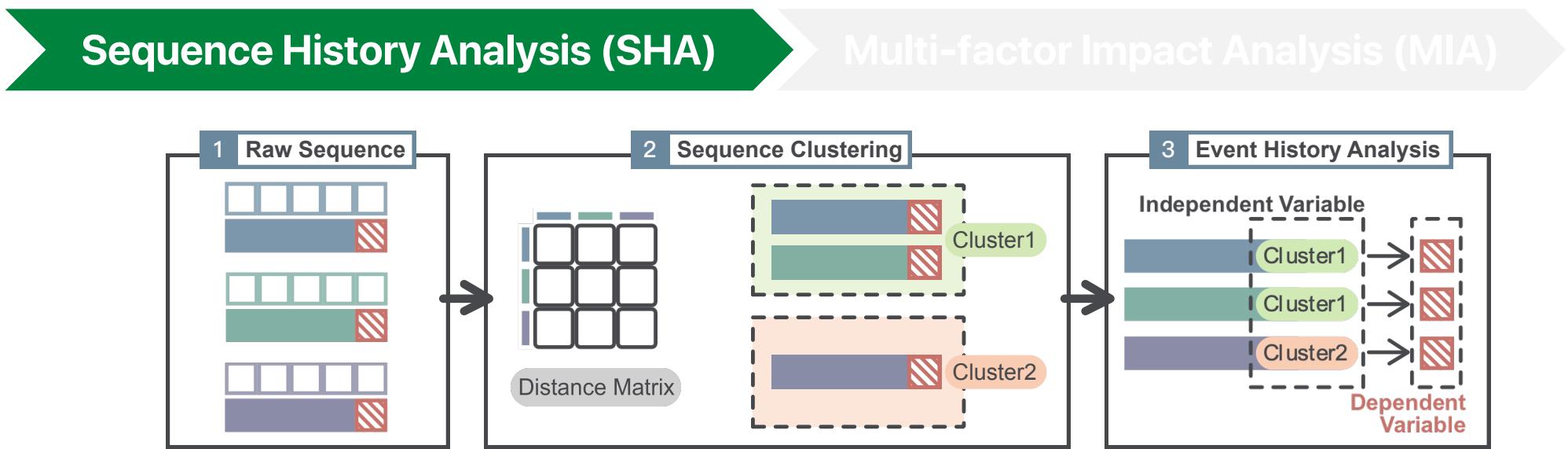


Data Model



[1] F. Rossignon, M. Studer, J.-A. Gauthier, and J.-M. Le Goff. Sequence history analysis (SHA): Estimating the effect of past trajectories on an upcoming event. In Sequence Analysis and Related Approaches, pp. 83–100. 2018.

Data Model

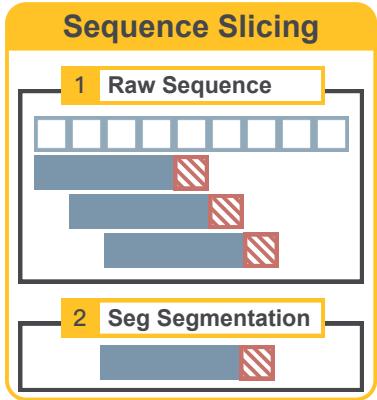


- **Limitations**
 - Can not analyzes the impacts in a **dynamic** manner
 - Can not support **multi-variate** analysis

Data Model

Sequence History Analysis (SHA)

Multi-factor Impact Analysis (MIA)



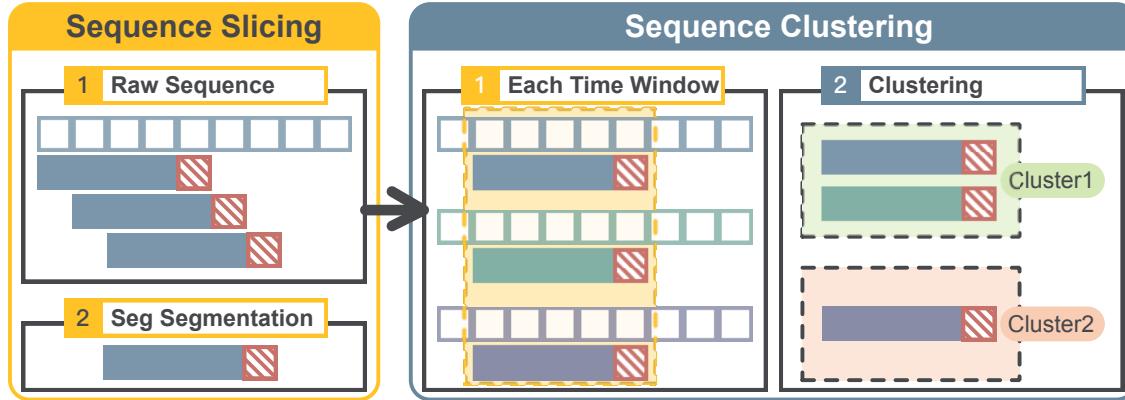
- **Sequence Slicing**

- The farther the historical event, the less relevant it is to the upcoming career performance
- To inspect **the time-varying impacts** of different factors: slice sequences into different time windows

Data Model

Sequence History Analysis (SHA)

Multi-factor Impact Analysis (MIA)

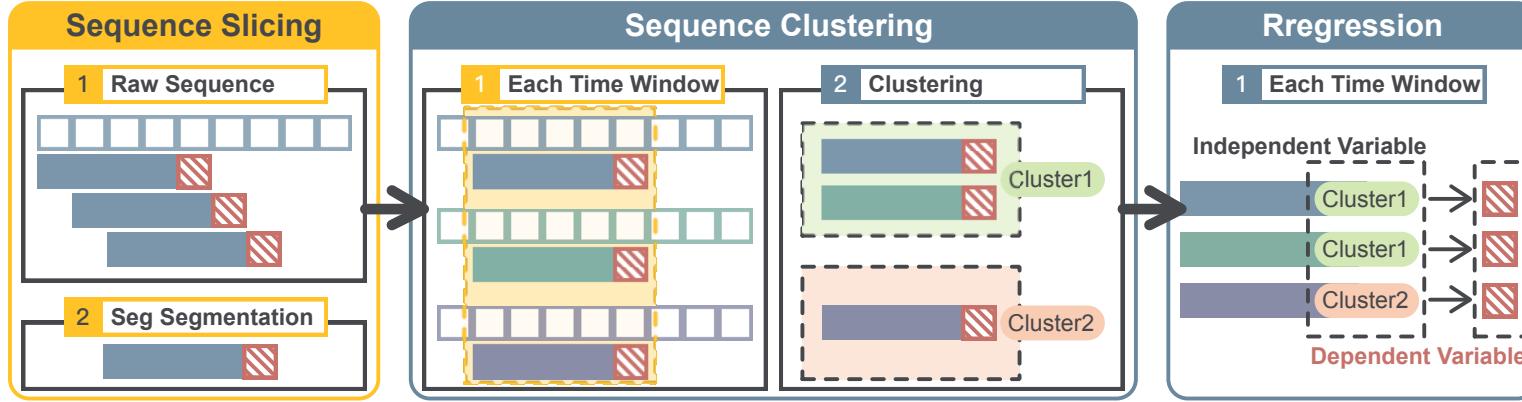


- **Sequence Clustering**

- For each sequence type (career, domain, and sector), for each time window: conduct sequence clustering

Data Model

Sequence History Analysis (SHA) → Multi-factor Impact Analysis (MIA)



- **Regression**

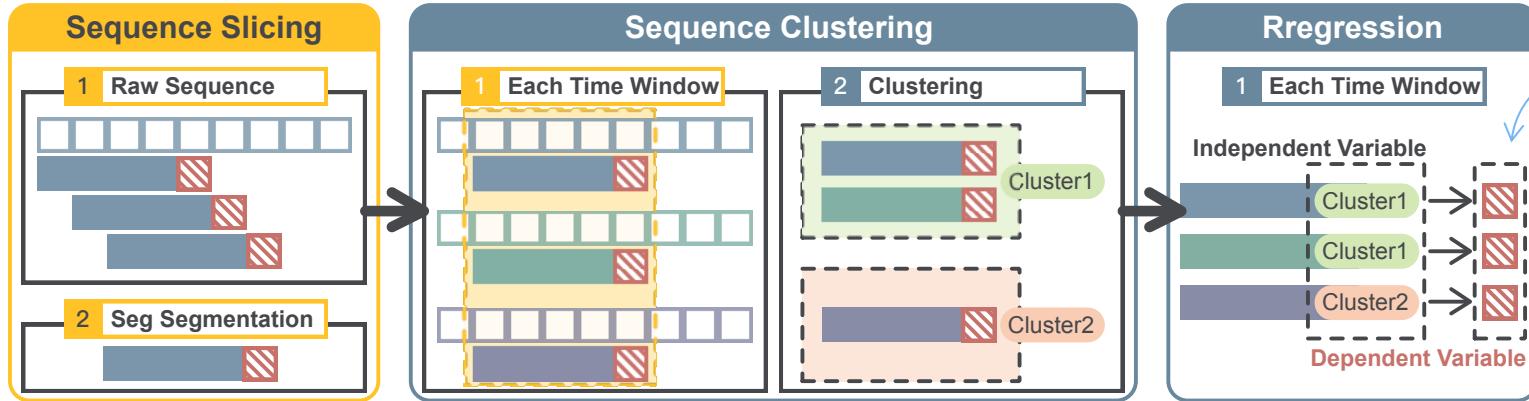
- For each sequence type (career, domain, and sector), for each time window:
conduct **Multivariate Linear Regression**:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_n x_{in} + \varepsilon_i, (i = 1, \dots, n) \quad (1)$$

Data Model

Sequence History Analysis (SHA)

Multi-factor Impact Analysis (MIA)



Outputs

- Coefficient +

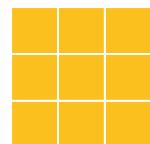
P-Value

R²: Impact of a Factor

Set each cluster as the reference category in turn
to enables pairwise comparisons among all
categories of a categorical IV as a panoramic view

• Regression

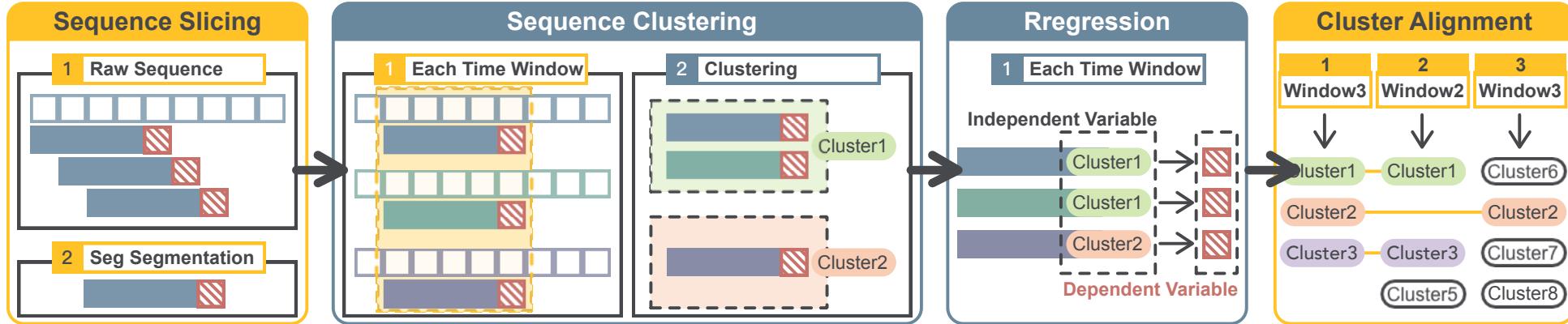
- For each sequence type (career, domain, and sector), for each time window:
conduct **Multivariate Linear Regression**
- Post-hoc analysis:** for categorical independent variables (i.e., sequence clusters)



Data Model

Sequence History Analysis (SHA)

Multi-factor Impact Analysis (MIA)



- **Cluster Alignment**
 - Align the same cluster in each sequence type **across time windows** to inspect its effect on career success **temporally**

ACSeeker System



ACSeeker System

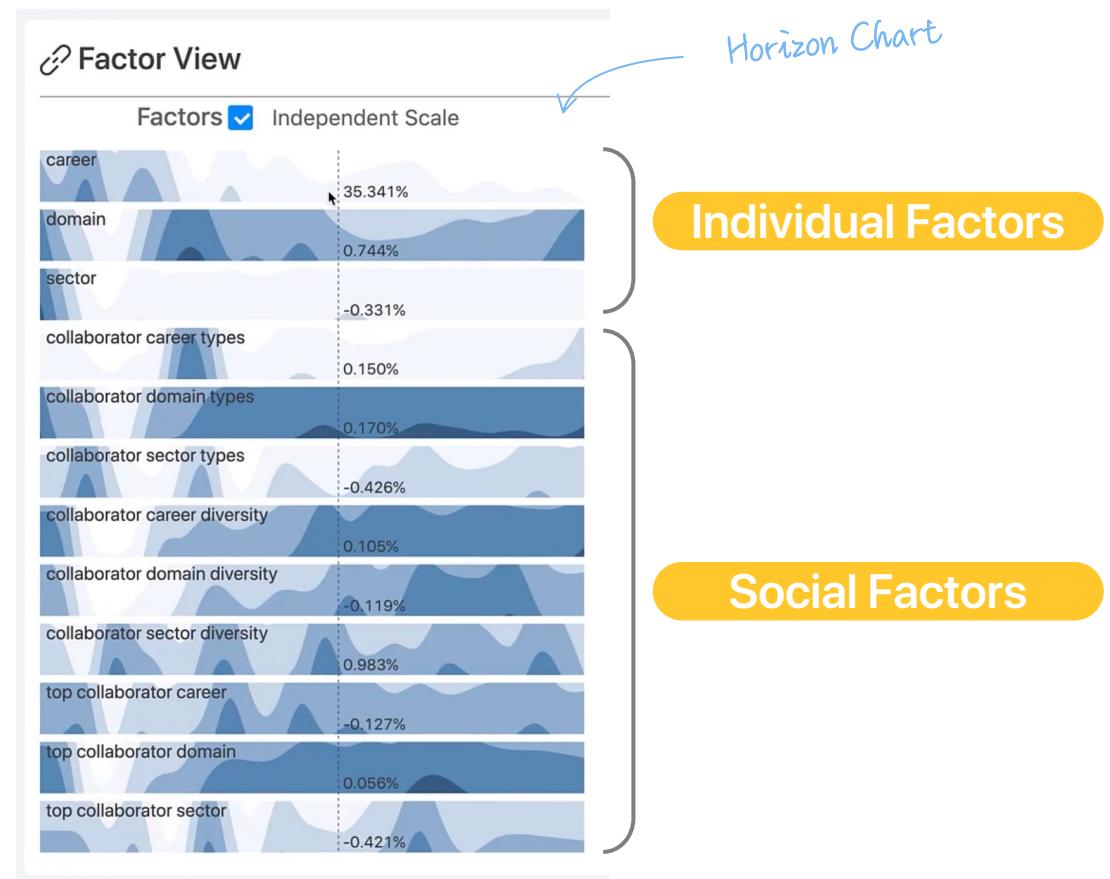


- **T1** How does a **factor** influence career success over time?
- **T2** How do multiple factors differ in their impacts on career success?

– Coefficient +

0 P-Value 1

R²: Impact of a Factor



ACSeeker System



- **T3** How does a **category** within a factor change over time to affect career success?
- **T4** How do the categories within a factor differ from affecting career success?

- Coefficient +

0 P-Value 1

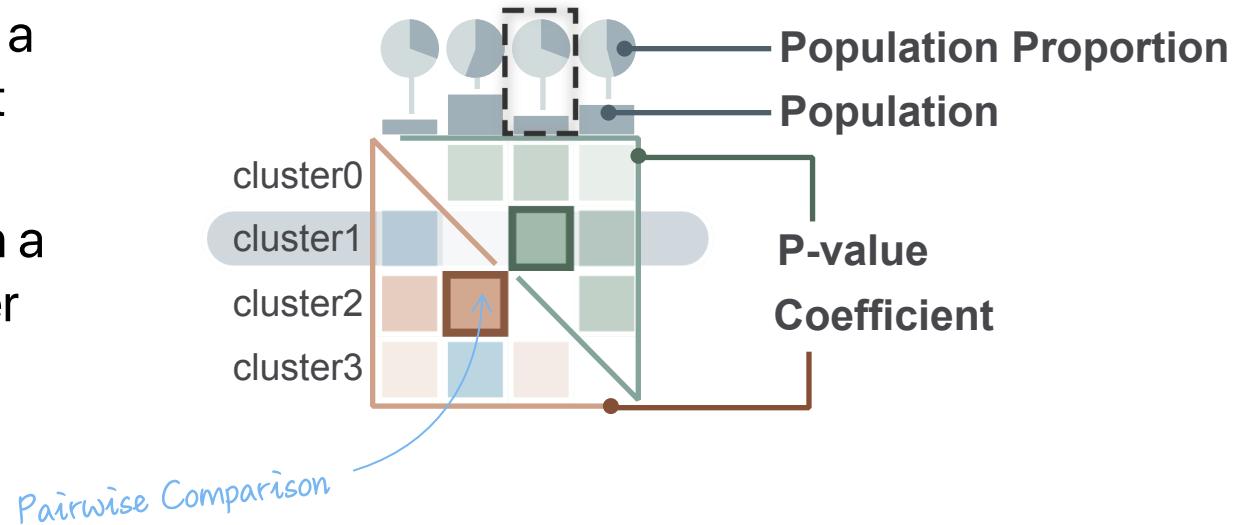
R²: Impact of a Factor



ACSeeker System



- **T3** How does a **category** within a factor change over time to affect career success?
- **T4** How do the categories within a factor differ from affecting career success?



- Coefficient +

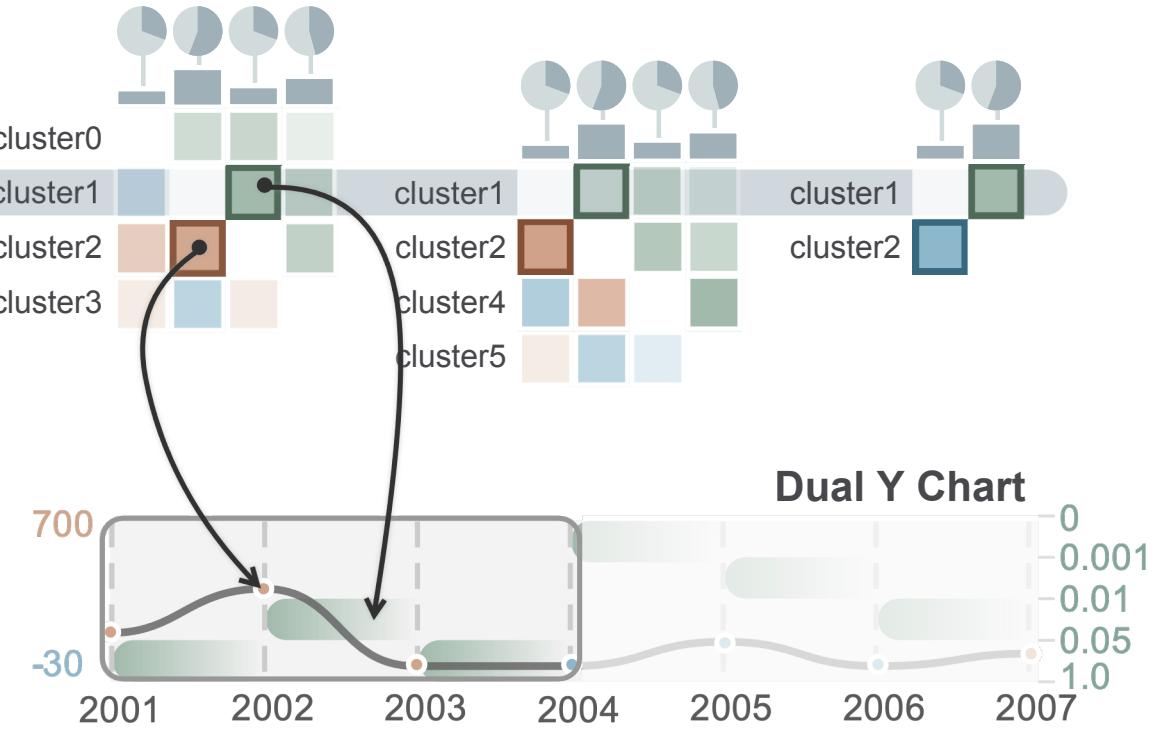
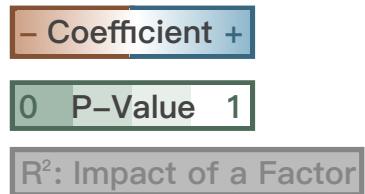
0 P-Value 1

R²: Impact of a Factor

ACSeeker System



- **T3** How does a **category** within a factor change over time to affect career success?
- **T4** How do the categories within a factor differ from affecting career success?



ACSeeker System

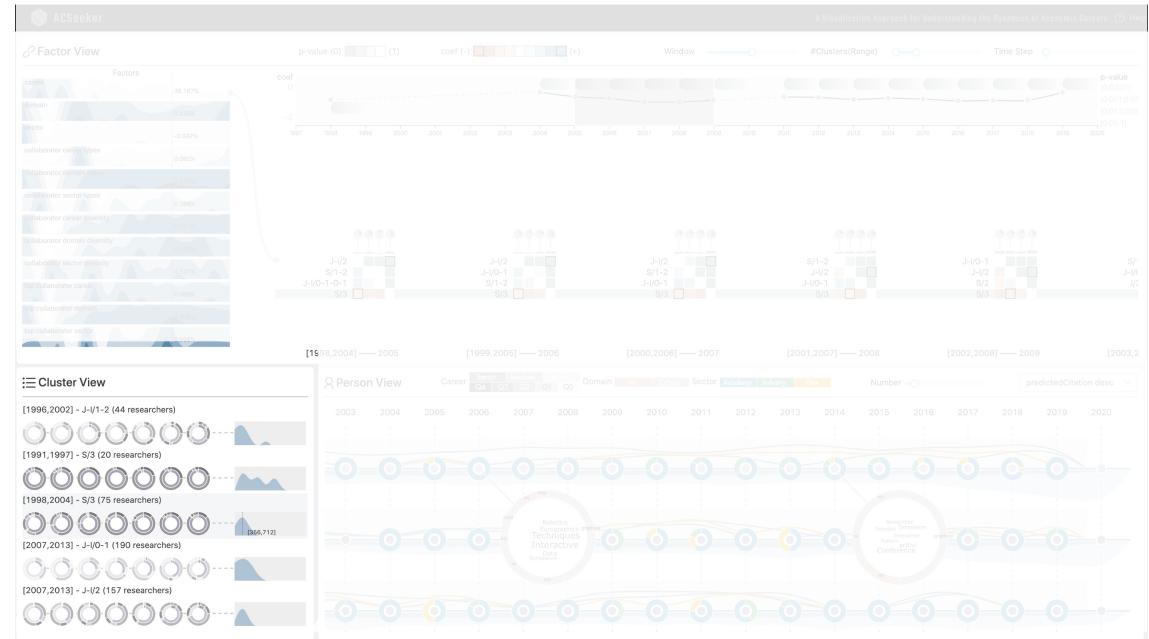


- **T3** How does a **category** within a factor change over time to affect career success?
- **T4** How do the categories within a factor differ from affecting career success?

- Coefficient +

0 P-Value 1

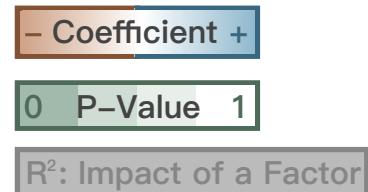
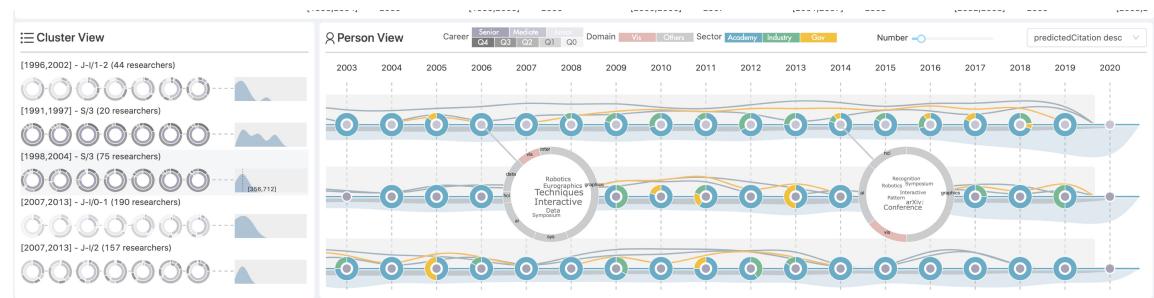
R²: Impact of a Factor



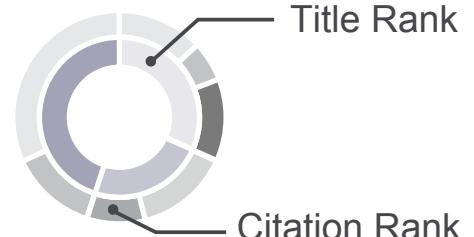
ACSeeker System



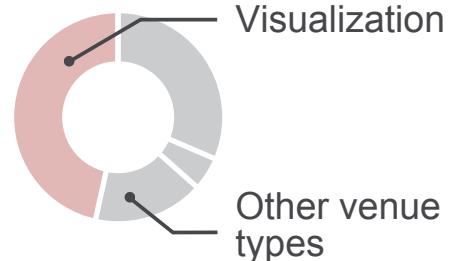
- **T3** How does a **category** within a factor change over time to affect career success?
- **T4** How do the categories within a factor differ from affecting career success?



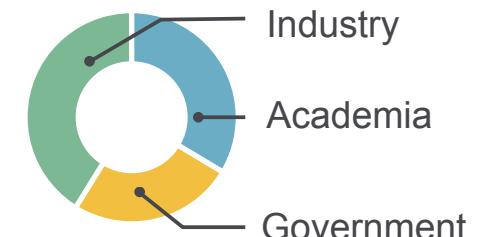
Career Glyph



Domain Glyph



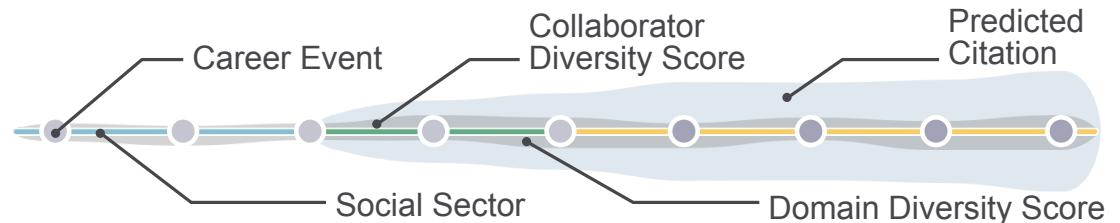
Sector Glyph



ACSeeker System



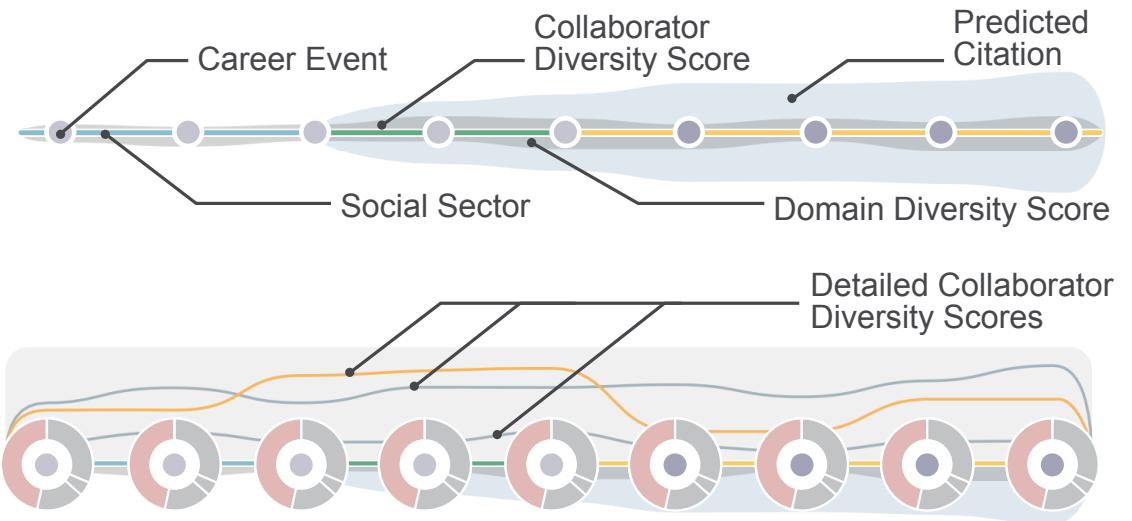
- **T5** How does a researcher's career path change over time?
- **T6** How do the different factors of a researcher change over time?



ACSeeker System

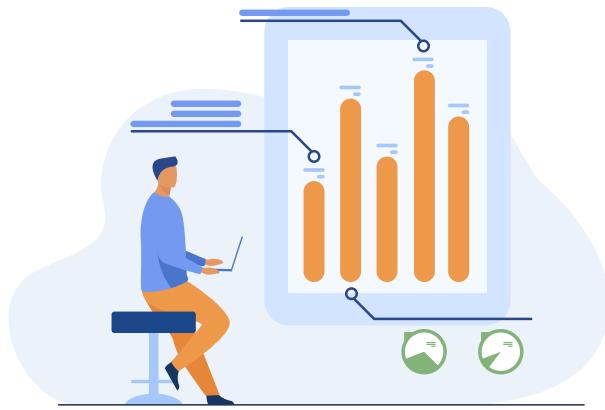


- **T5** How does a researcher's career path change over time?
- **T6** How do the different factors of a researcher change over time?



$$Diversity(k,t) = - \sum_{i=1}^n P(x_i,t) \log P(x_i,t), k \in \{career, sector, domain\}$$

Evaluation



Social Scientist

Case Study & Expert Interview



VIS Researchers

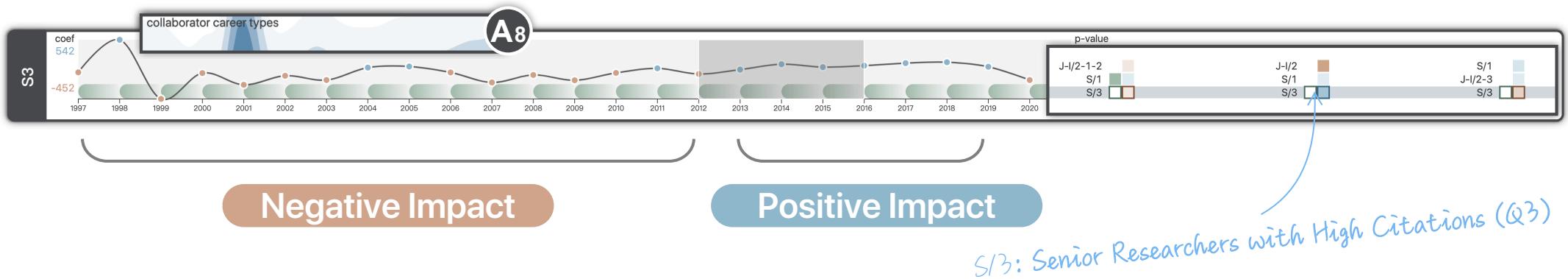
Free Exploration and Interview

Evaluation

Social Factor

Collaborators' Career Types

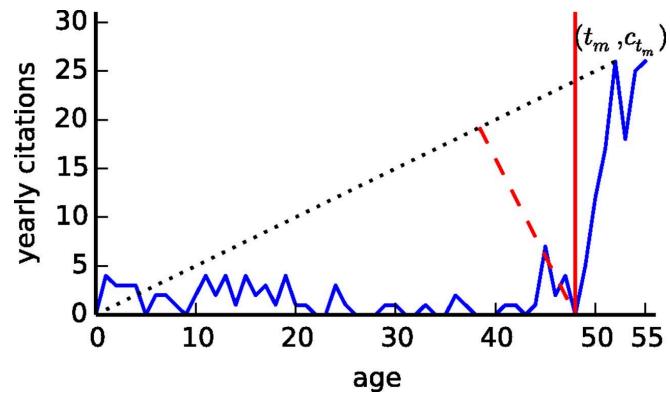
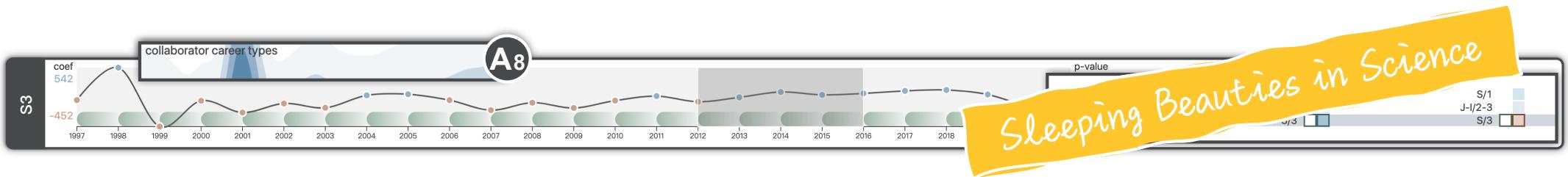
Working with who? Senior or Junior?



Evaluation

Social Factor

Collaborators' Career Types



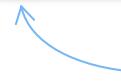
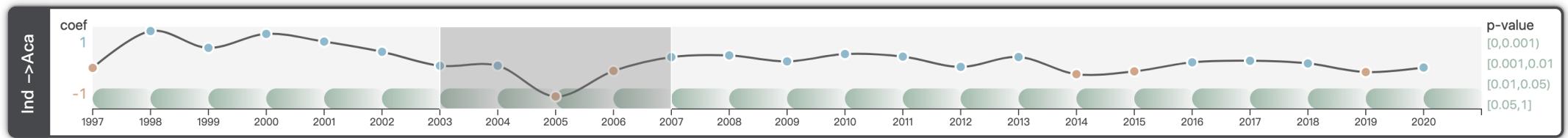
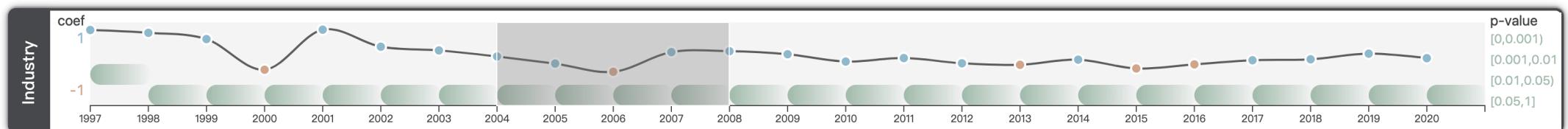
- [1] Ke, Qing, et al. "Defining and identifying sleeping beauties in science." Proceedings of the National Academy of Sciences 112.24 (2015): 7426-7431.
[2] Paper with delayed recognition (or a "sleeping beauty") : Wikipedia. https://en.wikipedia.org/wiki/Paper_with_delayed_recognition.

Evaluation

Social Factor

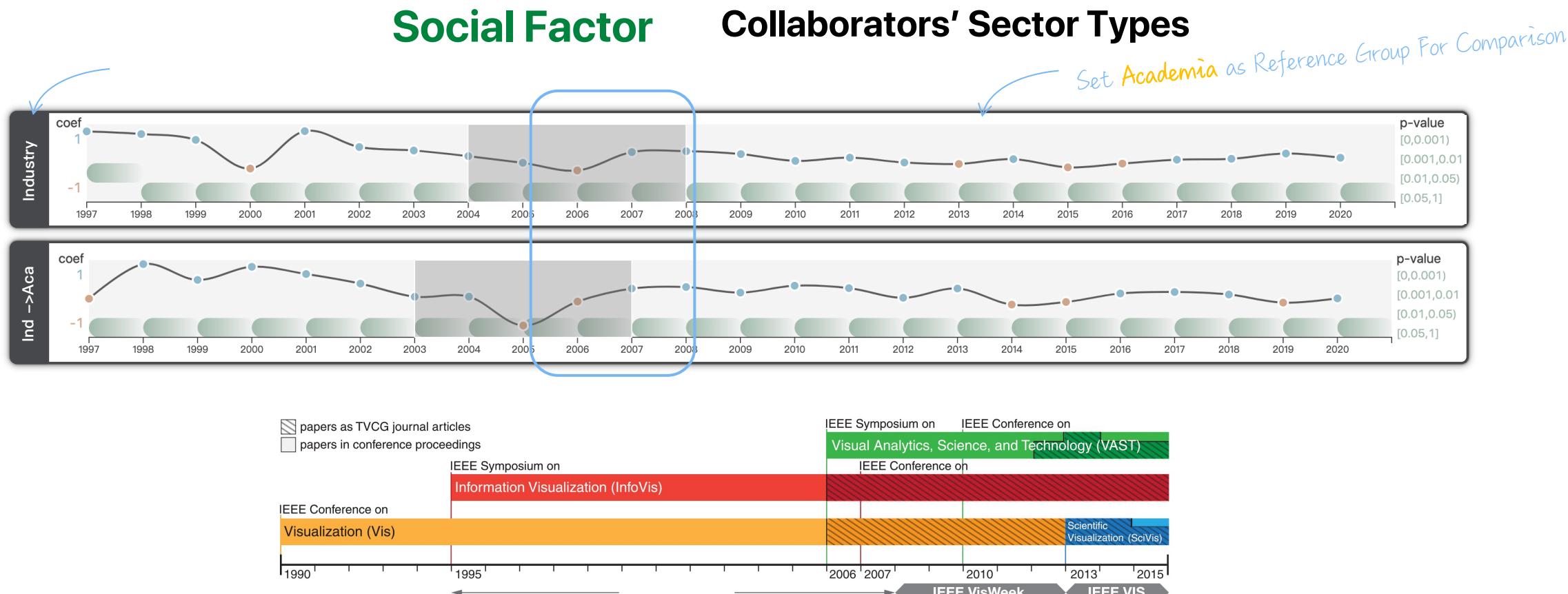
Collaborators' Sector Types

Working with who? Academia, Industry, or Gov Institutions?



Set Academia as Reference Group For Comparison

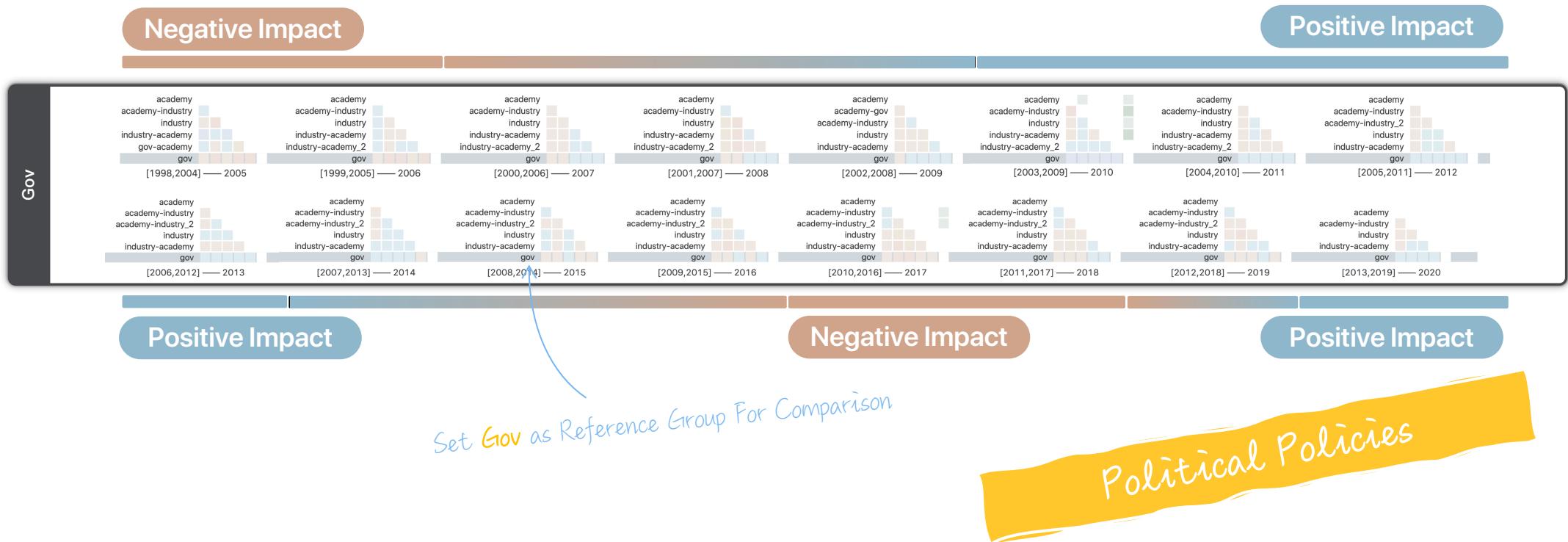
Evaluation



Evaluation

Social Factor

Collaborators' Sector Types



Conclusion

Contributions

- Characterize the problem domain of visual analytics of time-varying effects of multiple factors on academic career success
- Propose a novel framework that analyzes the effects of multiple factors on researchers' career success longitudinally
- Develop an interactive visualization system *ACSeeker*, for social scientists and other scholars to explore academic careers with multiple factors through novel designs
- Demonstrate the effectiveness and usability of *ACSeeker* with a dataset that involves more than 1,100 visualization researchers

Future Work

- Enhance the current MIA framework by setting dynamic lengths of sliding windows based on sequential trends
- Collect data from a border range (e.g., computer science field) to find more stories
- Gather requirements from general researchers to adapt *ACSeeker* for wider use

Seek for Success: A Visualization Approach for Understanding the Dynamics of Academic Careers



Yifang Wang^{1,3} Winson Peng²



Huihua Lu³



Haoren Wang³



Xiao Xie³



MICHIGAN STATE UNIVERSITY



浙江大學
THE ZJU UNIVERSITY

