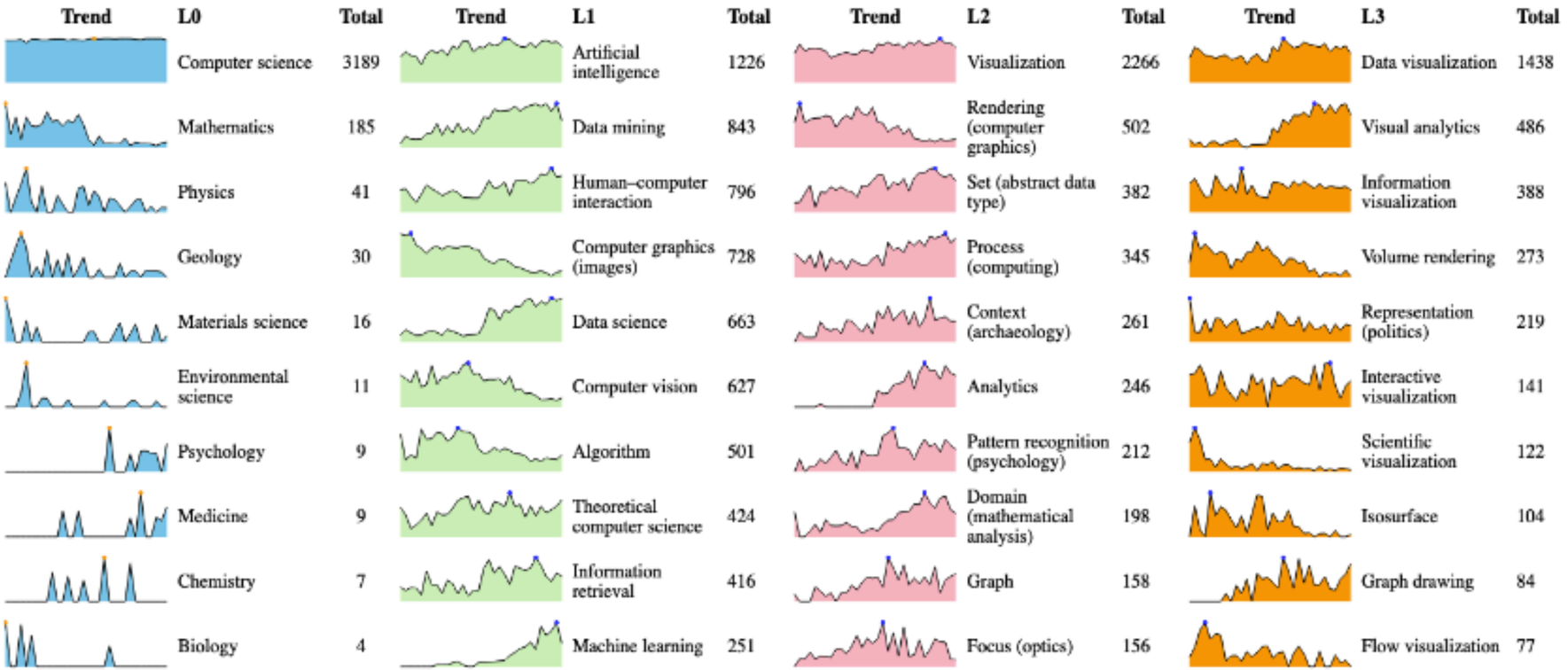
Thirty-Two Years of IEEE VIS: Authors, Fields of Study and Citations

Hongtao Hao, Yumian Cui, Zhengxiang Wang, and Yea-Seul Kim



Trend	L0	Total	Trend	L1	Total	Trend	L2	Total	Trend	L3	Total
	Computer science	3189		Artificial intelligence	1226		Visualization	2266		Data visualization	1438
	Mathematics	185		Data mining	843		Rendering (computer graphics)	502		Visual analytics	486
	Physics	41		Human–computer interaction	796		Set (abstract data type)	382		Information visualization	388
	Geology	30		Computer graphics (images)	728		Process (computing)	345		Volume rendering	273
	Materials science	16		Data science	663		Context (archaeology)	261		Representation (politics)	219
	Environmental science	11		Computer vision	627		Analytics	246		Interactive visualization	141
	Psychology	9		Algorithm	501		Pattern recognition (psychology)	212		Scientific visualization	122
	Medicine	9		Theoretical computer science	424		Domain (mathematical analysis)	198		Isosurface	104
	Chemistry	7		Information retrieval	416		Graph	158		Graph drawing	84
	Biology	4		Machine learning	251		Focus (optics)	156		Flow visualization	77

Fig. 1: Fields of study at different levels in 3,240 VIS papers across the past 32 years. From L0 to L3, granularity increases. "Trend" indicates the proportion of papers falling into a field of study against the total number of papers published in that year. The highest proportion for each field of study is highlighted. "Total" indicates the total number of VIS publications falling into a field of study. One paper may contain more than one field of study at the same level, and one field of study may have multiple parent fields. For example, Pattern Recognition belongs to both Computer Science and Psychology.

Where is VIS in science?

Introduction

For IEEE VIS 2021 and beyond, the conference changes how it groups shared research interests as part of the reviewing process. Previously, the conference consisted of three sub-conferences, each with similar but separate reviewing processes. Instead, the new area model groups different research topics in Visualization and Visual Analytics into six areas. This allows research papers on closely-related topics to be reviewed in a coherent manner. To ensure a high quality review process, two area paper chairs oversee the reviewing process for each area and draw program committee members from a large joint program committee (PC). The mechanism of a unified PC across all areas allows area co-chairs to have access to wider expertise than was previously the case.

This page provides guidance on how the area model affects authors, reviewers, paper chairs, and the reviewing process more generally. It also gives guidance on how to make an appropriate choice during paper submission.

https://ieeevis.org/year/2022/info/call-participation/area-model

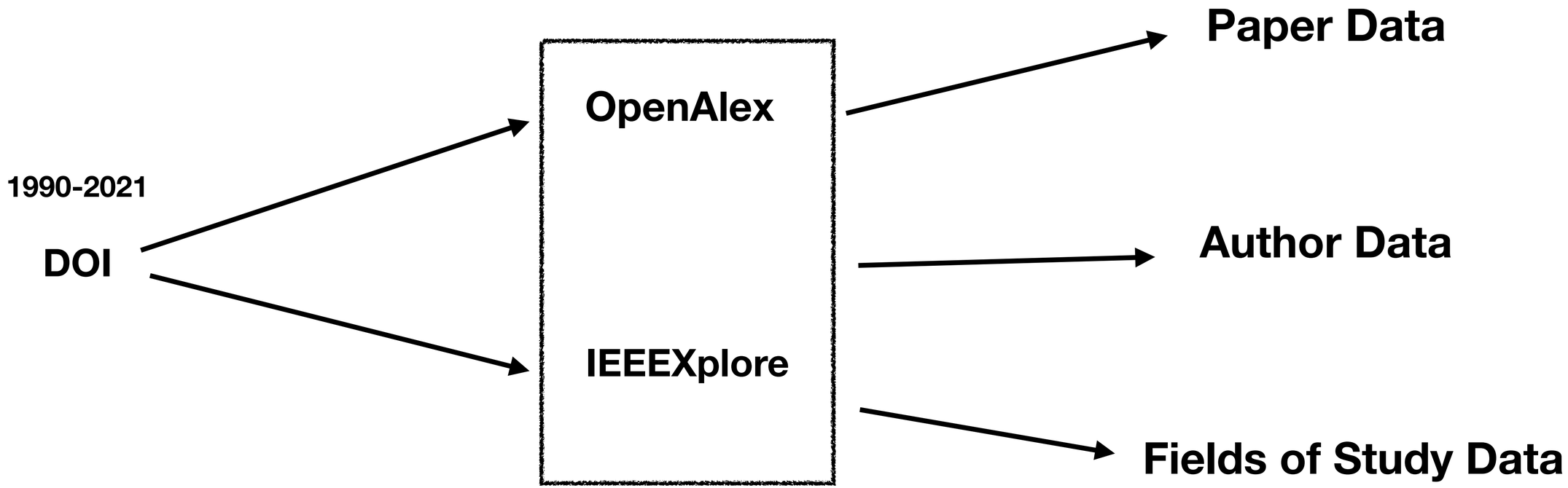
Affiliations? Countries?

“The conference will convene an international community of researchers and practitioners from universities, government, and industry to exchange recent findings ...”

Research questions

- Where does VIS stand in science? What are VIS topics, whom VIS cited and Who are citing VIS?
- Where are VIS authors from, in terms of their affiliation types and affiliation country origins? What have been their collaboration patterns?

Data


1990-2021
DOI
OpenAlex
IEEEXplore
Paper Data
Author Data
Fields of Study Data

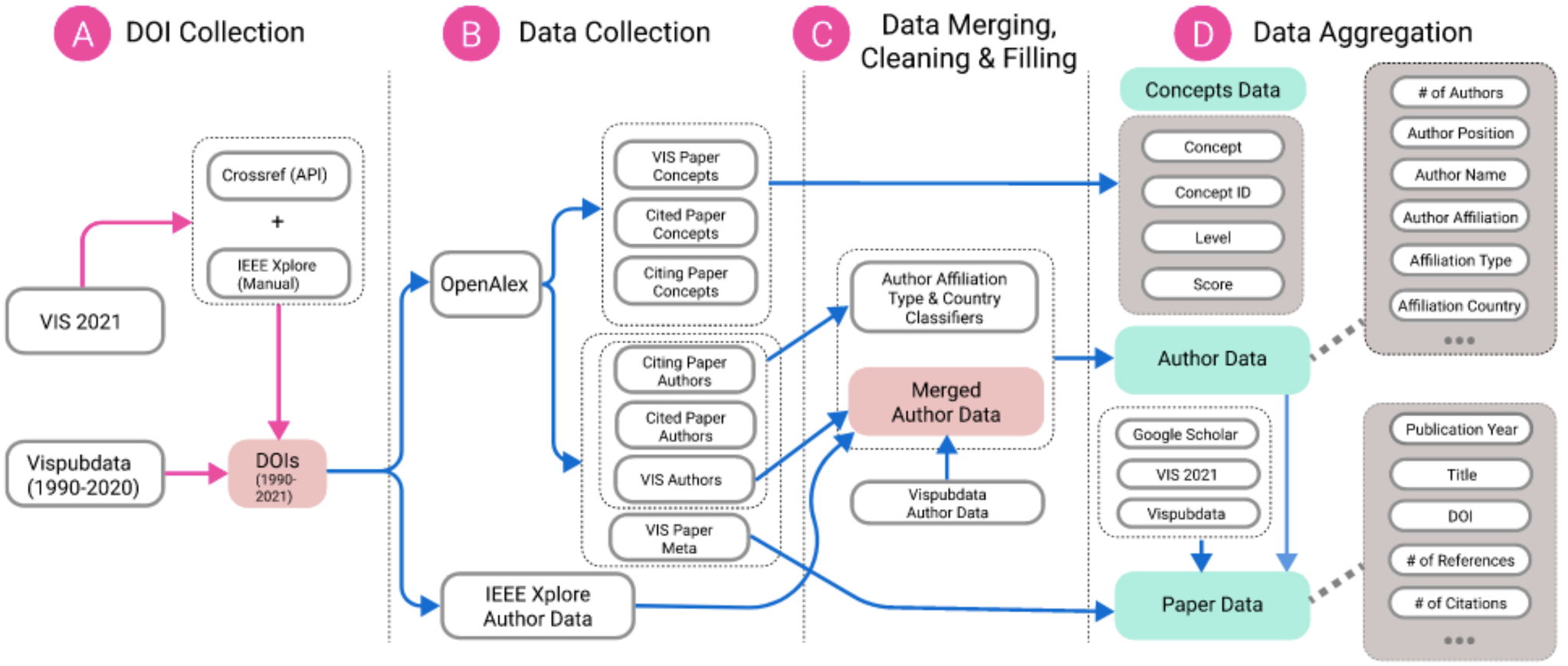


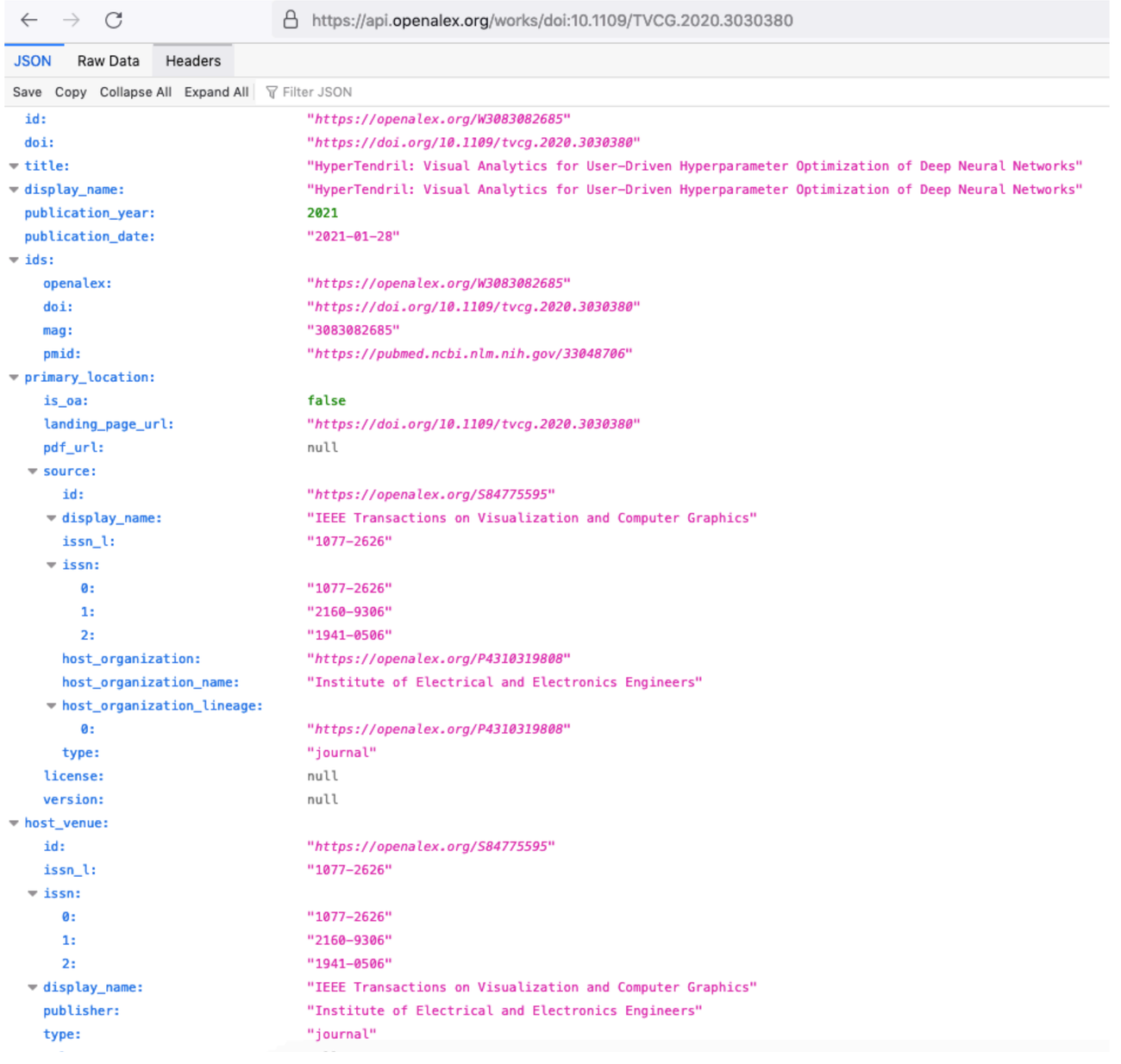
Fig. 2: Data processing pipeline. We started from Vispubdata and VIS2021 to get the paper DOIs, with which we obtained data on relevant papers from OpenAlex and IEEEXplore. We then merged and cleaned data and filled in missing data. Some of the paper data, for example, whether a paper is a cross-country or a cross-type collaborative paper, came from author data. The final outputs were three major data files: authors, concepts, and paper meta.

Tedious: 2 months

Finding the appropriate data source

Unreliable initial API

Data cleaning, filling & merging


https://api.openalex.org/works/doi:10.1109/TVCG.2020.3030380
JSON
Raw Data
Headers
Save
Copy
Collapse All
Expand All
Filter JSON
id: "https://openalex.org/W3083082685"
doi: "https://doi.org/10.1109/tvcg.2020.3030380"
title: "HyperTendril: Visual Analytics for User-Driven Hyperparameter Optimization of Deep Neural Networks"
display_name: "HyperTendril: Visual Analytics for User-Driven Hyperparameter Optimization of Deep Neural Networks"
publication_year: 2021
publication_date: "2021-01-28"
ids:
openalex: "https://openalex.org/W3083082685"
doi: "https://doi.org/10.1109/tvcg.2020.3030380"
mag: "3083082685"
pmid: "https://pubmed.ncbi.nlm.nih.gov/33048706"
primary_location:
is_oa: false
landing_page_url: "https://doi.org/10.1109/tvcg.2020.3030380"
pdf_url: null
source:
id: "https://openalex.org/S84775595"
display_name: "IEEE Transactions on Visualization and Computer Graphics"
issn_l: "1077-2626"
issn:
0: "1077-2626"
1: "2160-9306"
2: "1941-0506"
host_organization: "https://openalex.org/P4310319808"
host_organization_name: "Institute of Electrical and Electronics Engineers"
host_organization_lineage:
0: "https://openalex.org/P4310319808"
type: "journal"
license: null
version: null
host_venue:
id: "https://openalex.org/S84775595"
issn_l: "1077-2626"
issn:
0: "1077-2626"
1: "2160-9306"
2: "1941-0506"
display_name: "IEEE Transactions on Visualization and Computer Graphics"
publisher: "Institute of Electrical and Electronics Engineers"
type: "journal"

20 hours, >10k lines

https://github.com/hongtaoh/32vis/tree/master/workflow/scripts

deprecated	updating site on Tue Jun 28 15:32:24 CDT 2022	10 months ago
CLASS_country.py	updating site on Sat Jun 25 17:35:47 CDT 2022	10 months ago
CLASS_type.py	updating site on Sat Jun 25 17:35:47 CDT 2022	10 months ago
get_HT_cleaned_author_df.py	updating site on Wed Jun 22 22:06:13 CDT 2022	10 months ago
get_HT_cleaned_paper_df.py	my first commit	10 months ago
get_gscholar_data.py	updating site on Mon Jun 20 18:33:55 CDT 2022	10 months ago
get_ieee_author_and_paper_title.py	updating site on Sat Jun 25 17:35:47 CDT 2022	10 months ago
get_merged_author_df.py	updating site on Sun Jul 3 17:39:43 CDT 2022	10 months ago
get_openalex_citation_dfs.py	my first commit	10 months ago
get_openalex_dfs.py	my first commit	10 months ago
get_openalex_reference_dfs.py	my first commit	10 months ago
get_papers_to_study.py	my first commit	10 months ago
get_titles_2021.py	my first commit	10 months ago
get_vispd_openalex_match_1.py	my first commit	10 months ago
get_vispd_openalex_match_2.py	my first commit	10 months ago
get_vispd_plus.py	updating site on Sat Jun 25 17:35:47 CDT 2022	10 months ago
get_vispd_plus_good_papers.py	my first commit	10 months ago
get_wos_id.py	my first commit	10 months ago
plot_data_author_chord_diagram_data.py	my first commit	10 months ago
plot_sankey_data.py	my first commit	10 months ago
plot_top_concepts_trends.py	my first commit	10 months ago
plot_vis_concepts_cooccurance_data.py	my first commit	10 months ago
scrape_award_papers.py	my first commit	10 months ago

```
        # ieee author affiliation. The only difference is that if ieee is nan,
        # i get the data from openalex raw string
        DF['IEEE Author Affiliation Filled'] = np.where(
                DF['IEEE Author Affiliation'].notnull(),
                DF['IEEE Author Affiliation'],
                DF['Raw Affiliation String'],
        )
        # rename columns
        DF.rename(columns={
                'IEEE Year': 'Year',
                'IEEE DOI': 'DOI',
                'IEEE Title': 'Title',
                'IEEE Author Affiliation': 'IEEE Author Affiliation Updated',
                'First Institution Name': 'First Institution Name Updated',
                'Raw Affiliation String': 'Raw Affiliation String Updated',
                # 'First Institution Type': 'First Institution Type Updated',
                # 'First Institution Country Code': 'First Institution Country Code Updated',
        }, inplace=True)
        return DF

def main():
        ieee = update_ieee_orig(ieee_orig)
        diff_dois = get_diff_dois(ieee, alex)
        alex_new = get_alex_new(ieee, alex, diff_dois)
        ieee_sorted, alex_sorted = get_sorted_dfs(ieee, alex_new, papers)
        concat_df = get_concat_df(ieee_sorted, alex_sorted, papers)
        concat_df = update_with_vispubdata_author_data(vispd, concat_df)
        concat_df = update_concat_df(concat_df)
        manual_update_concat_df(concat_df)
        concat_df_filled = get_concat_df_filled(concat_df)
        concat_df_filled = recode_to_edu(concat_df_filled)
        concat_df_filled = get_alex_raw_string_correct(concat_df_filled)
        concat_df_filled = add_binary_type(concat_df_filled)
        concat_df_filled = check_delete_rename(concat_df_filled)
        return concat_df_filled

if __name__ == '__main__':
        vispd = pd.read_csv(VISPUBDATA)
        doi_year_dict, doi_title_dict = get_dicts(VISPUBDATA)
        ieee_orig = pd.read_csv(IEEE_AUTHOR)
        alex = pd.read_csv(OPENALEX_AUTHOR)
        papers = read_txt(PAPERS_TO_STUDY)
        df = main()
        df.to_csv(MERGED_AUTHOR_DF, index=False)
```

Affiliation type & country classification using ML

	aff	label_raw	label	binary_label_raw	binary_label
0	Computer Science Department, Stanford Universi...	education	0	education	0
1	Dept. of Comput Sci., Maryland Univ., College ...	education	0	education	0
2	IBM Sci. Center, Los Angeles, CA, USA	company	1	non-education	1
3	Technische Univ. Eindhoven	education	0	education	0
4	Los Alamos National Laboratory and Lawrence Li...	facility	2	non-education	1
5	Los Alamos national Laboratory	facility	2	non-education	1
6	Lawrence livermore National Laboratory	facility	2	non-education	1
7	Rwth Aachen	education	0	education	0
8	Comput. Sci. Div., California Univ., Berkeley, CA	education	0	education	0
9	School of Information Management and Systems, ...	education	0	education	0

Affiliation type accuracy: 95%

Affiliation country/region accuracy: 93%

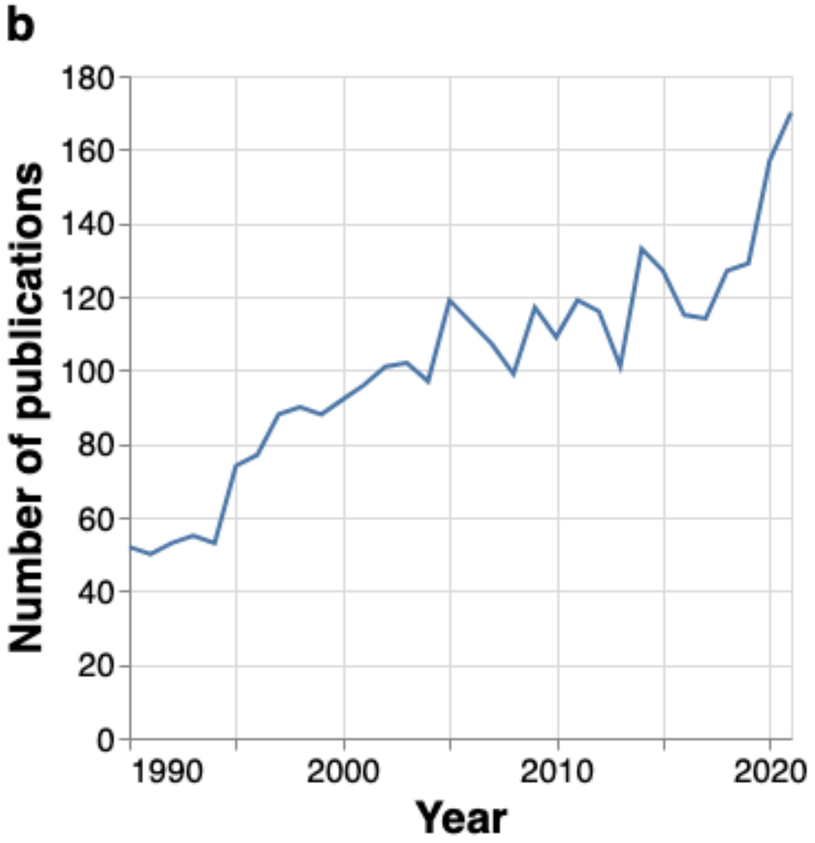
	aff	label_raw	label
0	Computer Science Department, Stanford Universi...	US	0
1	Dept. of Comput Sci., Maryland Univ., College ...	US	0
2	IBM Sci. Center, Los Angeles, CA, USA	US	0
3	Technische Univ. Eindhoven	NL	1
4	Los Alamos National Laboratory and Lawrence Li...	US	0
5	Los Alamos national Laboratory	US	0
6	Lawrence livermore National Laboratory	US	0
7	Rwth Aachen	DE	2
8	Comput. Sci. Div., California Univ., Berkeley, CA	US	0
9	School of Information Management and Systems, ...	US	0

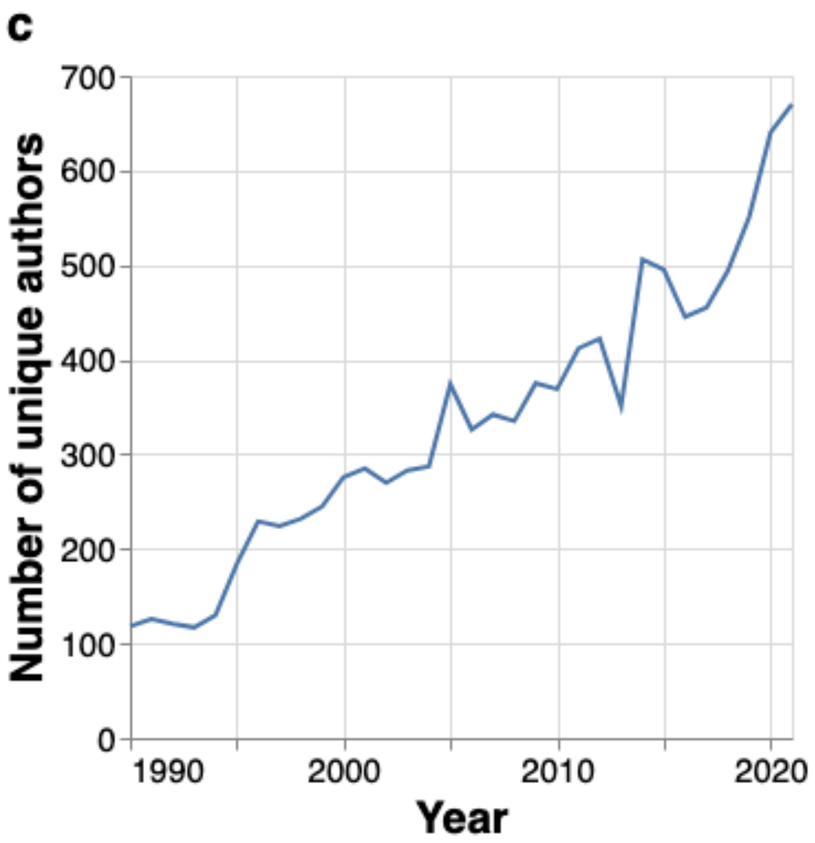
Results

Results: General Trends

Popular

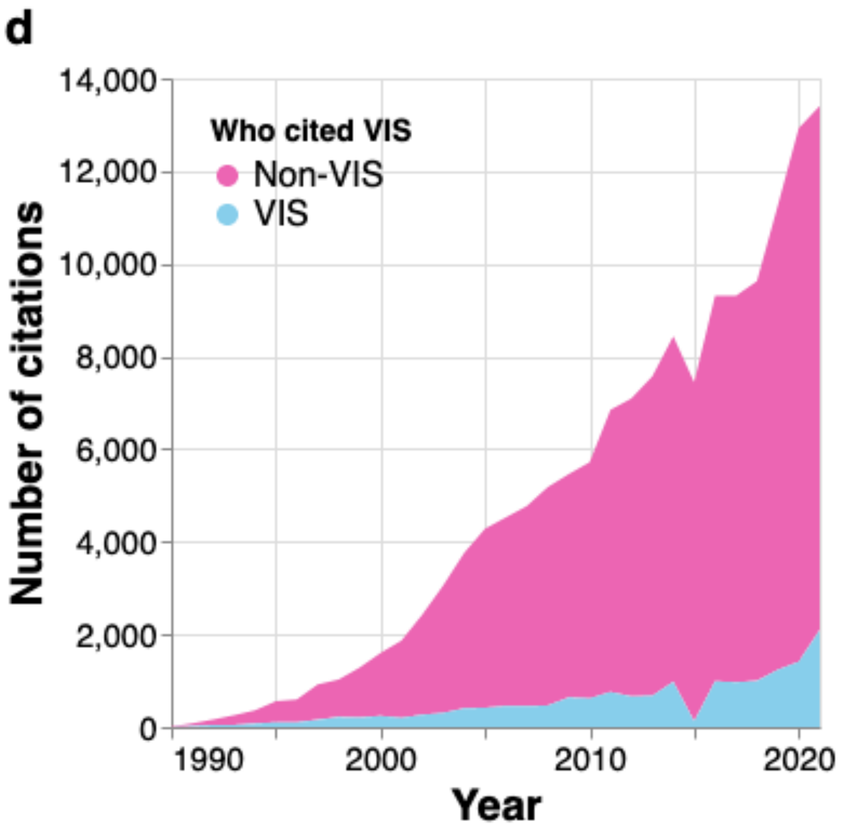
- 1. Pubs
- 2. Unique authors

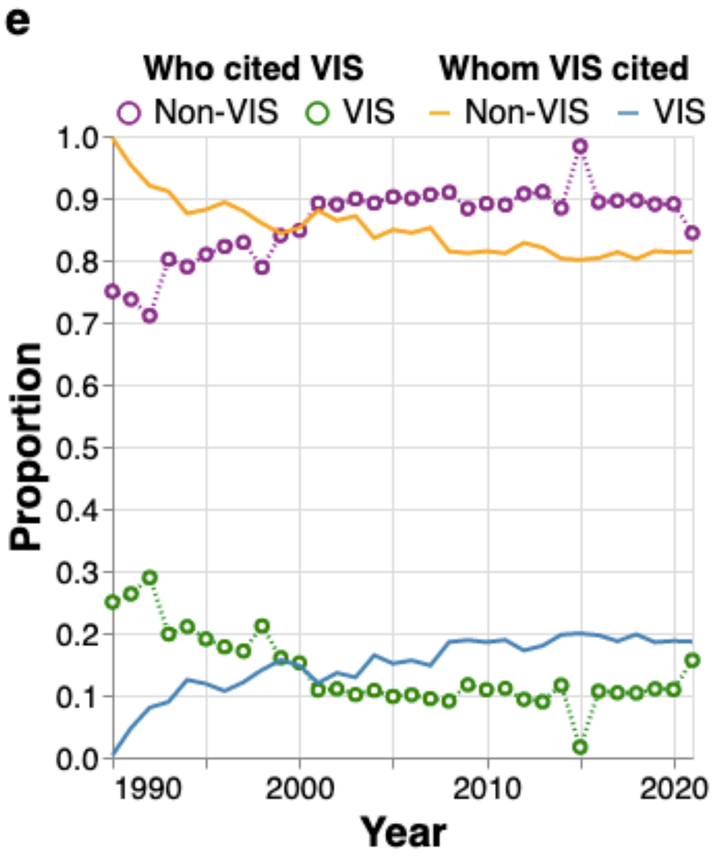




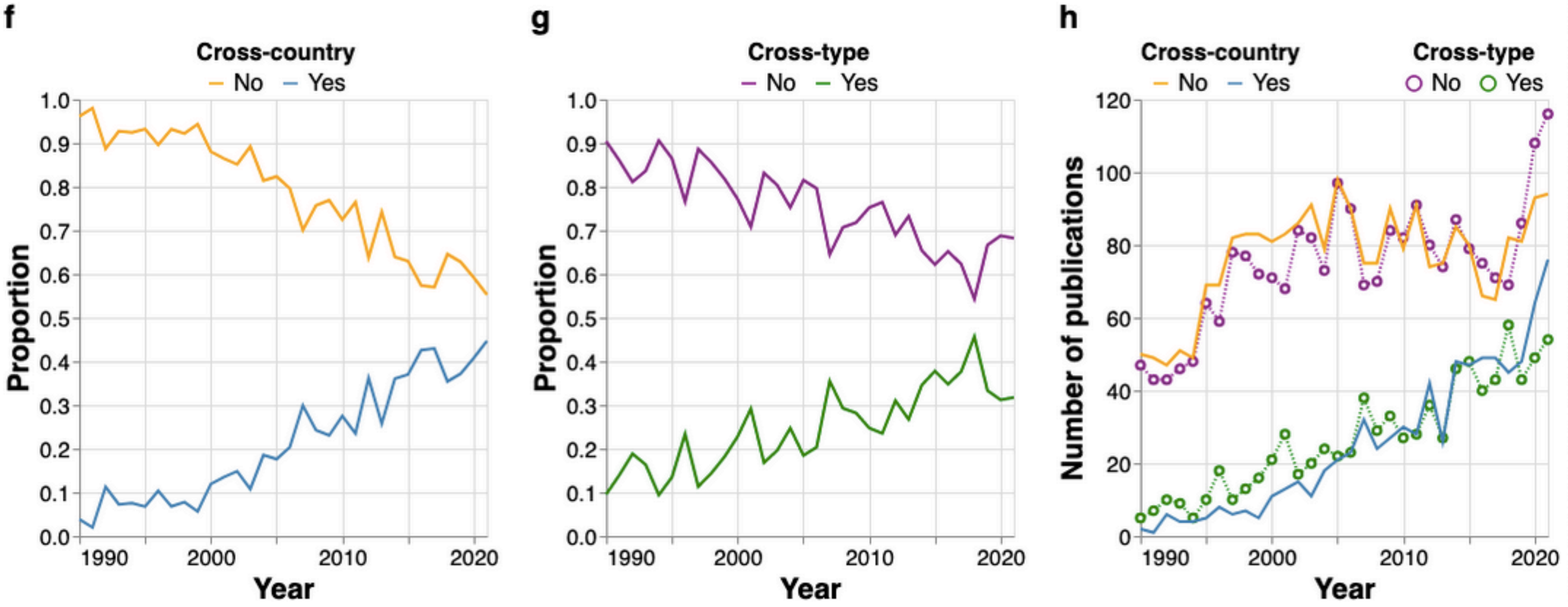
Impactful

- Increasing citations from non-vis

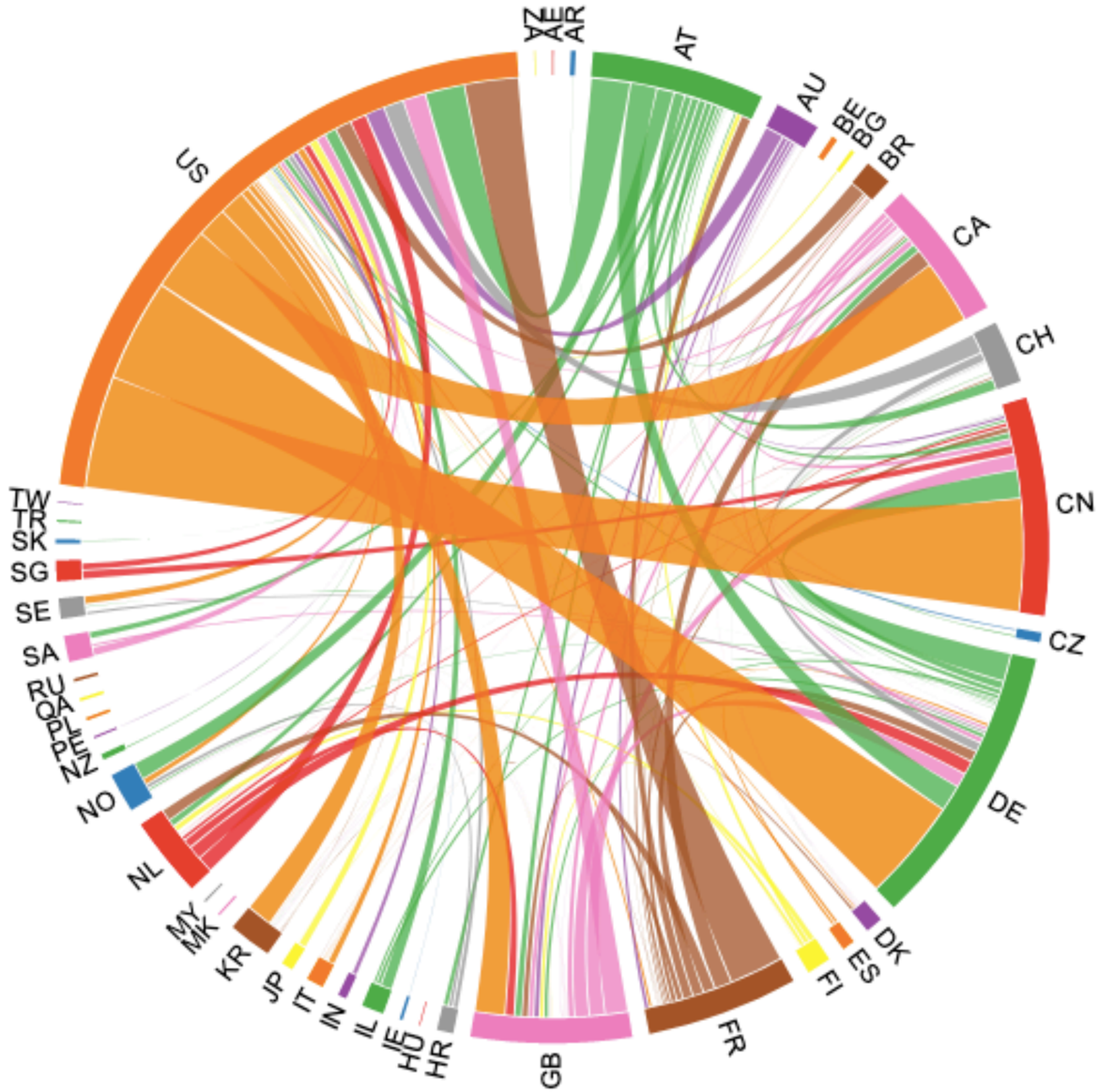




Collaborative



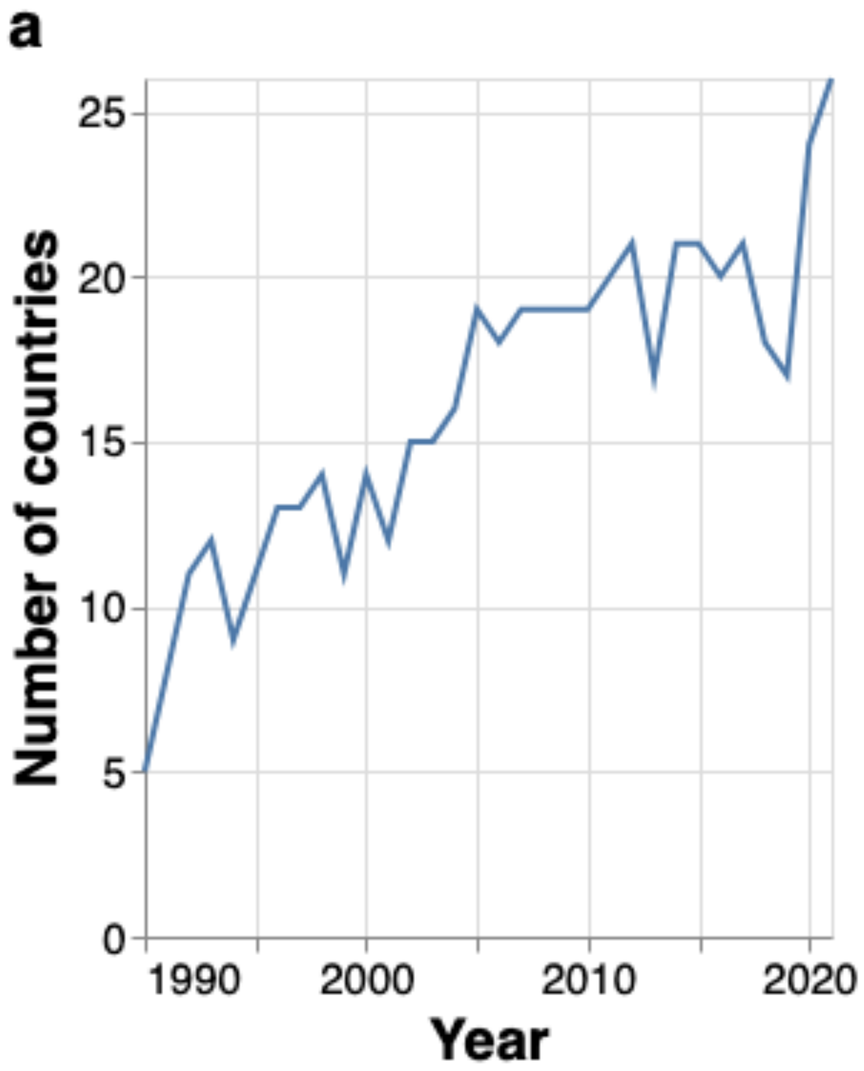
Cross-country collaboration, concentrated



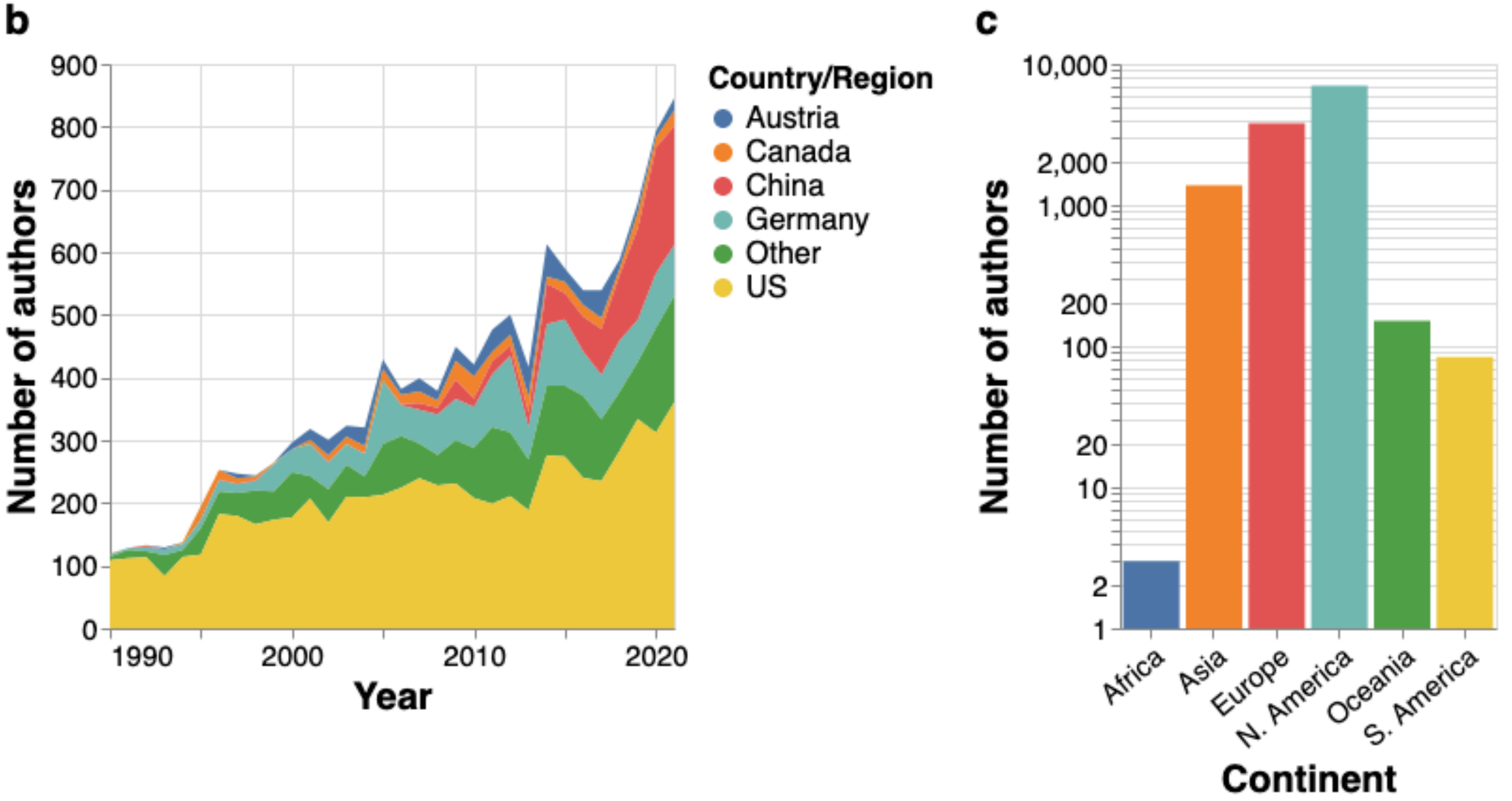
Results: Authors

Authors-Geographic

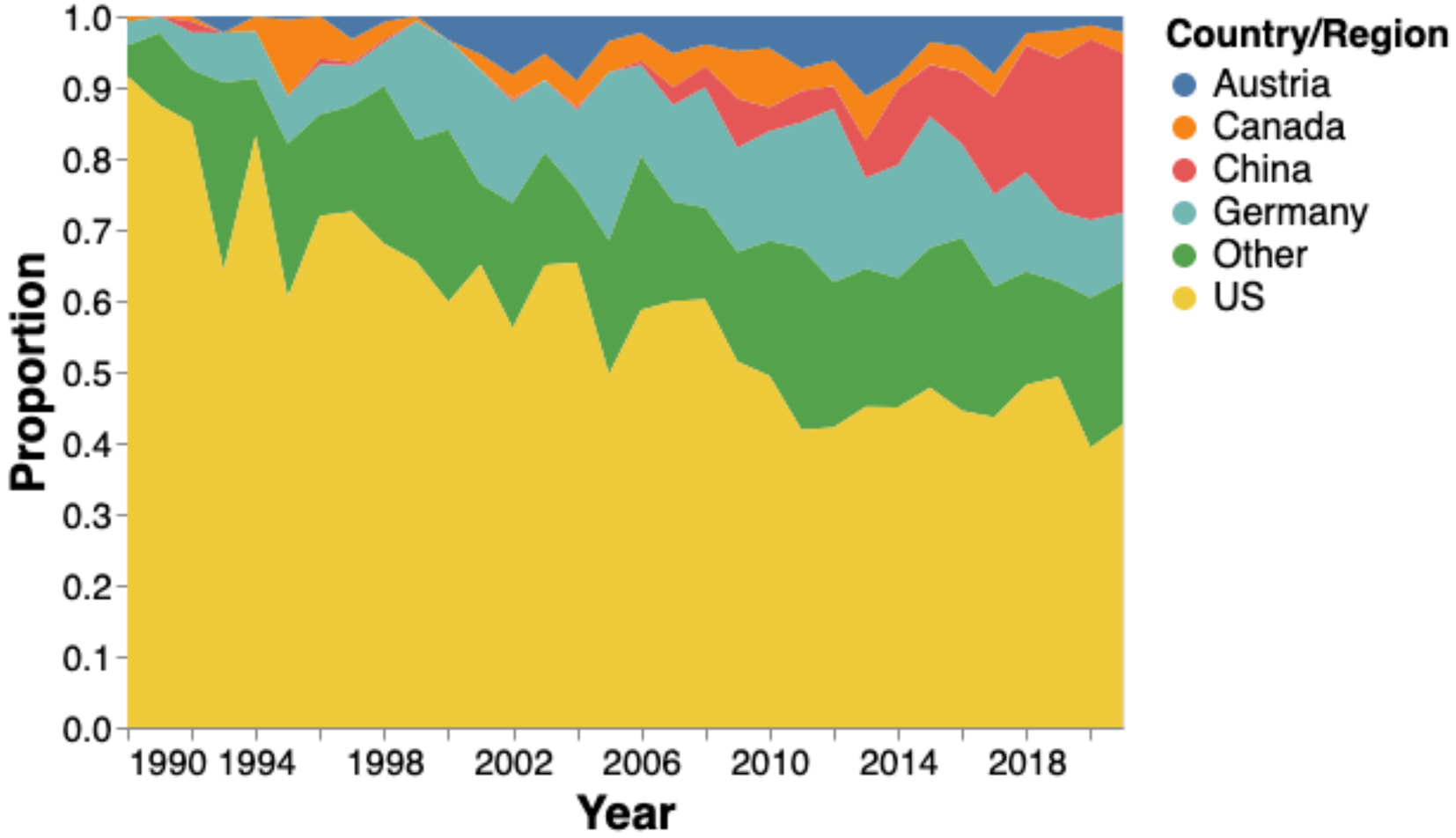
Increasing # of participating countries



Concentration

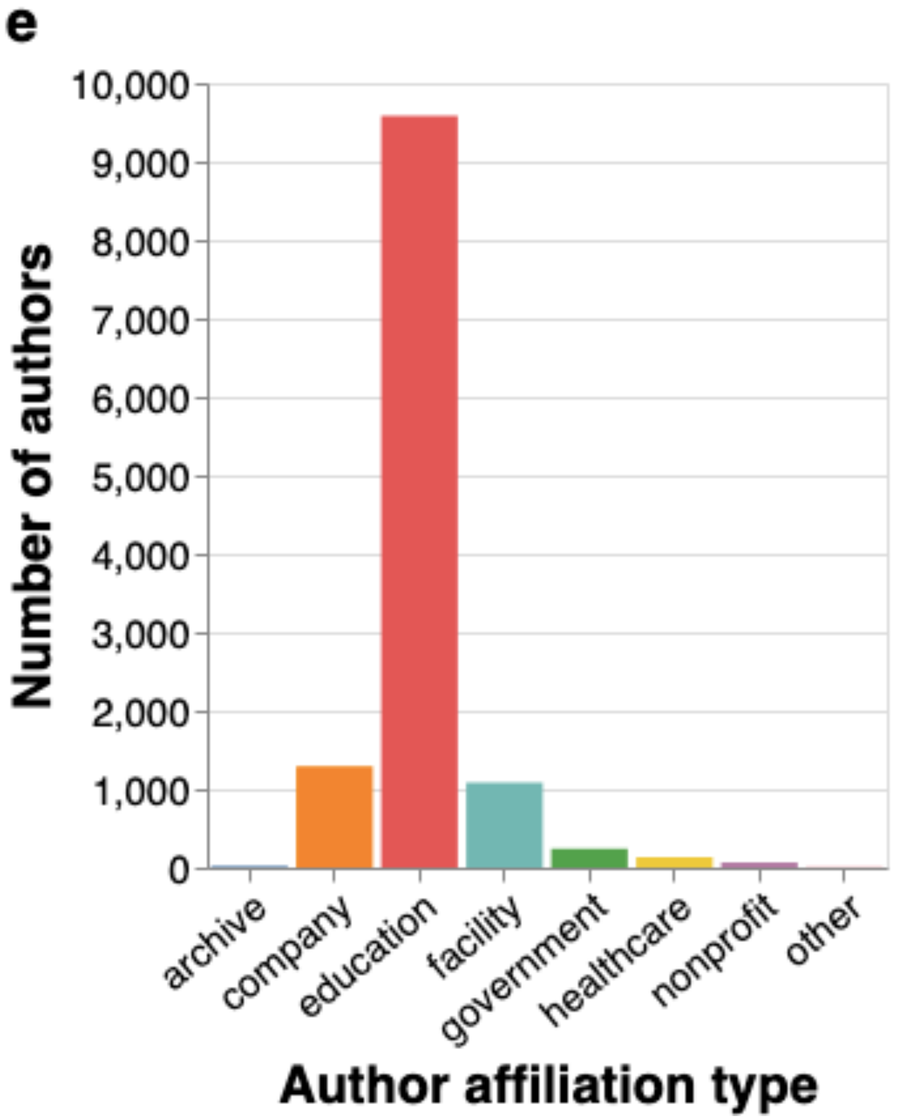


Redistributions

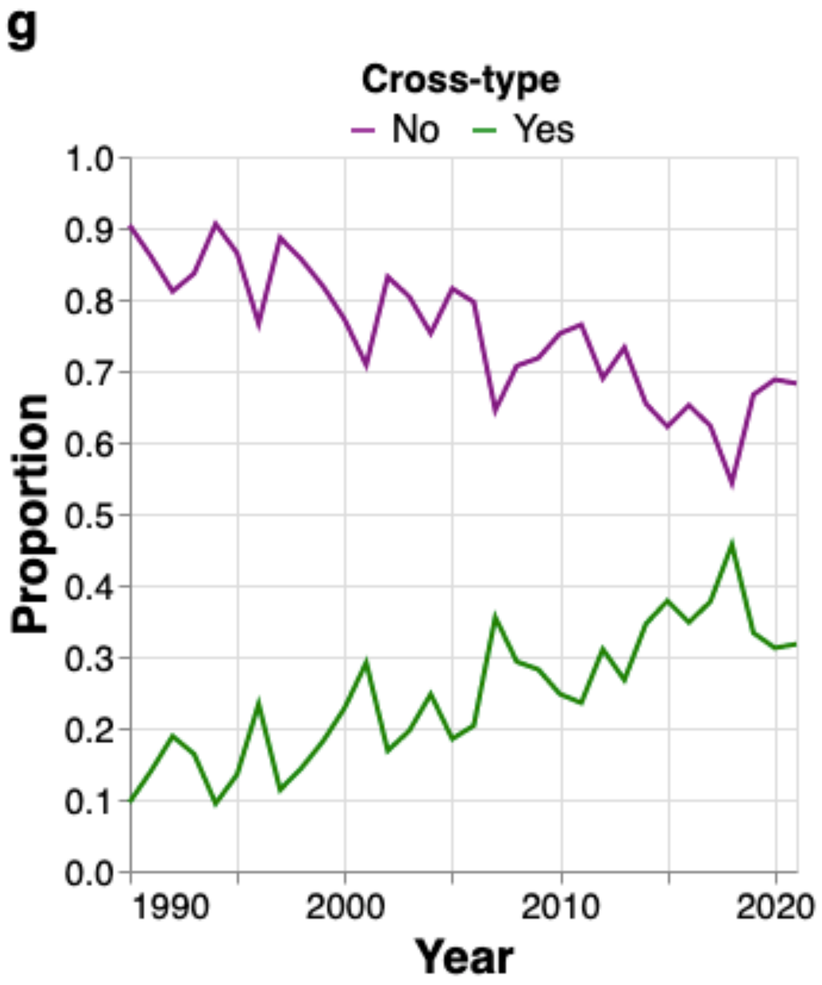


Authors-Affiliation types

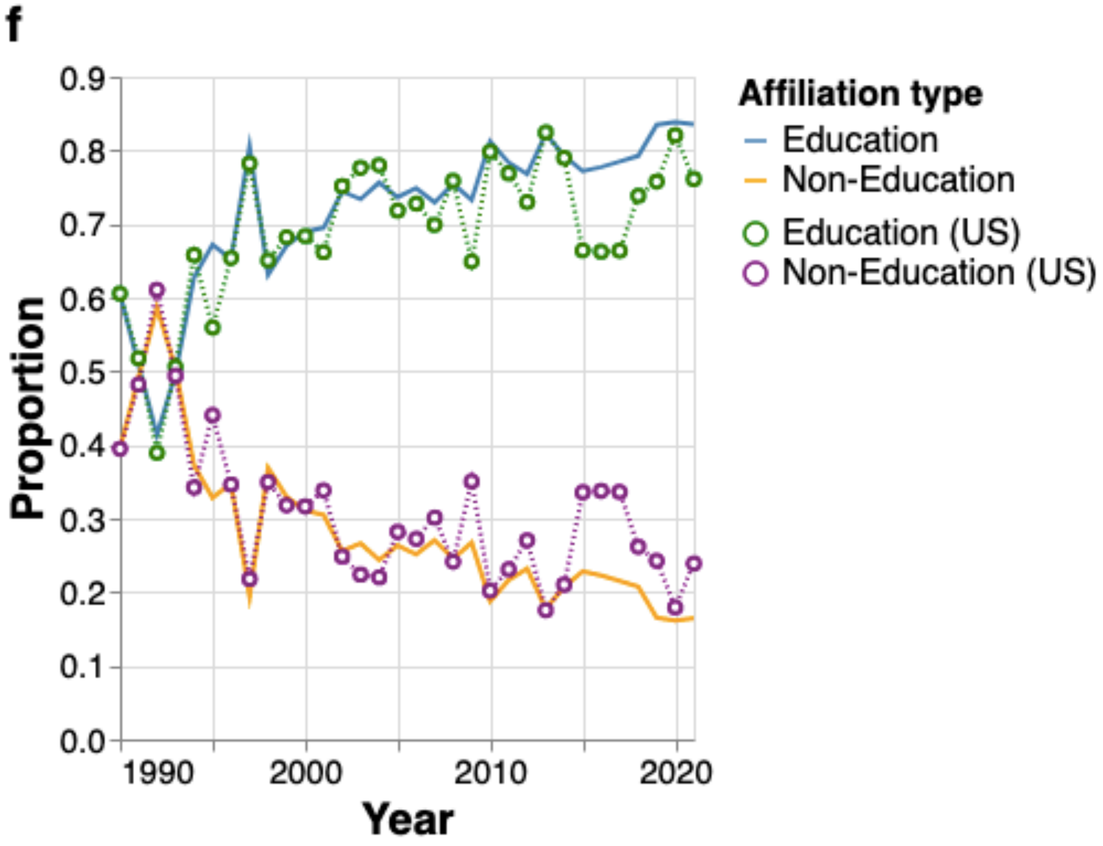
Universities dominate (77%)

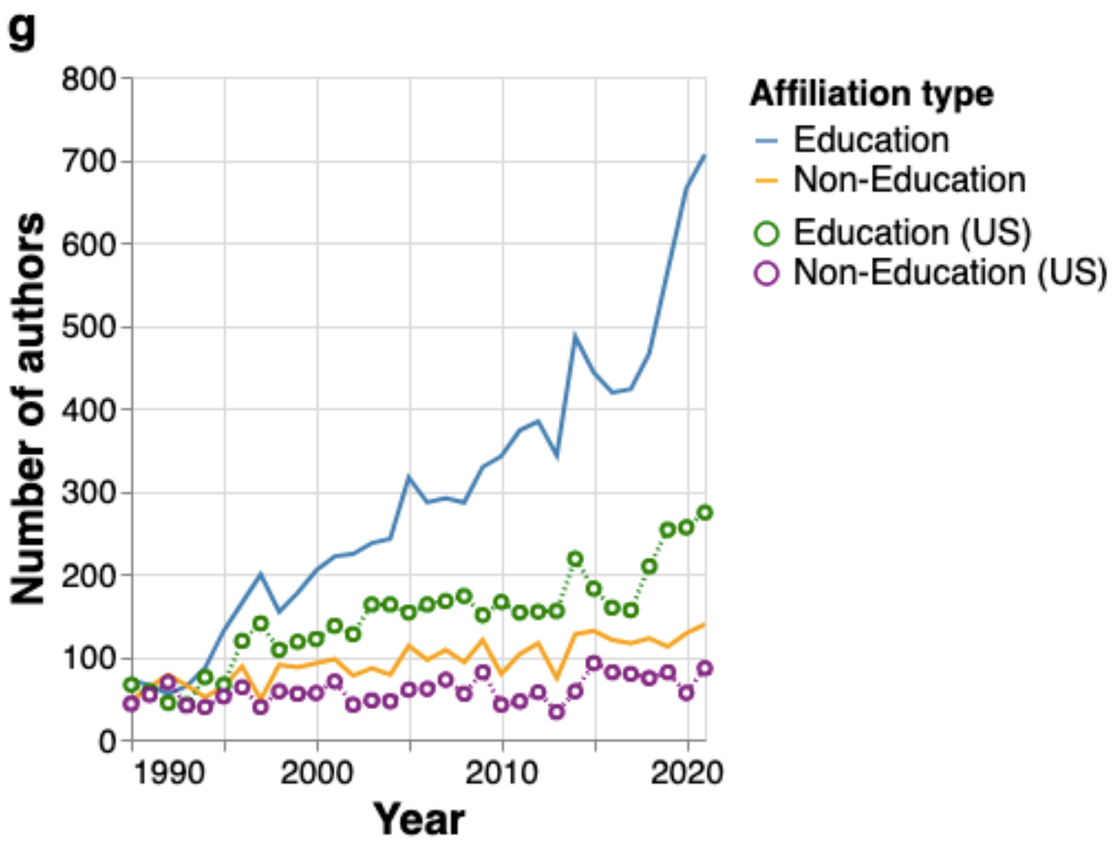


Collaborations Increase but non-edu stagnate



Collaborations Increase but non-edu stagnate





Results: Fields of Study

About, built upon and impacting CS and Math

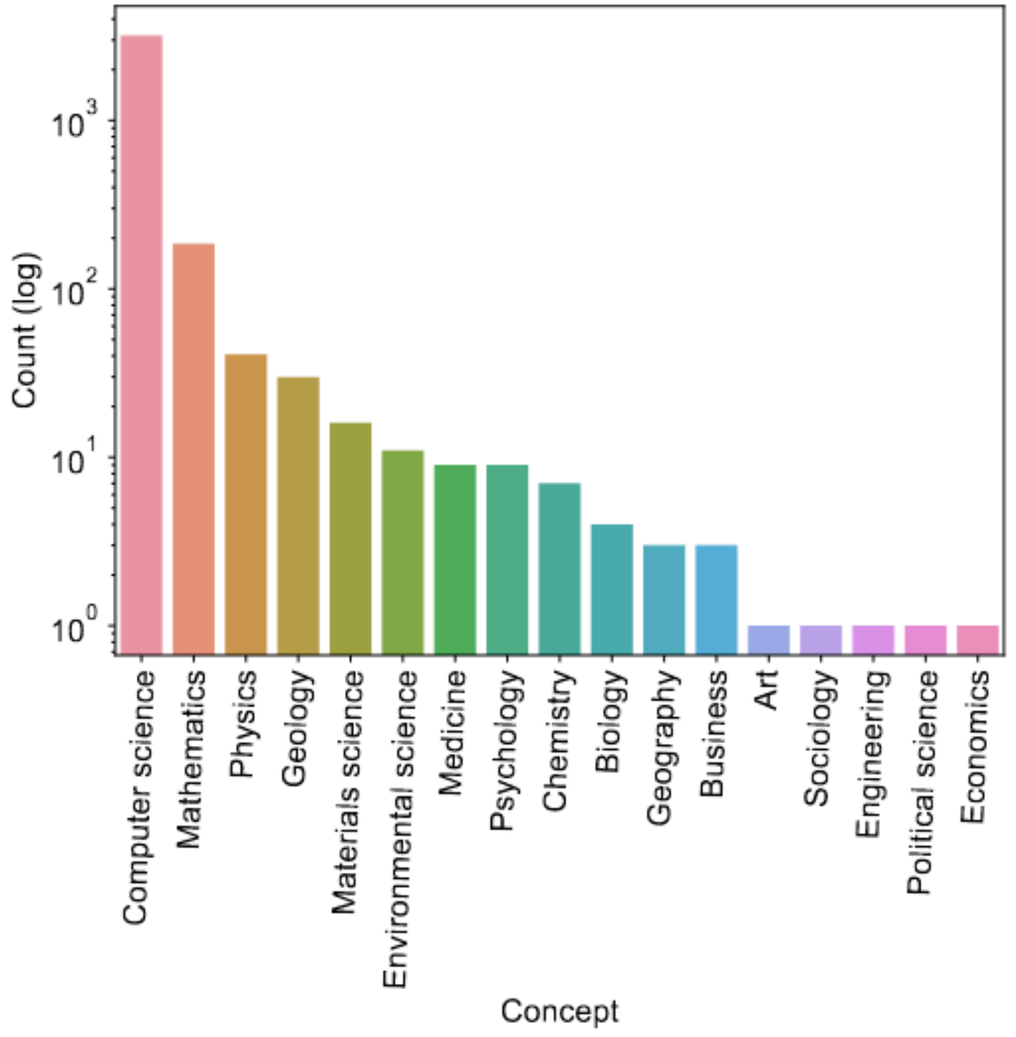


Figure 2: Distribution of L0 concepts among VIS papers

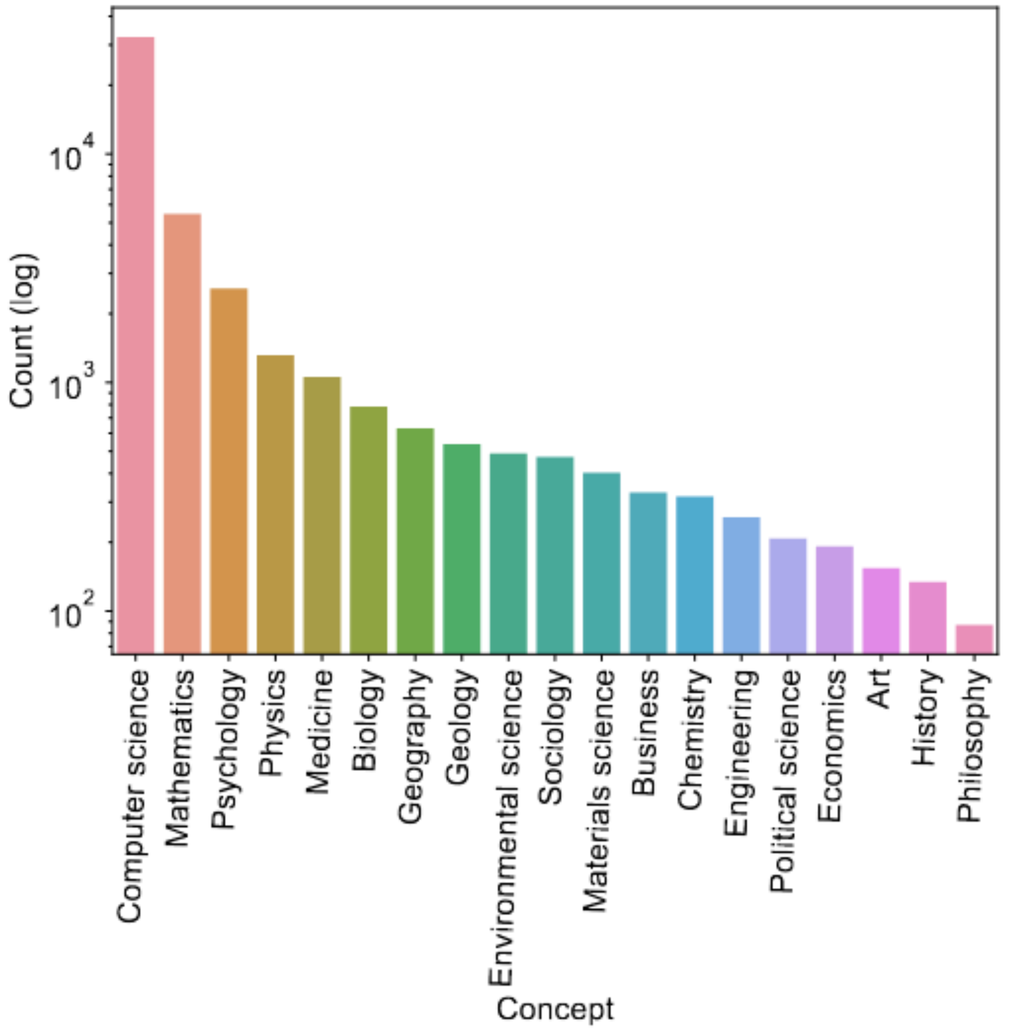


Figure 3: Distribution of L0 concepts among referenced papers

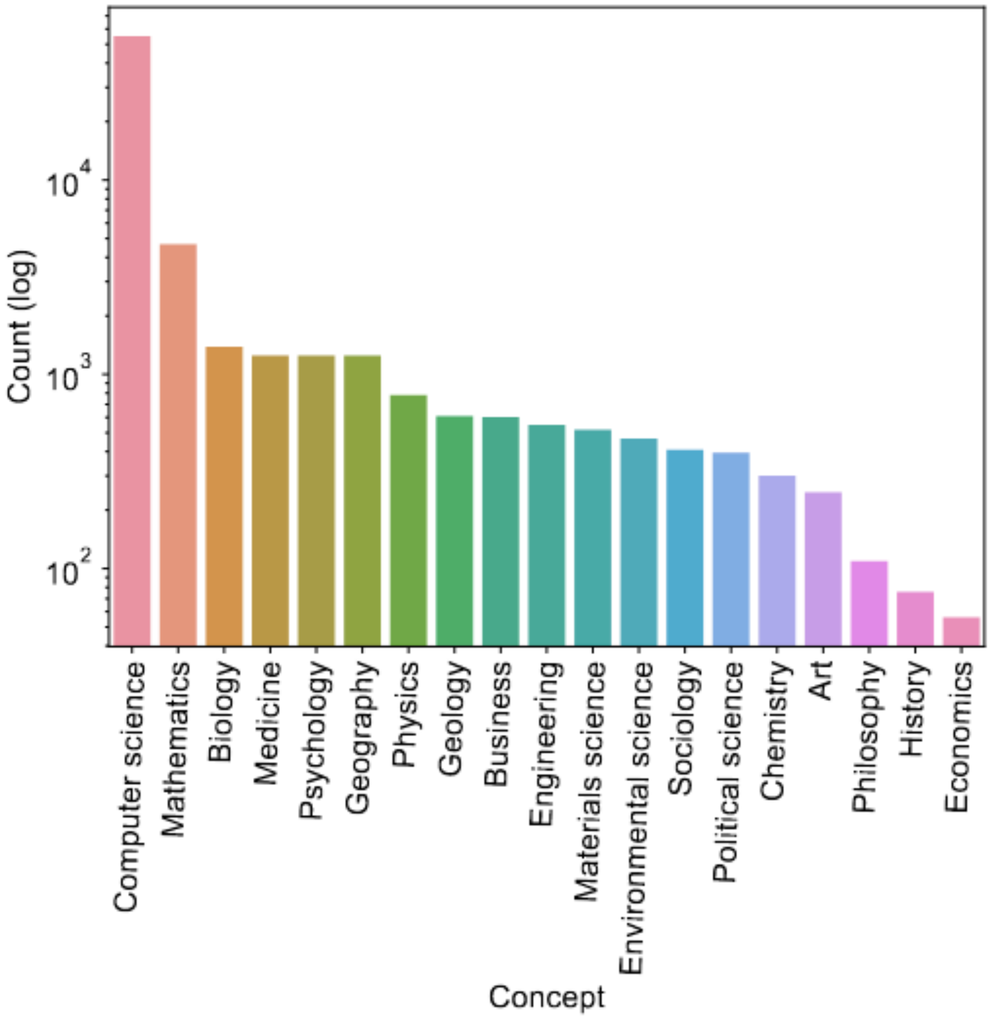


Figure 4: Distribution of L0 concepts among citing papers

Results: Citations

Citation concentration

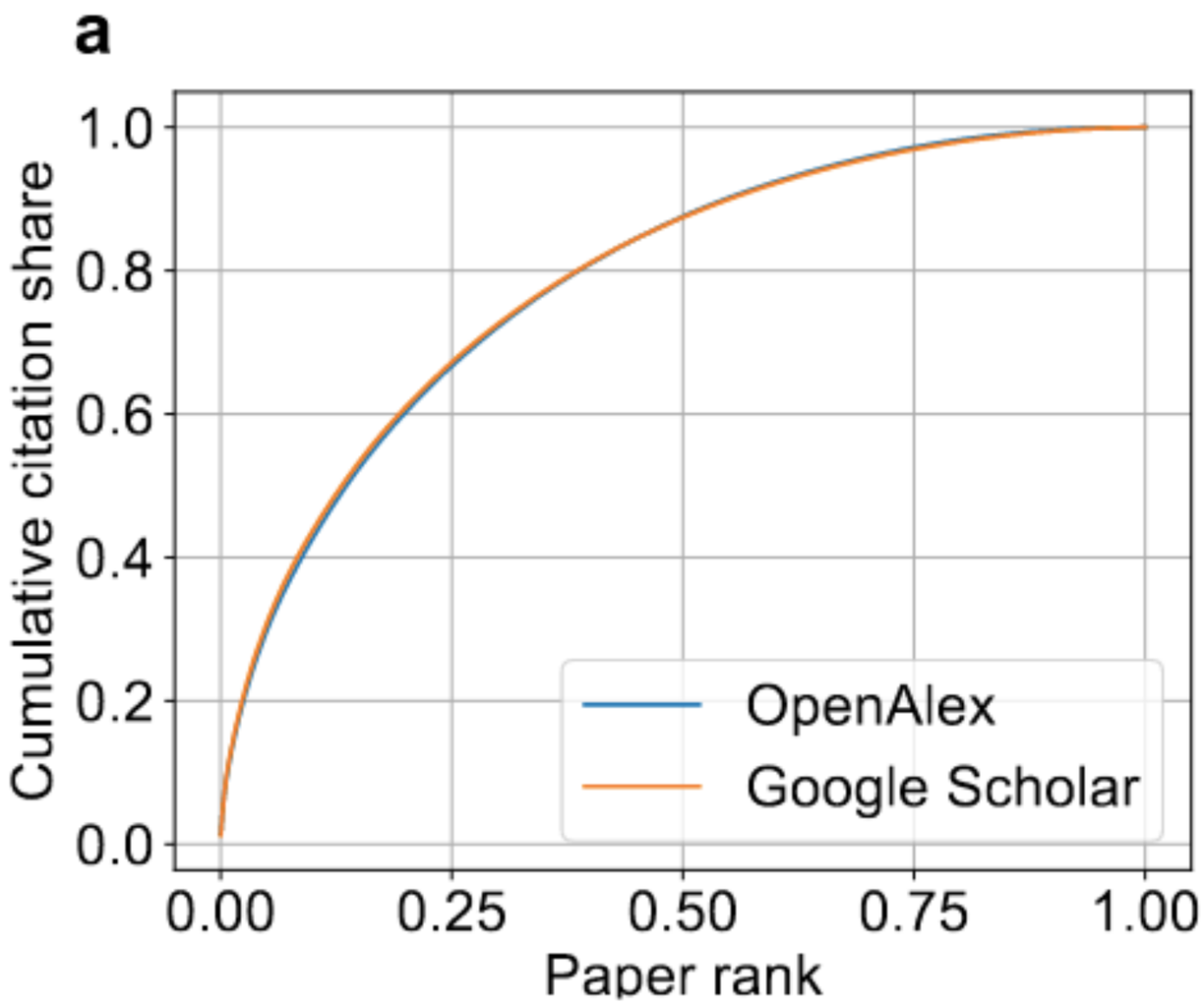


Table 1: Regression results with OpenAlex citations. Estimates are unstandardized coef cients with standard error and p values.

	Dependent variable:
	Number.of.Citations
Year.Distance.from.2020	2.507***
	(0.370)
ConferenceSciVis	−42.747***
	(5.646)
ConferenceVAST	−14.599***
	(4.919)
ConferenceVis	−37.252***
	(4.298)
PaperTypeJ	22.250***
	(4.866)
Number.of.Authors	−0.289
	(0.884)
Cross.type.CollaborationTrue	−1.323
	(3.630)
Cross.country.CollaborationTrue	−3.952
	(3.826)
With.US.AuthorsTrue	4.049
	(3.076)
AwardTrue	28.179***
	(6.705)
Constant	28.486***
	(8.196)
Observations	3,070
R^2	0.057

Recap

- VIS becoming popular, impactful and collaborative
- Authors: geographically, diverse and yet concentrated
- Authors: aff types, universities dominate
- Fields of study: VIS about, built upon and impacting CS and math
- Citations: earlier works, journal papers, award-winning papers had more citations.

32vis.hongtaoh.com

Thank you!