IETF 127 · RASPRG · SAN FRANCISCO · 19 NOVEMBER 2026

Who Wrote the RFCs

Authorship, Affiliation and Community, 1969-2026

Carlos A. Kamienski
Federal University of ABC (UFABC), Brazil

Marcelo Santos
Federal Institute of Sertão Pernambucano (IFSertãoPE), Brazil

24,343	9,848	6,580	66	3,020
person-authorships	RFCs	people	countries	organizations

MOTIVATION

Our goal: a standards process that can be understood

1

Understand the past

Who wrote the RFCs, from where, and for whom

2

Correct the record

Recover authorship data that was never entered

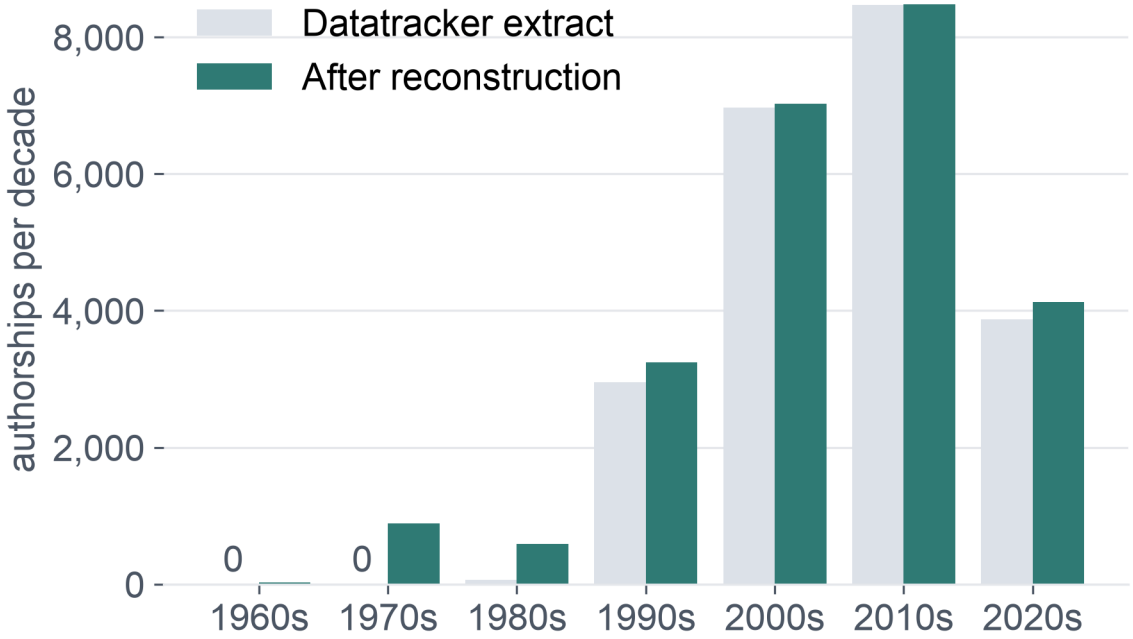
3

Improve the future

Suggest how authorship should be recorded and measured

THE PROBLEM

The Datatracker's structured authorship data begins in 1980



691 RFCs

published 1969-1979 (the ARPANET era) are a subset of legacy RFCs with no Datatracker author record

- The RFCs and the RFC Editor index name their authors
- Analyses built on the Datatracker see a community that begins in 1980

DIAGNOSIS

Not an extraction bug: the records were never entered

1,167 RFCs

- Legacy authorship records were never loaded
- Pre-1990 RFCs predate the Datatracker's document model
- Their authorship exists only in the documents

53 RFCs

- Authorship is keyed to an Internet-Draft, not the published RFC
- A missing link returns no RFC author record

Consequence

- Total: 1,220 RFCs have no API author record
- Solution: look into the documents
- Dataset: we can offer the recovered records back to the Datatracker

METHODOLOGY · DECISION 1 OF 3

Where the API is silent, the RFC text decides

WHY

- 1,220 RFCs had no API author record; name search alone identified under half

WHAT WE DID

- For these RFCs, the authors come from the document
- Evidence is read in this order, with name last:

92%
of authorships come straight from the Datatracker

8%
were rebuilt from the RFC text

1. RFC text
2. E-mail
3. Postal address
4. Affiliation
5. Era window
6. Name

METHODOLOGY · DECISION 2 OF 3

Country = place of work when the RFC was published

WHY

- Not nationality, not current residence: the question is where standards work was done

WHAT WE DID

- Every value records its source, so inferred data stays distinguishable from primary records, as draft-irtf-rasprg-analysing-ietf-data (work in progress) recommends

19%

of country values are inferred, not declared, and flagged as such row by row



■ Declared in the RFC 14,890 ■ E-mail ccTLD 3,180 ■ Manual research 1,561 ■ Inferred 4,712

METHODOLOGY · DECISION 3 OF 3

Judgment calls, stated and recorded

BBN is a company

- A private firm on ARPA contract
- The largest entity of the era

1970-1974 RFCs - BBN status
- **Research: 91% academic**
- **Company: 65% academic**

SRI is not Stanford

- Formally separated from the university in 1970

82 of its 86 authorships postdate the separation

Independents are not companies

- One-person consultancies are classified apart from firms

Rising series
- **1985-89 → 1.2%**
- **Today → 9.1%**

Dataset: every judgment is documented, with its evidence

RESULT

A complete dataset: traceable to its source

24,343
person-authorships

9,848
RFCs: all in the RFC Editor's list

20 RFCs
where the list names an organization, the dataset names the people

100%
with country and affiliation

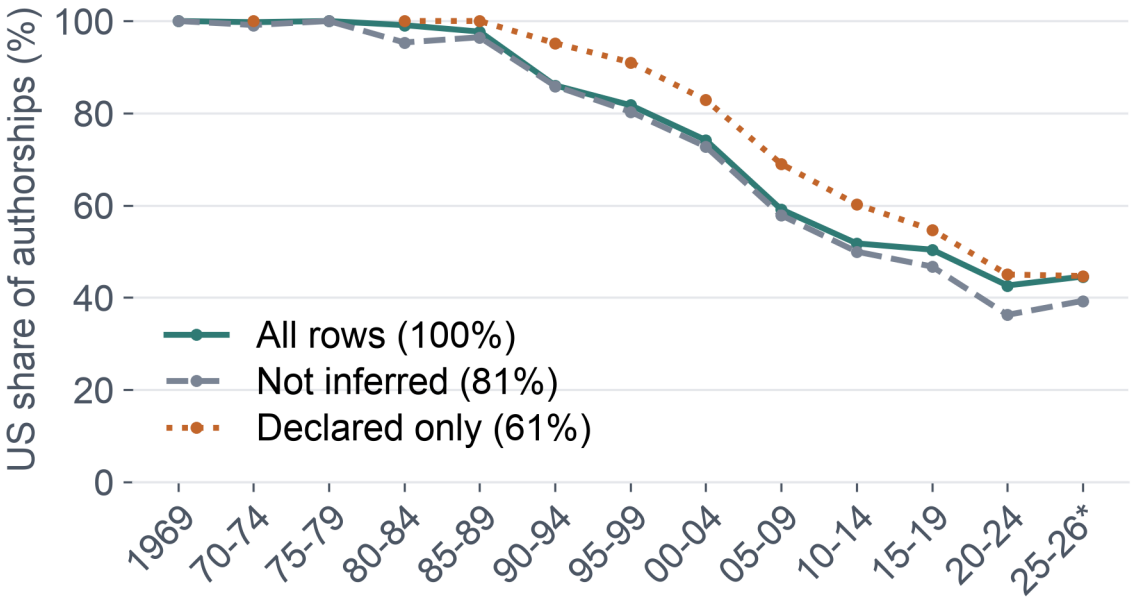
908
ARPANET-era authorships recovered, where there had been none

467
people absent from the dataset before the reconstruction

3.7%
of affiliations not read from text: 522 propagated, 222 researched, 156 from e-mail domain

COVERAGE IS NOT ACCURACY

How much do the inferences move the results?



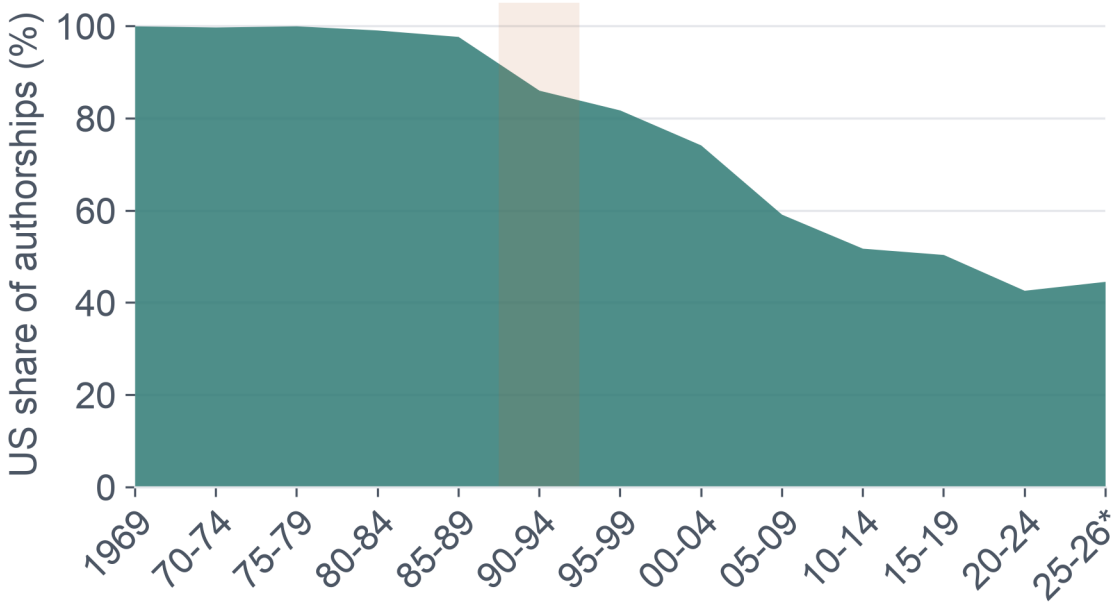
**2025-26 is partial, to October 2026*

Same curve, three standards of evidence

- DATA UNCERTAINTY · Declared-only records retain the broad decline; timing shifts
- 1990-94: 86% US with or without inferred country values
- CLASSIFICATION UNCERTAINTY · BBN as research vs company: 91% to 65% academic

FINDING 1 · GEOGRAPHY

The RFCs had US authors for twenty years



**2025-26 is partial, to October 2026*

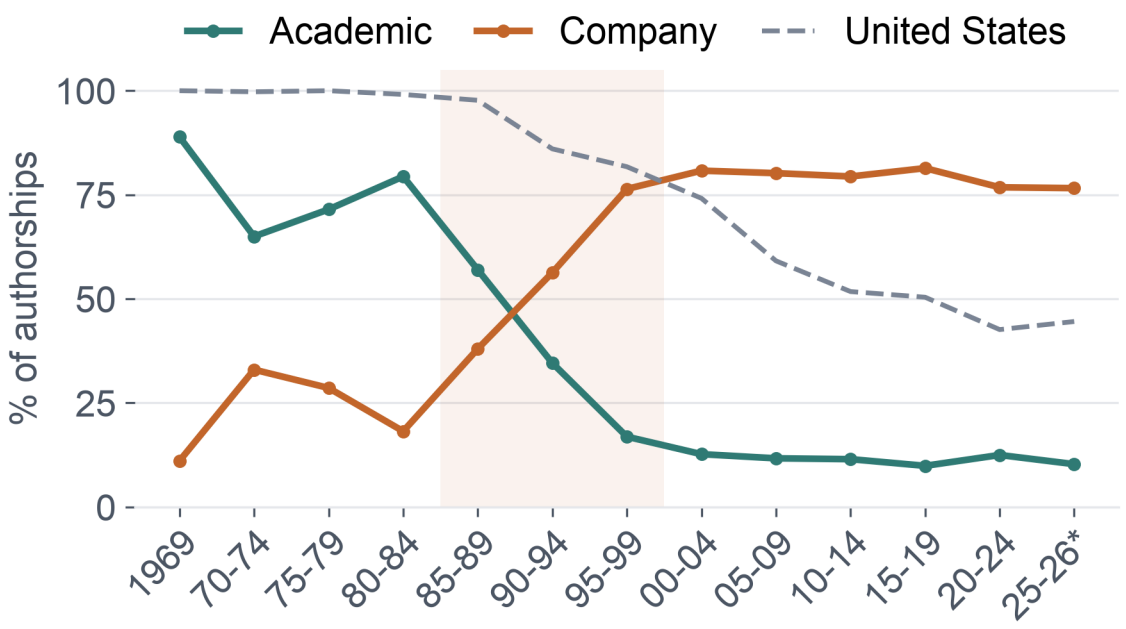
98-100%
US authorships, 1969-1989

43%
US authorships, 2020-24

3 → 19
contributing countries,
1985-89 → 1990-94

FINDING 1 · INSTITUTIONS

Less academic, then less American, in fifteen years



Academic = university + research institute

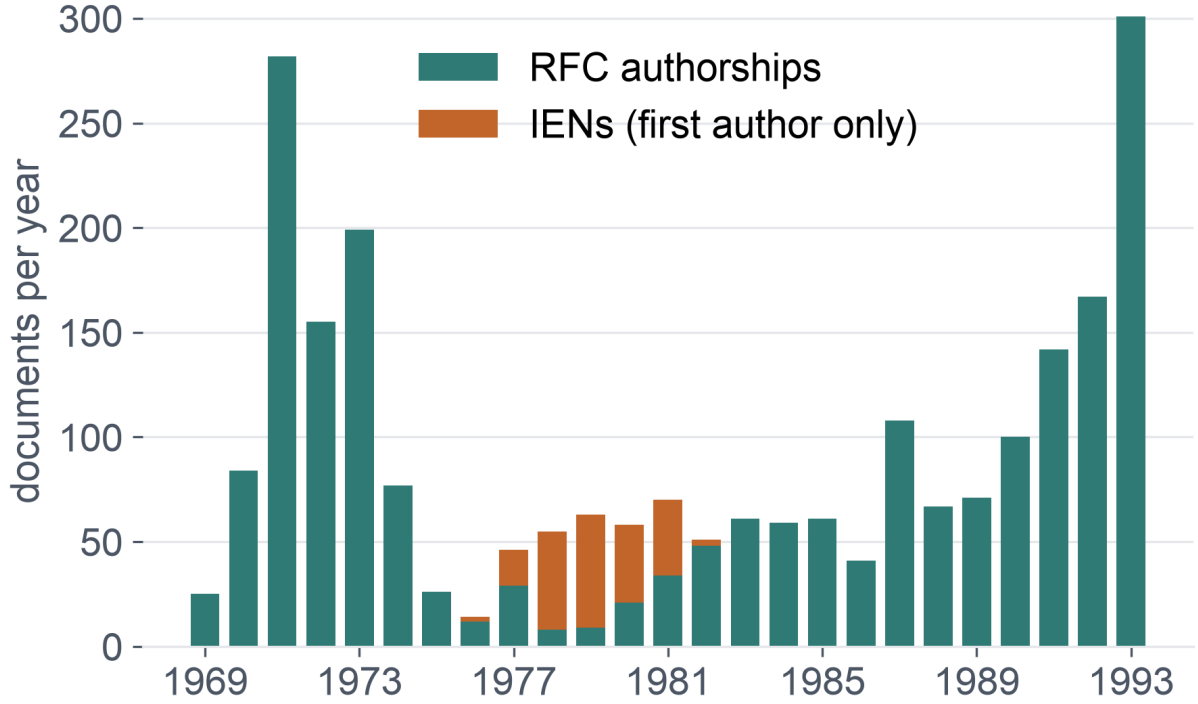
15 years

1980-84 to 1995-99
Academic 79% → 17%
Corporate 18% → 76%
US 99% → 82%

- The academic decline starts one window before the US one
- Commercialization driving both is a hypothesis, not a result

FINDING 2 · EARLY RFC AUTHORS

IENs became an important venue during the RFC slowdown



204 IENs

published from 1976 to 1982, by 57 first authors

- IEN = Internet Experiment Note · a venue for TCP/IP work
- Only 10 of the 57 were early RFC authors
- 22 never wrote an RFC

FINDING 2 · EARLY RFC AUTHORS

Most early RFC authors did not reappear after 1982

10 of 167

early RFC authors appear as first author of an IEN (1976-82)

14 of 167

early RFC authors wrote an RFC again from 1983 on

78.6%

of 1983-92 RFC authorships came from people who debuted in 1983 or later

The change of people happened during the IEN years, before the RFC series resumed

Early RFC authors: first RFC by 1974 · IENs counted by first author only

FINDING 3 · HOW WE COMPARE

Do people who write together share a working group?

Grouping 1: who writes with whom

- Communities found in the co-authorship network (Louvain algorithm)
- 20 large communities hold 80% of the people

Grouping 2: working group

- Each person gets one label
- The working group where they published most

How well do the two groupings match? AMI = Adjusted Mutual Information

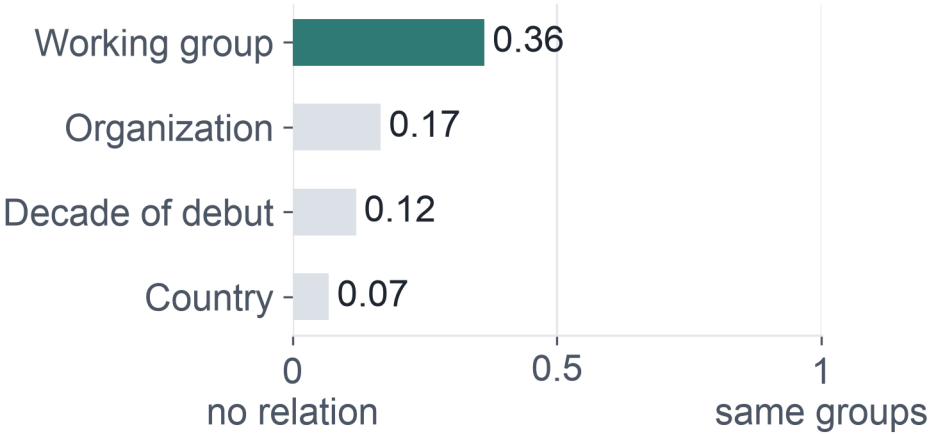


The groupings overlap only partly (AMI = 0.36)

FINDING 3 · COMMUNITIES

Communities do not map one-to-one to working groups

How well each attribute matches the communities



6,580 people · 22,098 co-authorship ties
20 large communities hold 80% of the people

~20 yrs
active span of a large community

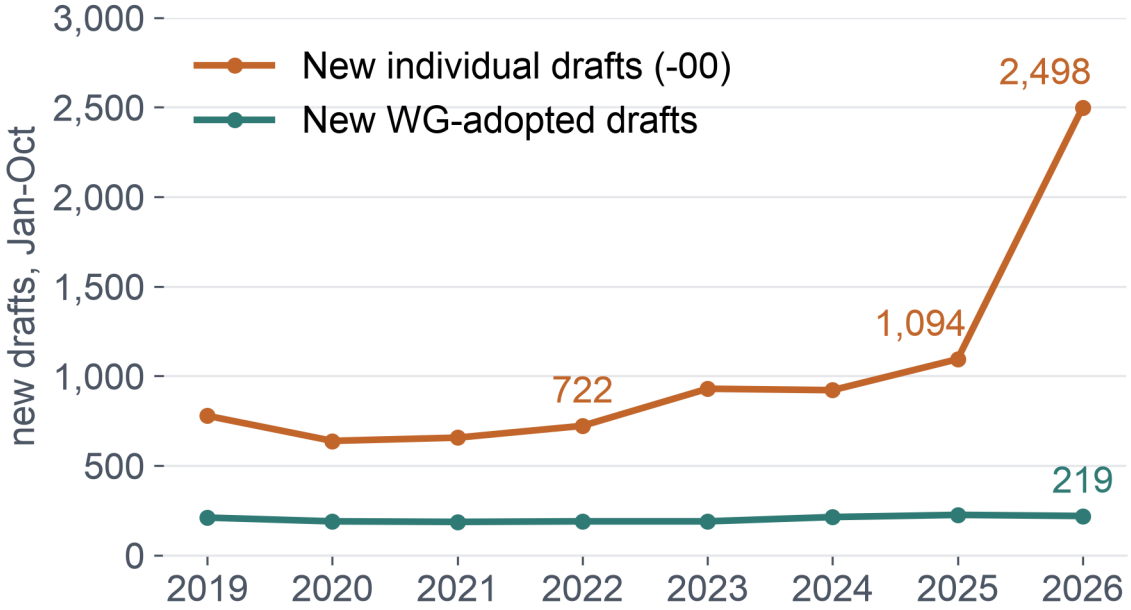
3 yrs
median lifespan of a working group

- Across the 57-year graph, a median large community spans 96 working groups
- Largest communities center on topics such as mpls, sec, sip, ippm, netext and dhc

Large = 50 or more members
A 57-year graph favors long-lived communities

FOR DISCUSSION · 2026

Fewer active RFC authors, more individual draft submissions



January to October of each year · says nothing about AI authorship

-17%

active RFC authors per decade, from 2005-14 to 2015-24

- 52% of 2026 individual drafts have a first-time first author (15-19% in 2019-24)
- 26% carry AI or agent terms in the title
  - Without them, still 2.5× the 2022 volume

BEFORE USING THESE NUMBERS

What this data is and is not

Authorships, not people

- A person with twelve RFCs counts twelve times
- 53% of authors wrote exactly one

Authorship, not participation

- Says nothing about who attends, reviews, chairs or argues on the list

Place of work, not nationality

- A researcher at a US lab counts as the United States

The RFC series only

- Streams are pooled: IETF=80%, Independent=4%, IRTF+IAB=3% (Legacy=12%)
- IENs are not covered

Some labels are inferred

- 19% of countries
- 39% of sector labels come from a generic classifier

Personal data

- 6,580 named people with employer and country
- Names or pseudonyms in the release: an open question

TAKEAWAYS

The RFC series did not begin in 1980

1. Between the late 1980s and late 1990s the RFC series became less academic and less American

2. Most early RFC authors did not reappear after 1982, and IEN authorship is still incomplete

3. Co-author communities overlap with, and extend beyond, working groups

To be open: pipelines and judgments with evidence
The dataset too: with names or pseudonymized is a question for this RG

NEXT STEPS

What should RASPRG take forward?

1

UNDERSTAND THE PAST

Shared definitions for authorship metrics

RASPRG focus

2

CORRECT THE RECORD

Metadata errata and backfilling of historical authorship

Coordinate with RSWG

3

IMPROVE THE FUTURE

Dedicated IDs for RFCXML authors and organizations

Coordinate with RSWG

Which research question should RASPRG develop first?

carlos.kamienski@ufabc.edu.br · marcelo.santos@ifsertao-pe.edu.br

IETF 127 · RASPRG · SAN FRANCISCO · 19 NOVEMBER 2026

Thank you!

Who Wrote the RFCs: Authorship, Affiliation and Community, 1969-2026

Carlos A. Kamienski
Federal University of ABC (UFABC), Brazil

Marcelo Santos
Federal Institute of Sertão Pernambucano (IFSertãoPE), Brazil

BACKUP · CANDIDATE DRAFTS

Seven candidate Internet-Drafts

1	**Metadata errata and backfilling of historical authorship**	RSWG
2	**Dedicated IDs for RFCXML authors and organizations**	RSWG
3	Machine-readable contribution roles	RSWG
4	**Shared definitions for authorship metrics**	RASPRG
5	Provenance model for reconstructed records	RASPRG
6	Privacy guidance for participant datasets	RASPRG
7	Requirements for Datatracker authorship statistics	Tools team