50 Years
# of Standardisation in the Internet

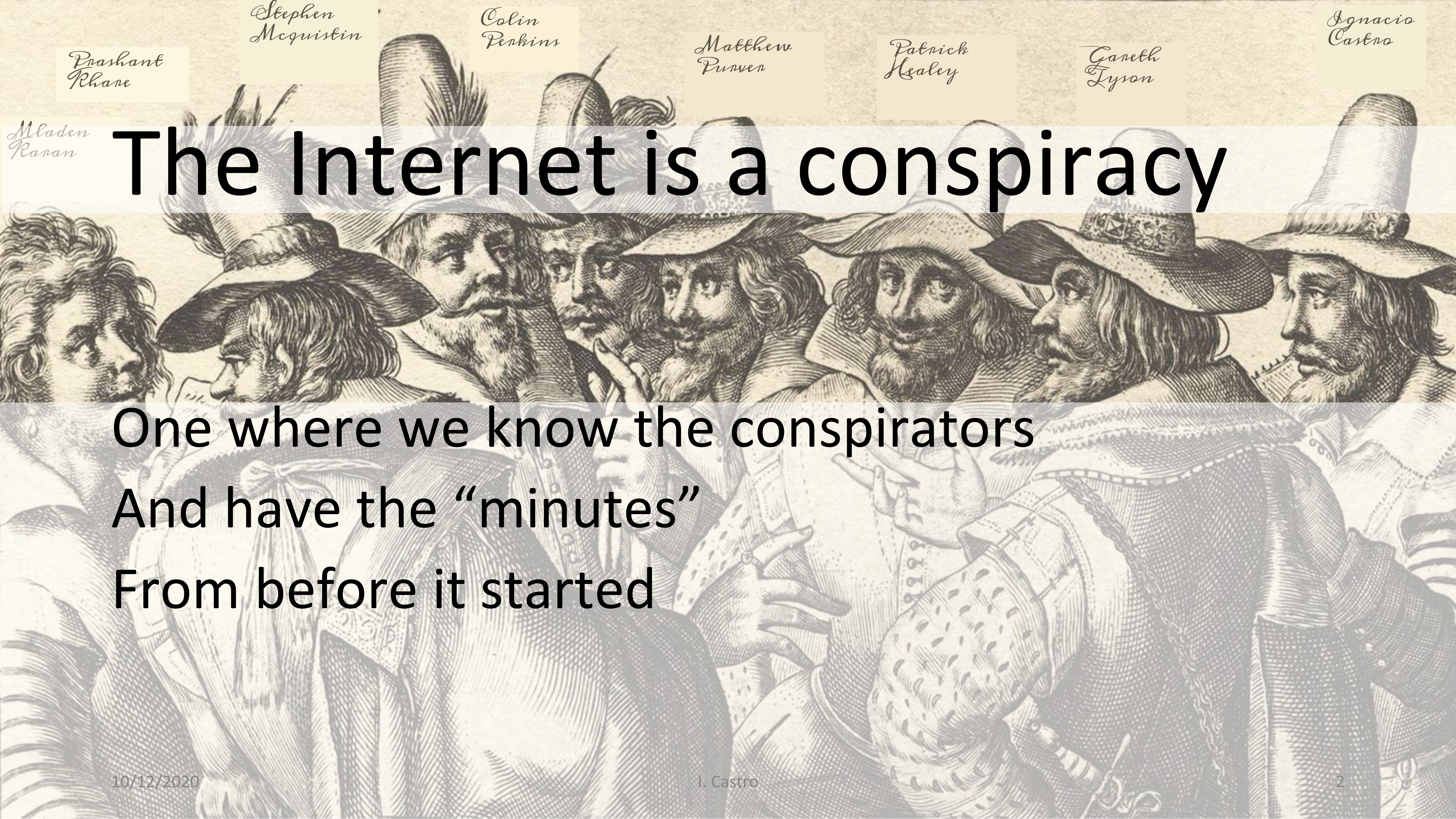
Ignacio Castro +

[Colin Perkins, Gareth Tyson, Matthew Purver, Patrick Healey,

Mladen Karan, Prashant Khare, Stephen McQuistin, Waleed Iqbal]

EPSRC - Streamlining Social Decision Making for Improved Internet Standards (SODESTREAM)

The Internet is a conspiracy

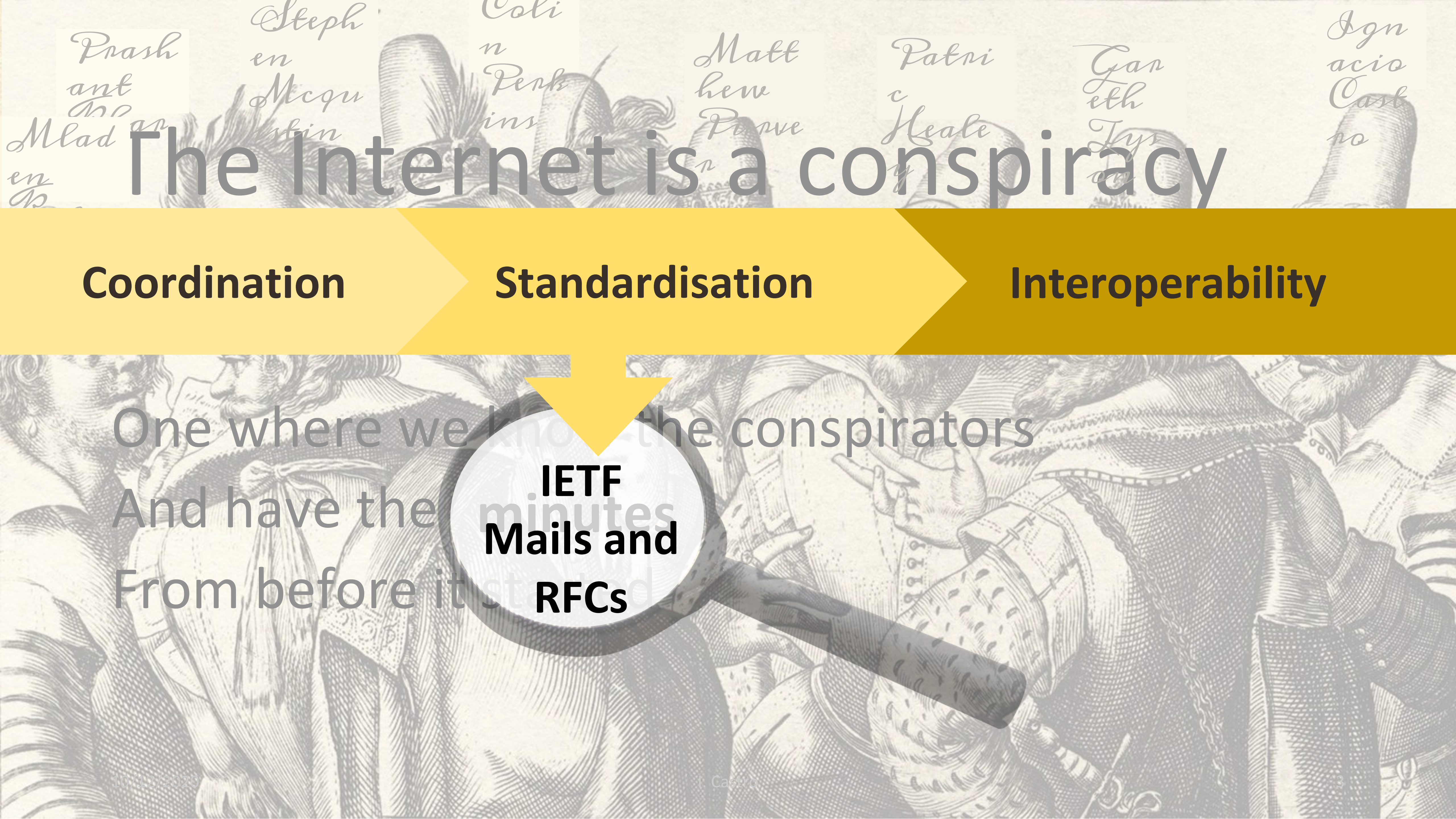

Prashant Khare
Stephen Mcquistin
Colin Perkins
Matthew Purver
Patrick Healey
Gareth Tyson
Ignacio Castro
Mladen Karan

One where we know the conspirators

And have the “minutes”

From before it started

The Internet is a conspiracy


Prashant
Mladen
Stephen Mcquistin
Colin Perkins
Matthew Purver
Patrick Healey
Gareth Tyson
Ignacio Castro

Coordination → **Standardisation** → **Interoperability**

One where we know the conspirators

And have the minutes

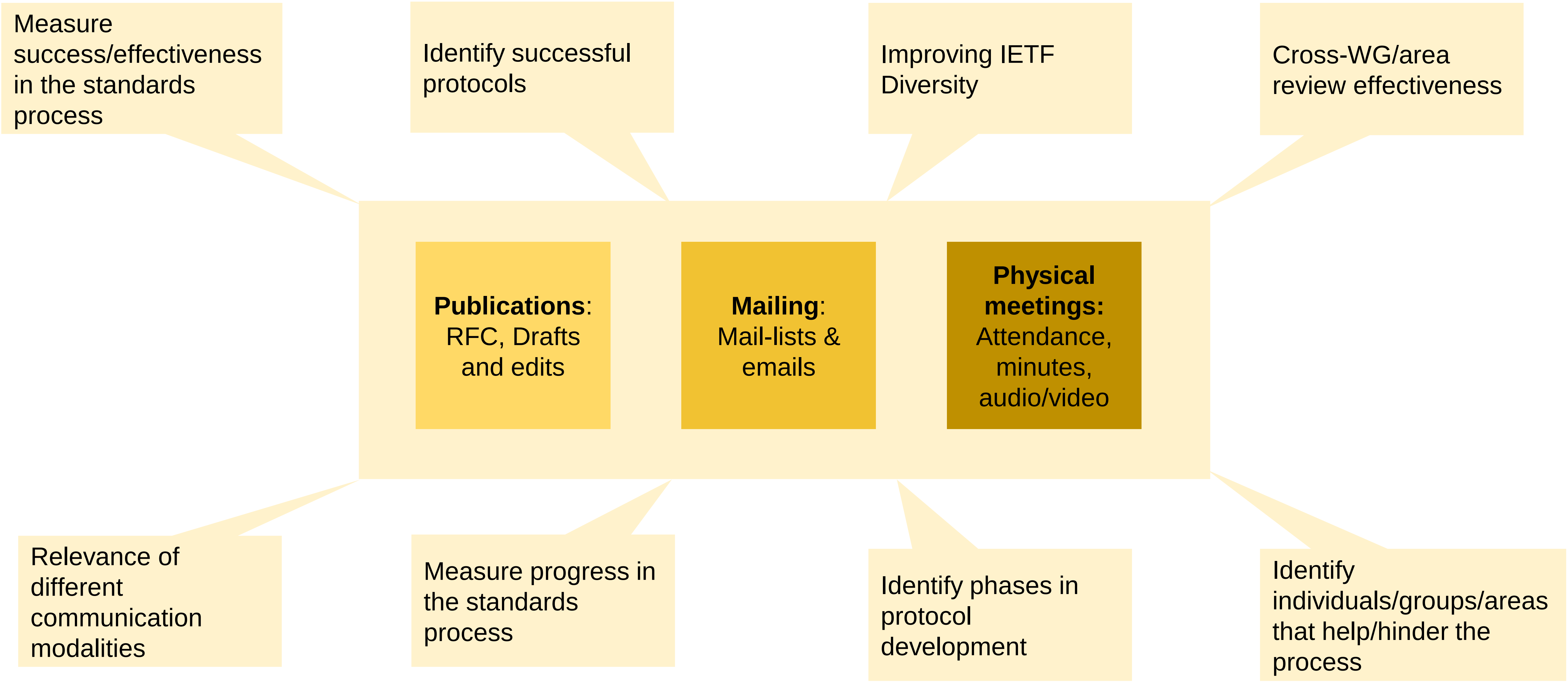
From before it started

IETF Mails and RFCs

Mladen
Karan
Prashant
Khare
Stephen
Mcquistin
Colin
Perkins
Matthew
Purver
Patrick
Healey
Gareth
Tyson
Ignacio
Castro

- How has the Internet evolved?
- How to “coordinate” better?

Data and Objectives



Measure success/effectiveness in the standards process

Identify successful protocols

Improving IETF Diversity

Cross-WG/area review effectiveness

Publications: RFC, Drafts and edits

Mailing: Mail-lists & emails

Physical meetings: Attendance, minutes, audio/video

Relevance of different communication modalities

Measure progress in the standards process

Identify phases in protocol development

Identify individuals/groups/areas that help/hinder the process

Preliminary analysis: trends in RFCs production

Do RFC publications reflect the underlying Internet ecosystem?

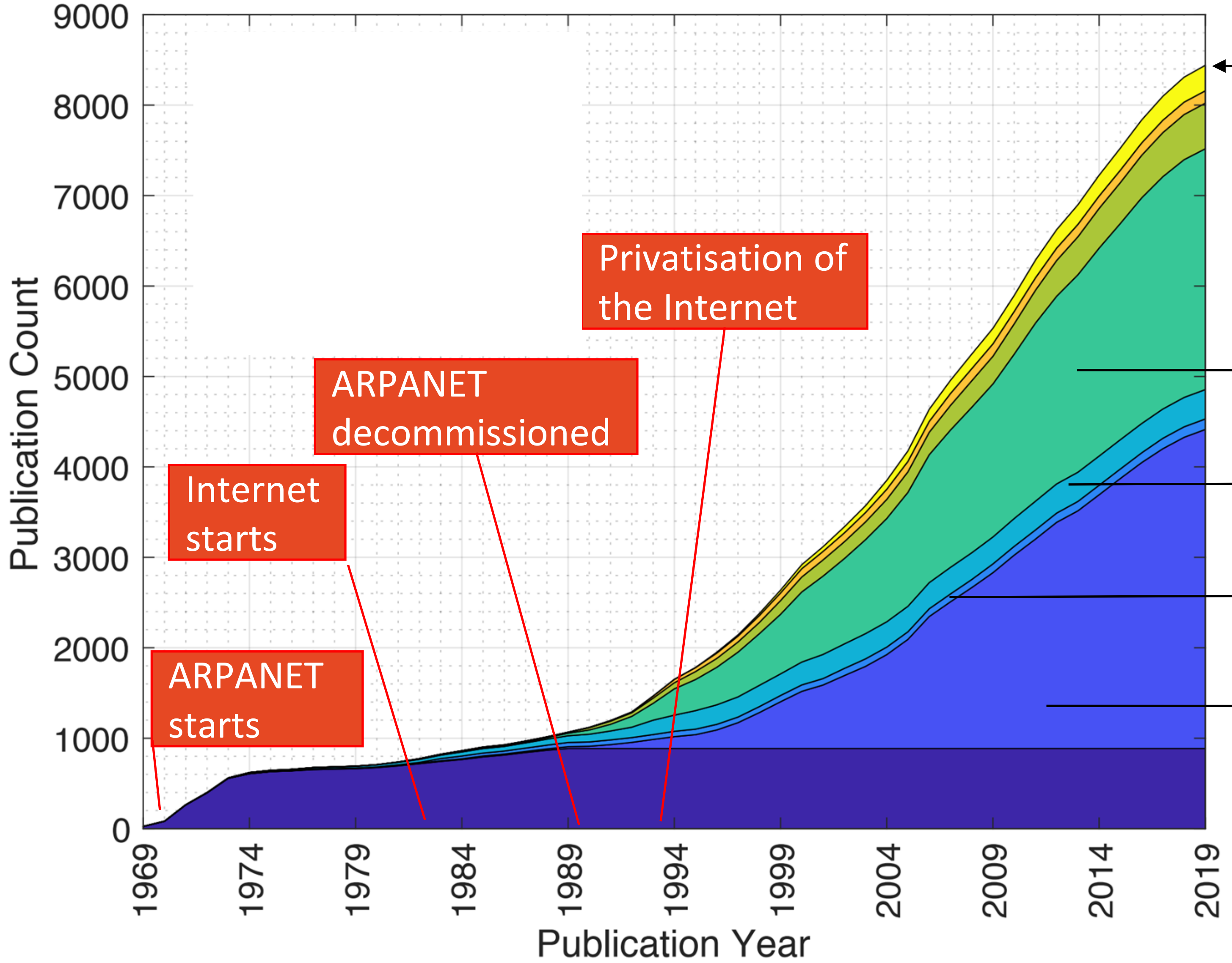
Publications: RFC, Drafts and edits

Mailing: Mail-lists & emails

Physical meetings: Attendance, minutes, audio/video

Is it increasingly harder to publish an RFC?

RFCs reflect the underlying ecosystem



ARPANET starts

Internet starts

ARPANET decommissioned

Privatisation of the Internet

RFC publication growth reduction

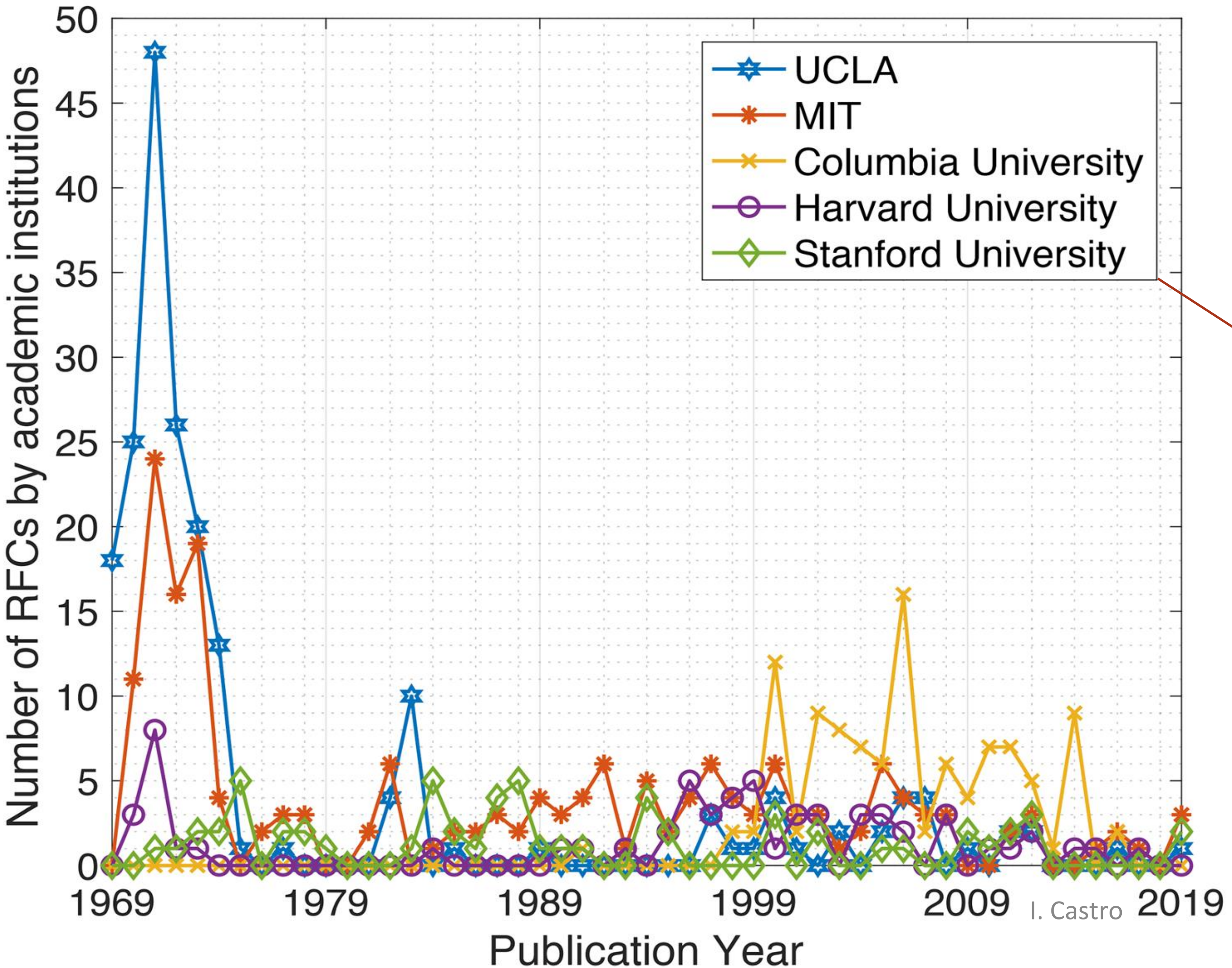
Many informational RFCs

Some RFCs become obsolete

Few are advanced to full standards

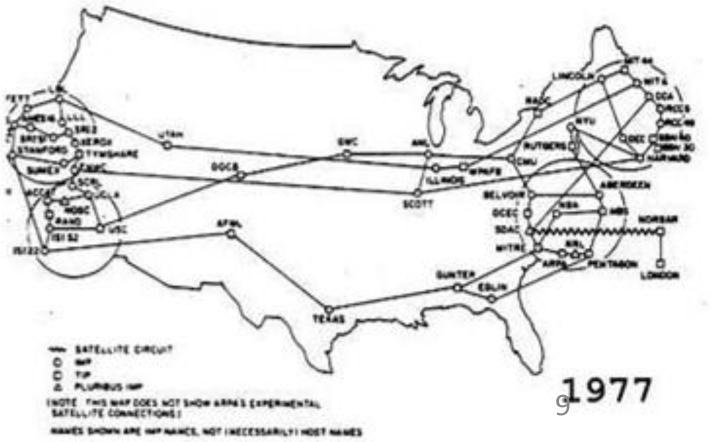
Many standards are proposed

Evolution: from research

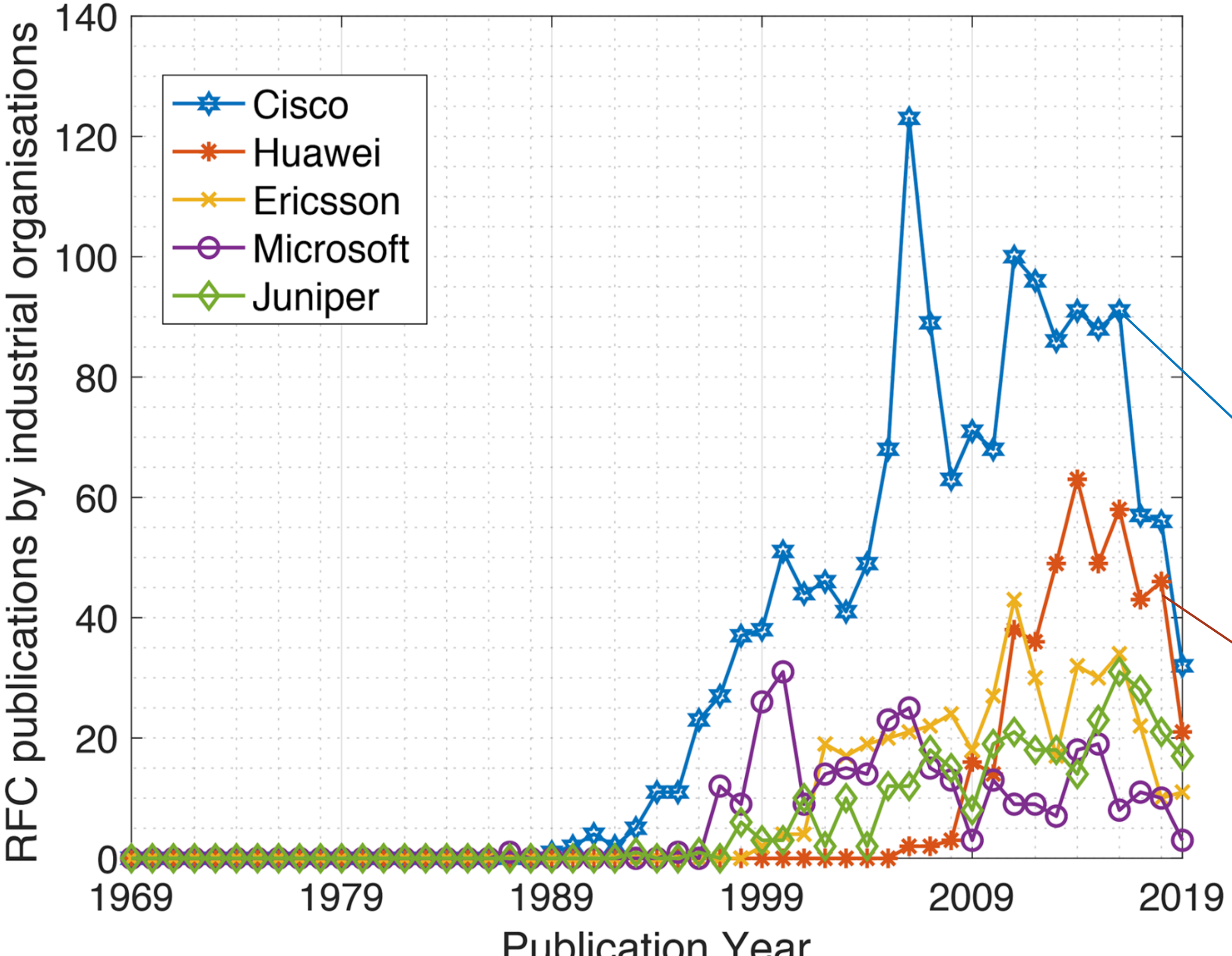


- RFC publication reflects the underlying stakeholders
  - From academic institutes

Nodes in the early ARPANET



Evolution: to industry



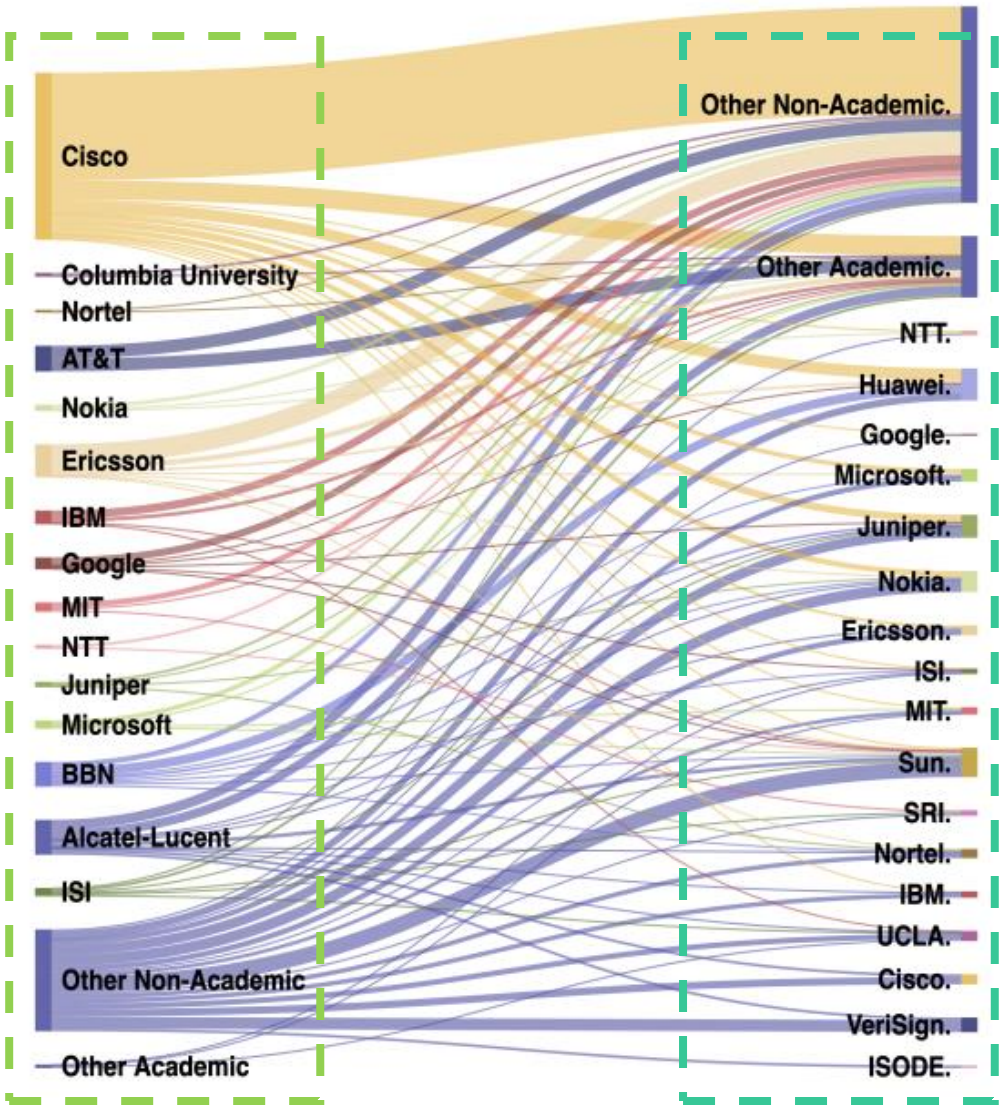
- RFC publication reflects the underlying stakeholders
  - From academic institutes
  - To commercial giants

Rise of Internet giants: CISCO

Rise of new giants: Huawei

Competition for talent (RFC authors)

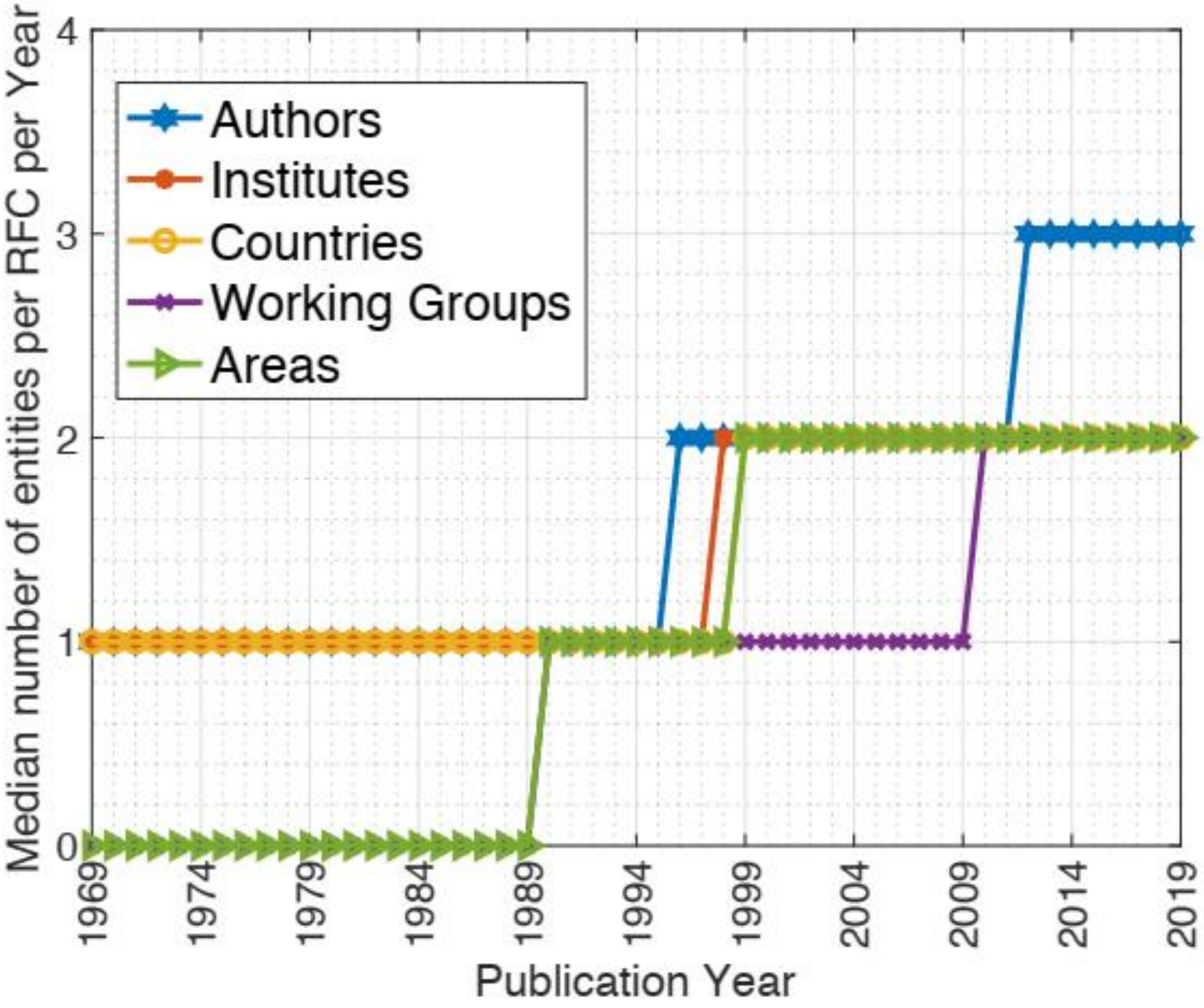
First observed affiliation

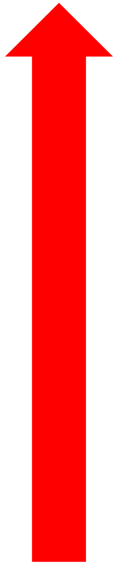


Last observed affiliation

Is publishing an RFC becoming harder?

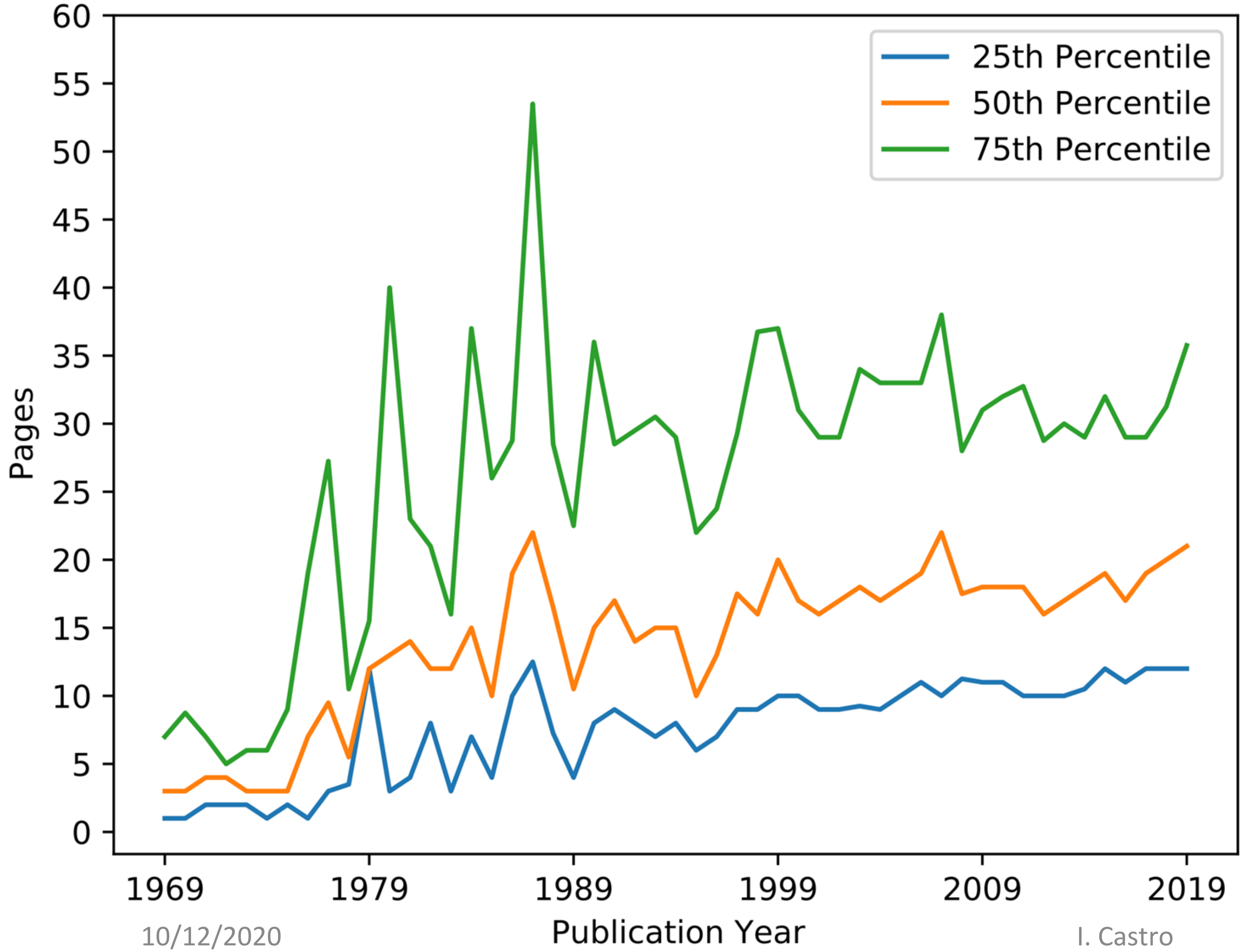
RFCs involve more…

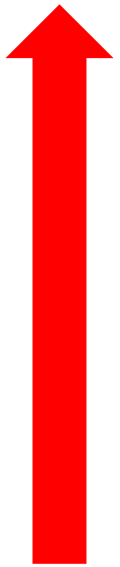




More:
- People
- Institutions
- Countries
- Dependencies (WGs/Areas)

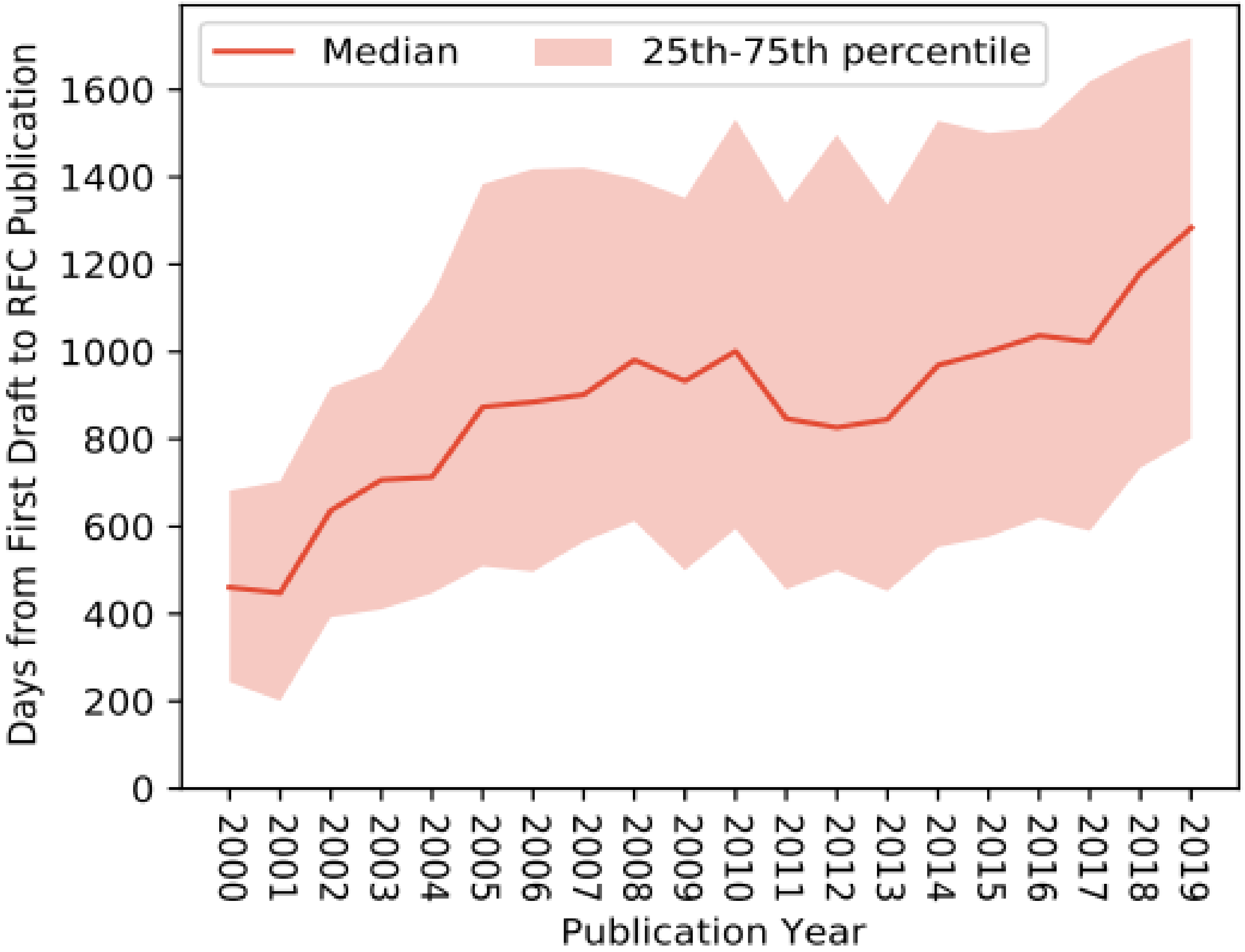
RFCs are longer

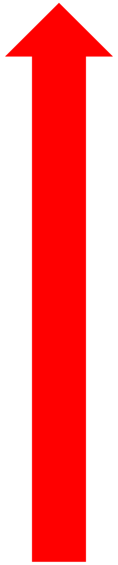




More pages

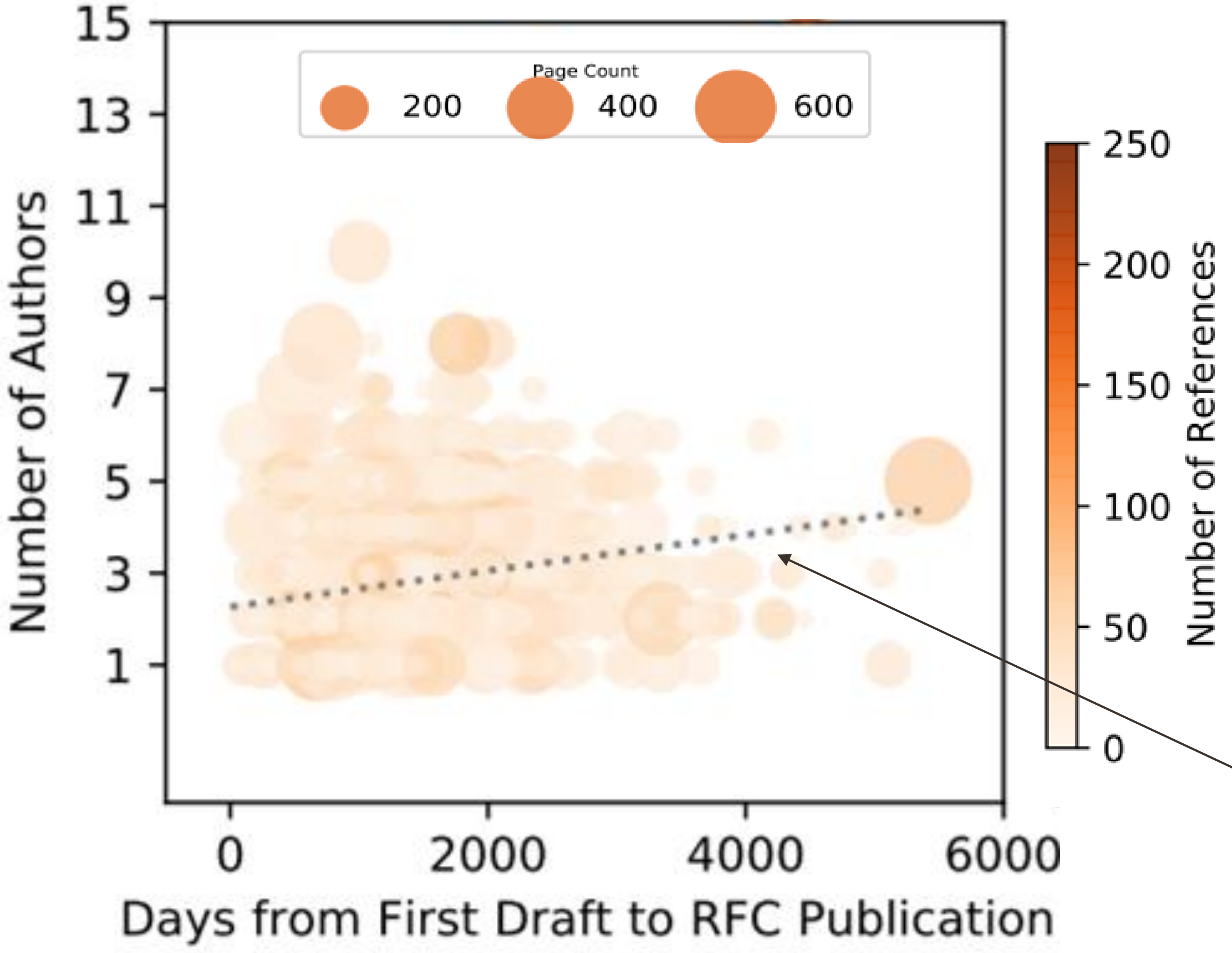
RFCs take longer (to be published)





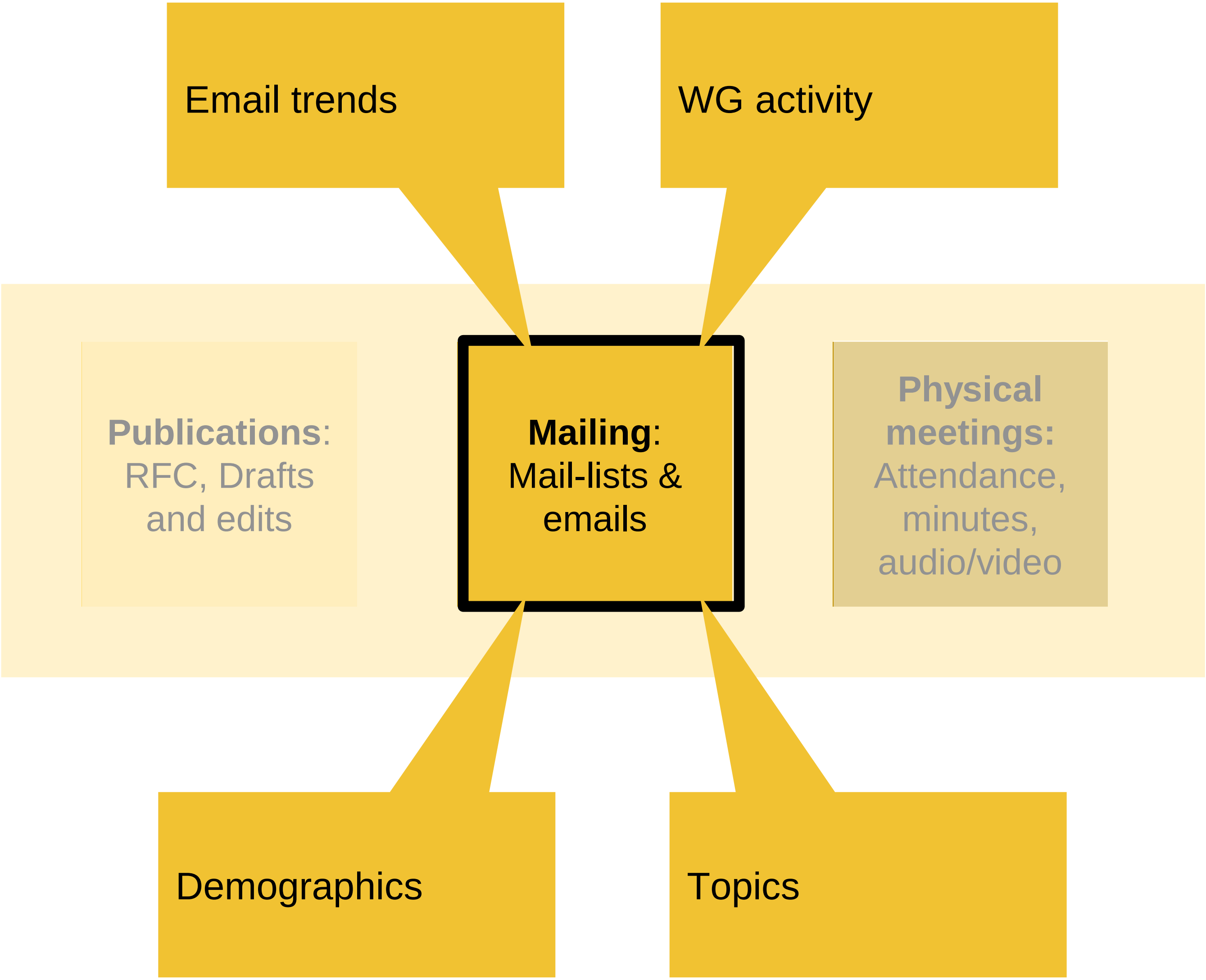
More days from draft to RFC publication

Is it harder to publish an RFC and why?



Documents become longer, include more references, and authors, is that making the publication process longer?

Preliminary analysis: trends in RFCs production



Email trends

WG activity

Publications: RFC, Drafts and edits

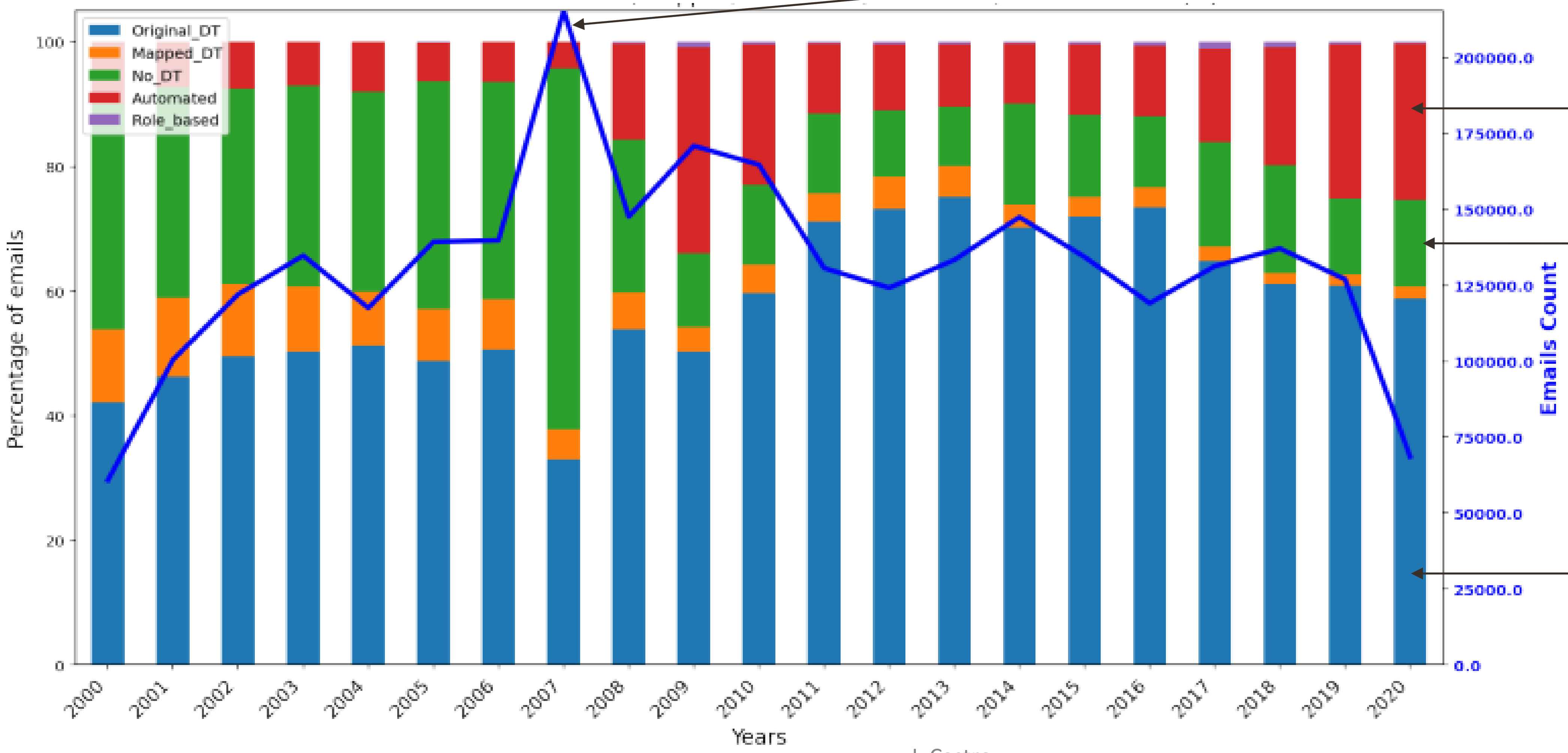
Mailing: Mail-lists & emails

Physical meetings: Attendance, minutes, audio/video

Demographics

Topics

Datatracker users are responsible for most emails



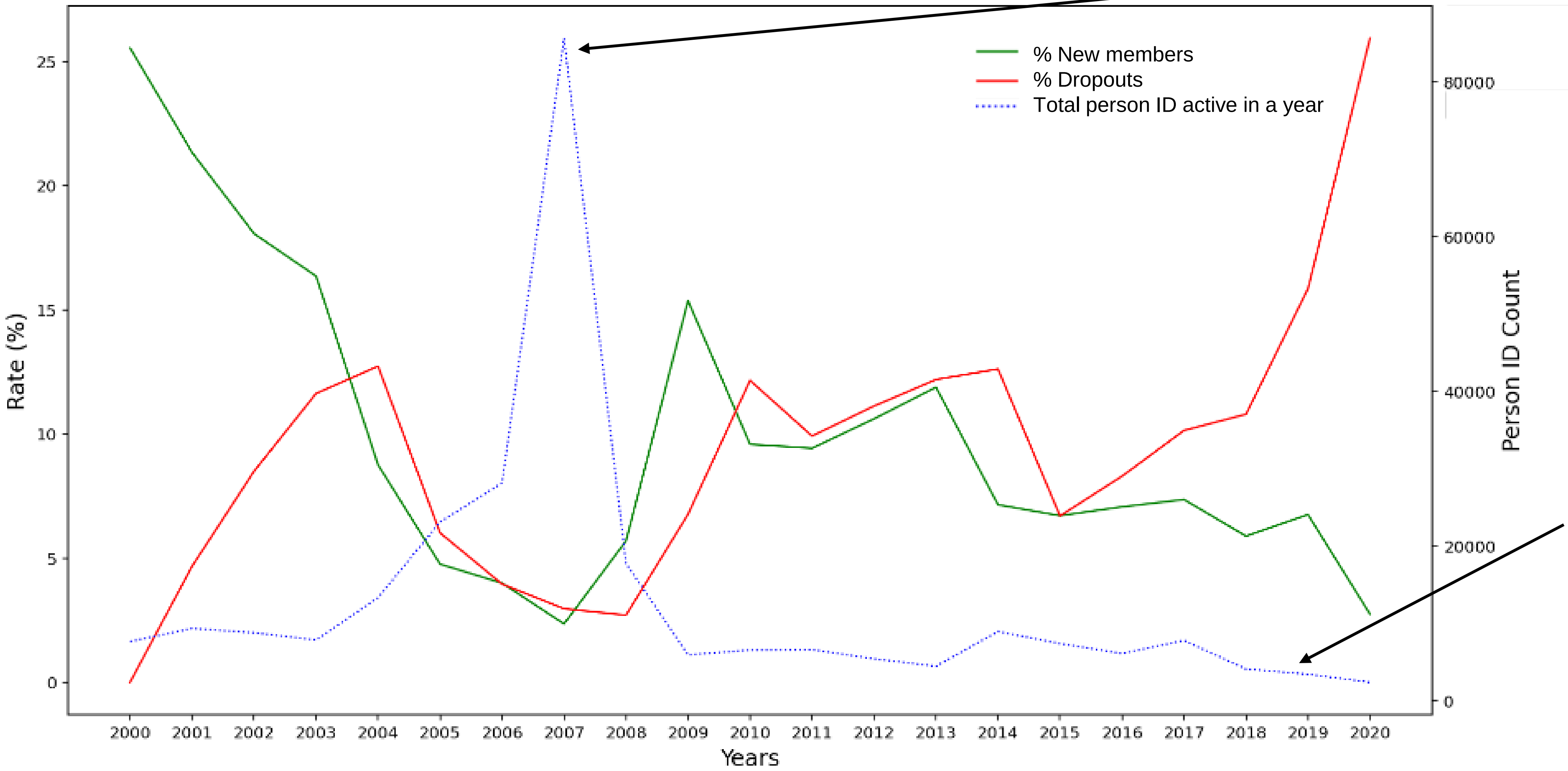
Outburst in 2006?

Growing percentage of automated emails

Small percentage of emails with no ID

Most emails have a datatracker ID

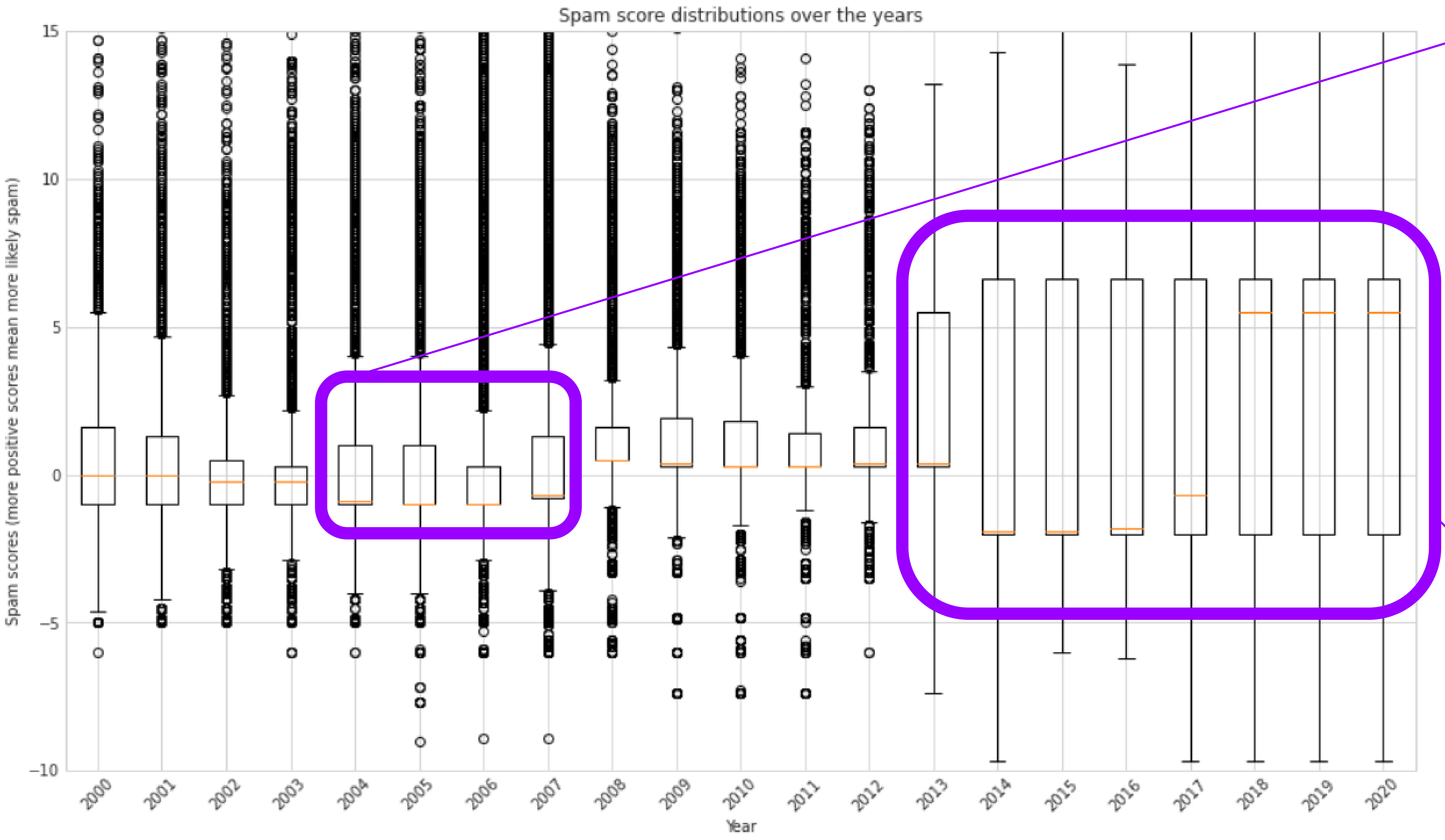
Is the IETF growing?



Outburst in 2006?

Dropping numbers?

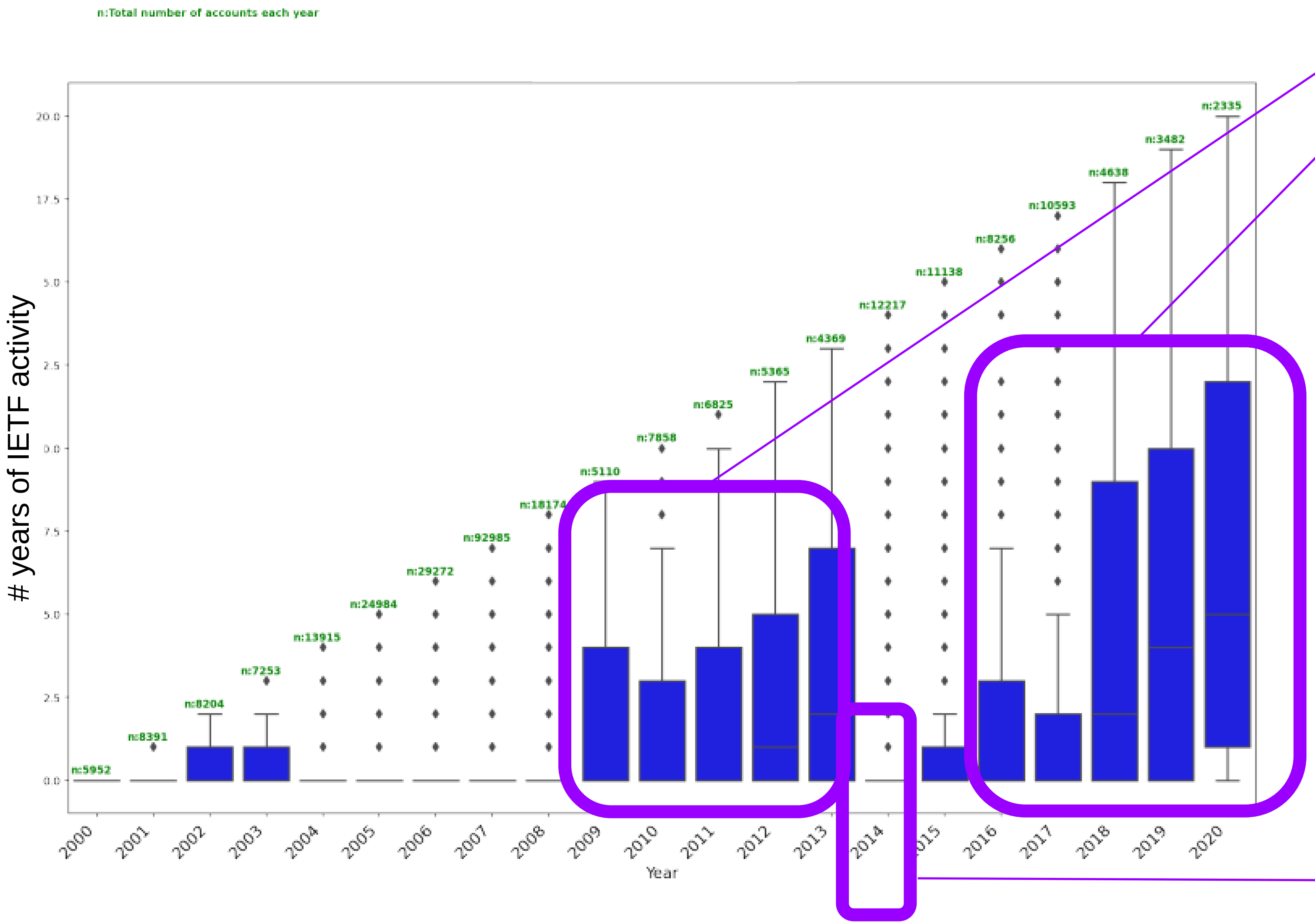
Spam? Unlikely according to SpamAssasin



Lower average

Change of policy in 2013?

Is the IETF aging?

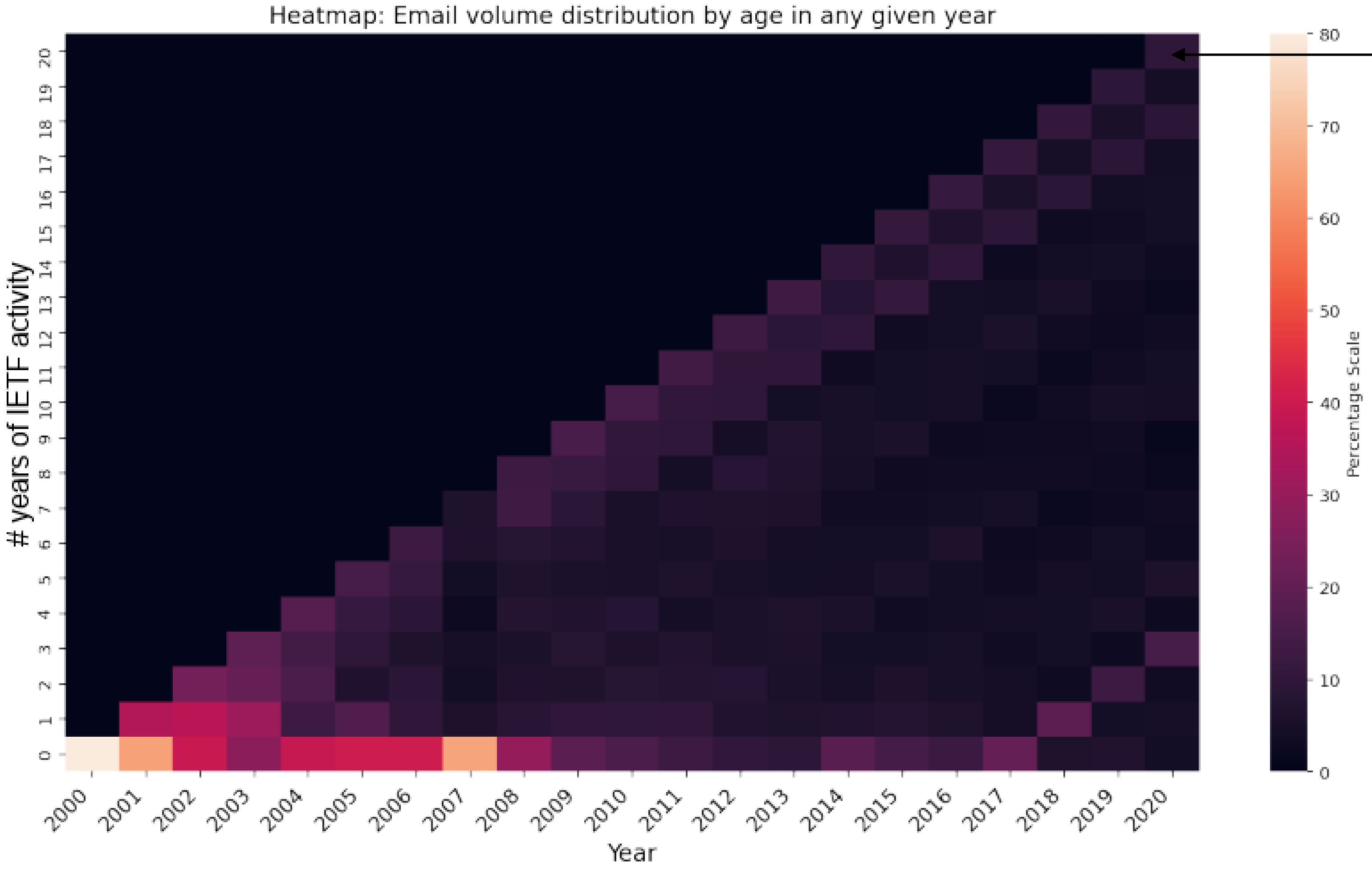


Why this waves?

- Age = IETF-age = # years of email activity of a person
- 2020=0

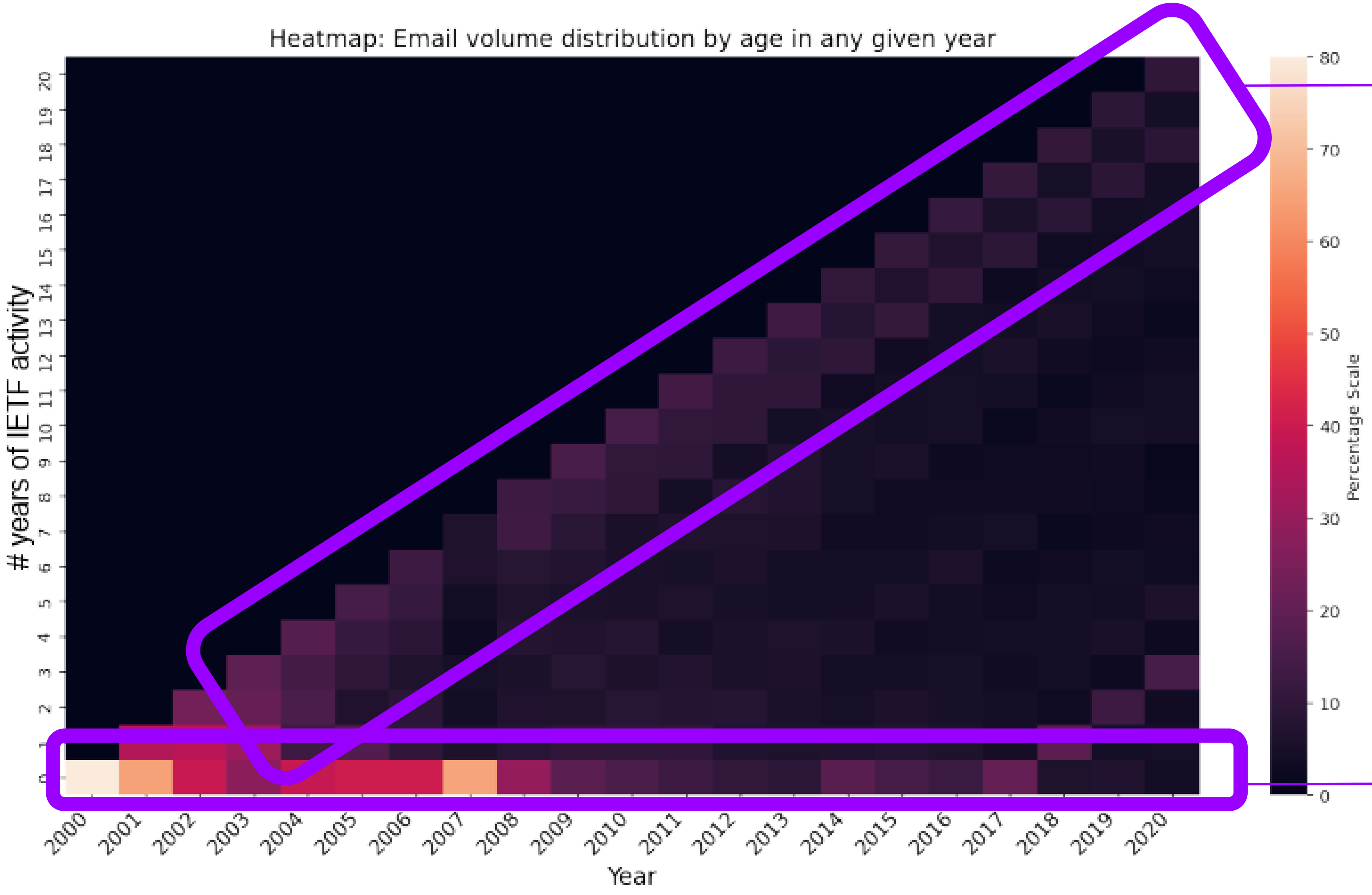
What happened in 2014?

Elders and newbies are most active



% of emails sent in 2020
(100%=all emails sent in 2020)
by IETFers of IETF-age=20
(ie, 20 years of IETF email-activity)

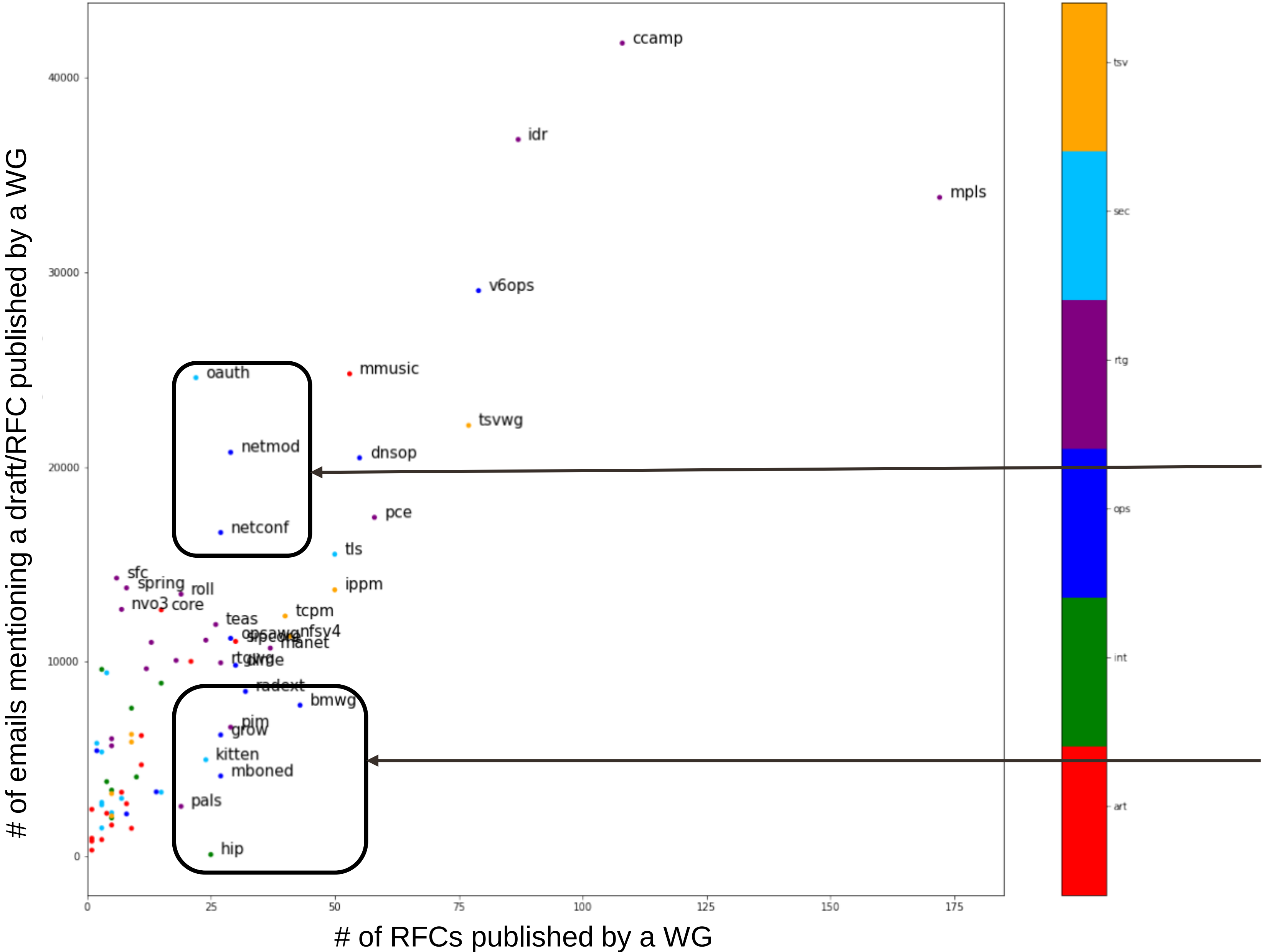
Elders and newbies are most active



Are most seniors dominating the conversation?

Are newcomers discouraged after the initial excitement?

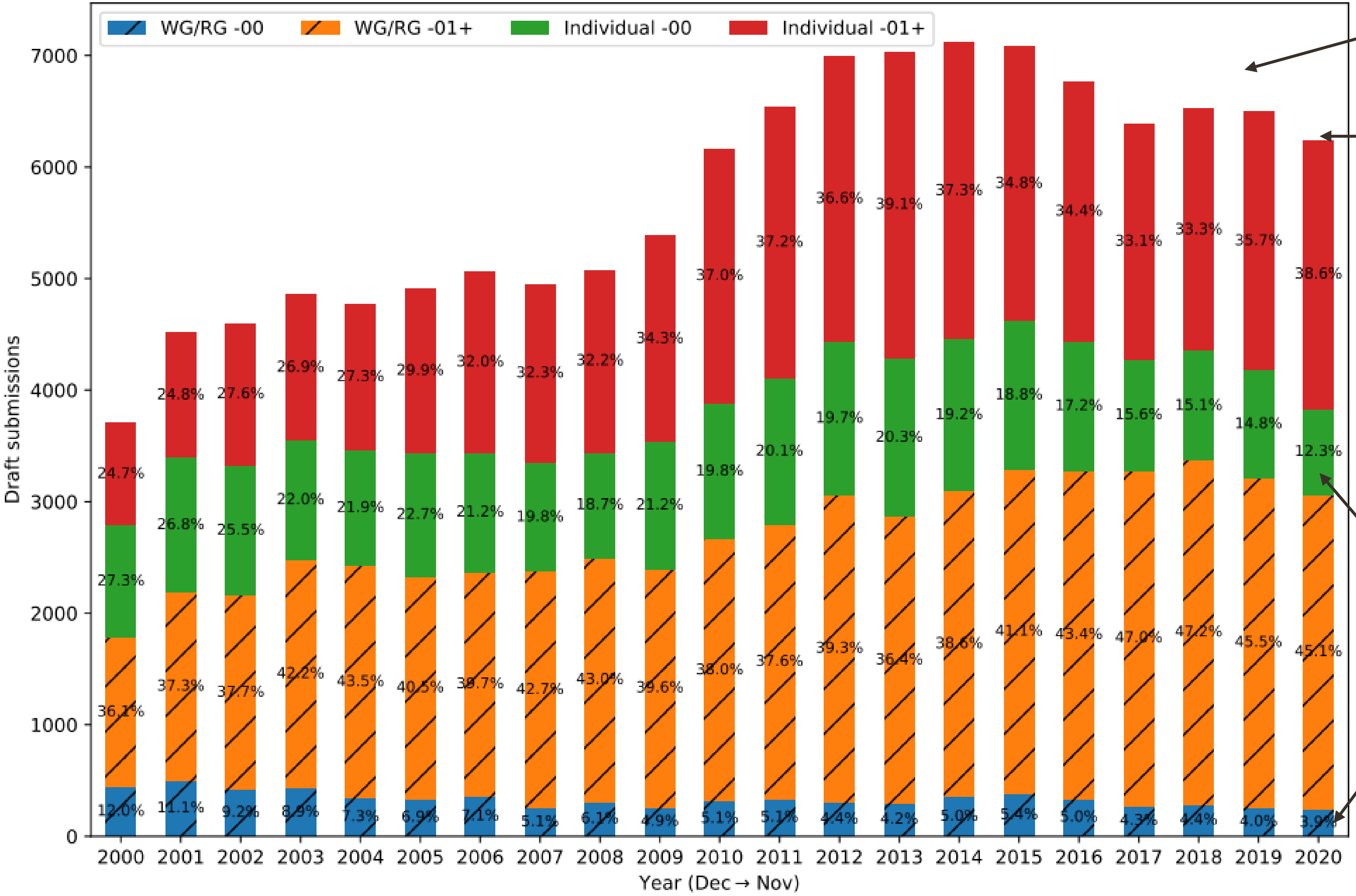
Mailing activity reflect RFC production by WGs



Many emails, few RFCs

Few emails, many RFCs

Recent decrease of draft submissions

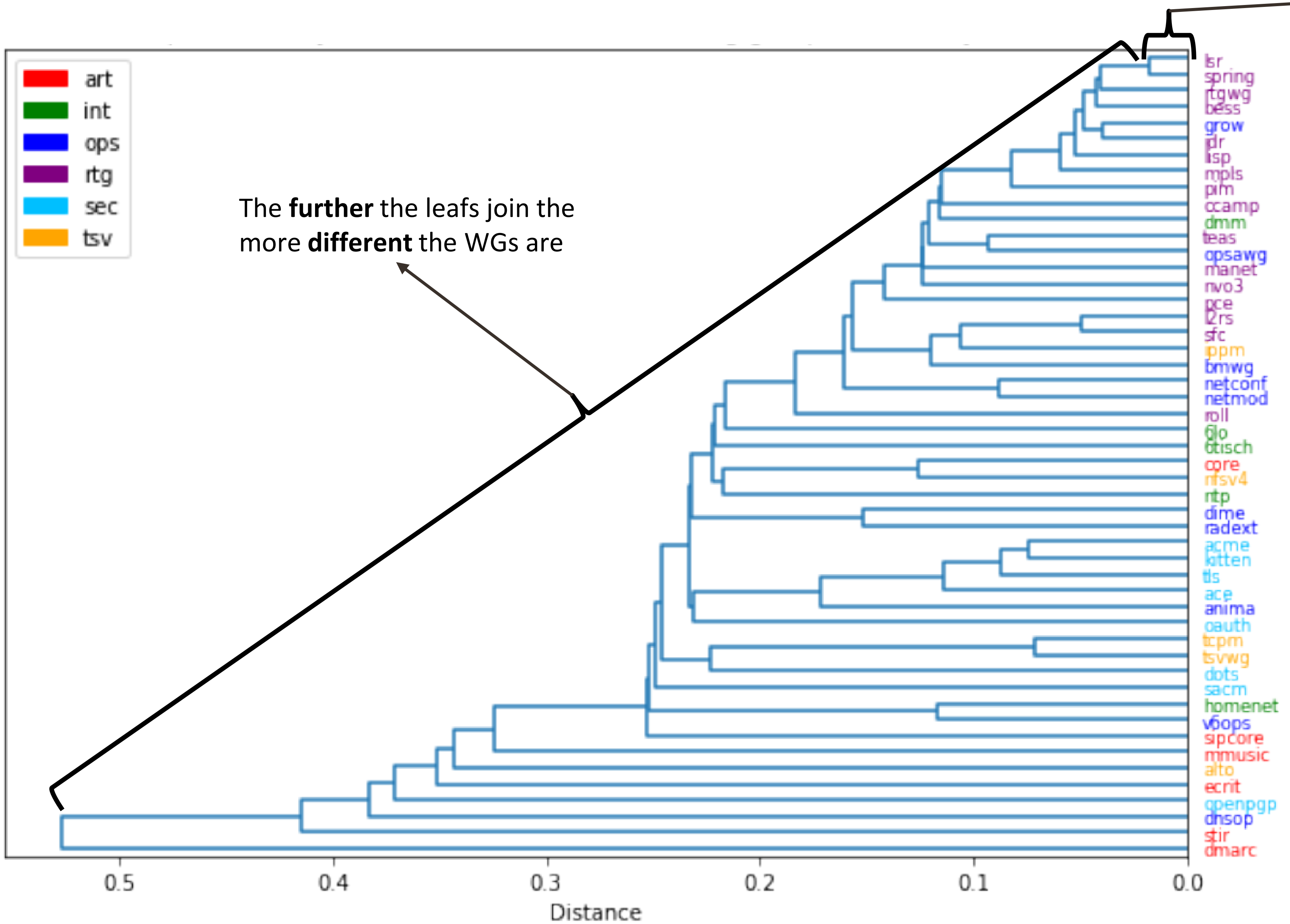


Decreasing number of RFC submissions

Covid-19?

Decreasing # of first draft submissions

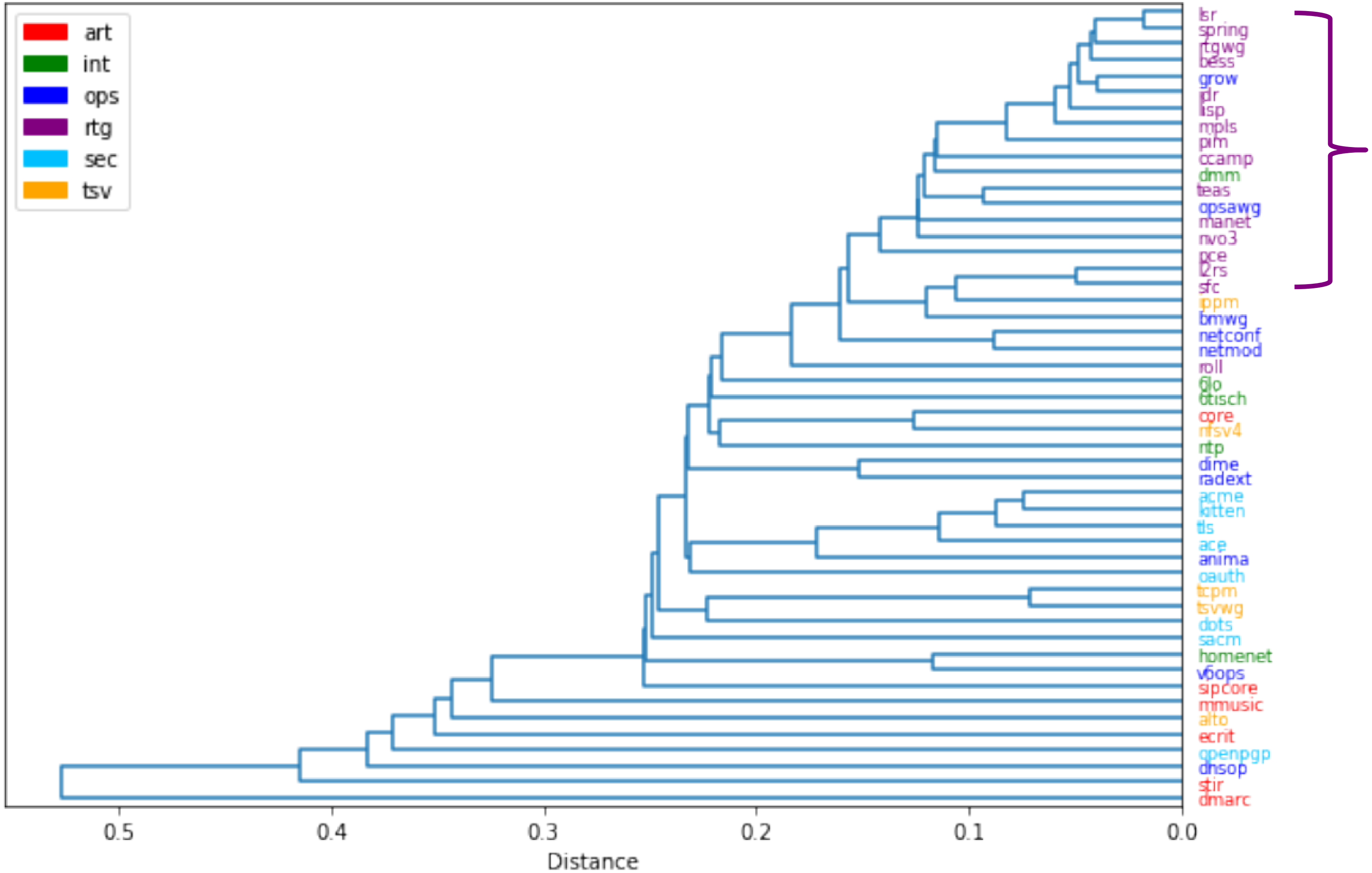
Topic similarity (of the emails) across WGs



The **sooner** the leafs join the more **similar** the WGs are

The **further** the leafs join the more **different** the WGs are

Topic similarity (of the emails) within areas



Open questions

- How can we improve the standards process?
    - Is diversity an issue?
    - What is the relevance of affiliation
    - How to keep participants? An ossified or a dynamic forum?
    - Enough cross area collaboration?
    - What makes a successful WG?
- How does the IETF stay ahead of the game?
    - Covid impact
- Are there other IETF datasets
    - E.g., historic mail-list subscriber)?
- How can these finding be communicated?
    - E.g., what tooling would be useful

What other aspects/problems we should be looking at?

i.castro@qmul.ac.uk