Queen Mary
University of London

# Behind the drafts: IETF
## The evolution of RFCs' publication through their discussions.

Hugo Ramirez
July 2025

**Internet Engineering Task Force**

Develops Standards for Internet
eg: HTTPS, DNS or TCP/IP

Publishes Request for comments **(RFC)**

Open participation

Consensus driven

No mandatory implementation for RFCs

# Declining number of RFCs



Honey in Decline

**Declining** number of **RFCs** per year

We are **looking** into the **conversations** to analyze this trend.

Conversations around the drafts (RFC) reflect the development process.

Queen Mary
University of London

# Conversation around the same draft

The IETF mainly works in mailing lists (ML)

**Conversation:**
Groups of emails that discuss the same draft

Characteristics (Remove outliers and complete conversations)
- At least 30 messages
- Conversations after 2000 before 2023

Total 6,971 Conversations.
4,621 ends in publication of RFC
2,350 do not end in publication

# Conversations and RFC Publication how they relate?



**Hypothesis:**
We can extract **structural parts** of the **conversation** to find interaction patterns.

And we can **use** them to **predict** potential conversations that can result in the publication of an RFC

We extracted **timestamps** from the messages

**Time** models the pace and interaction
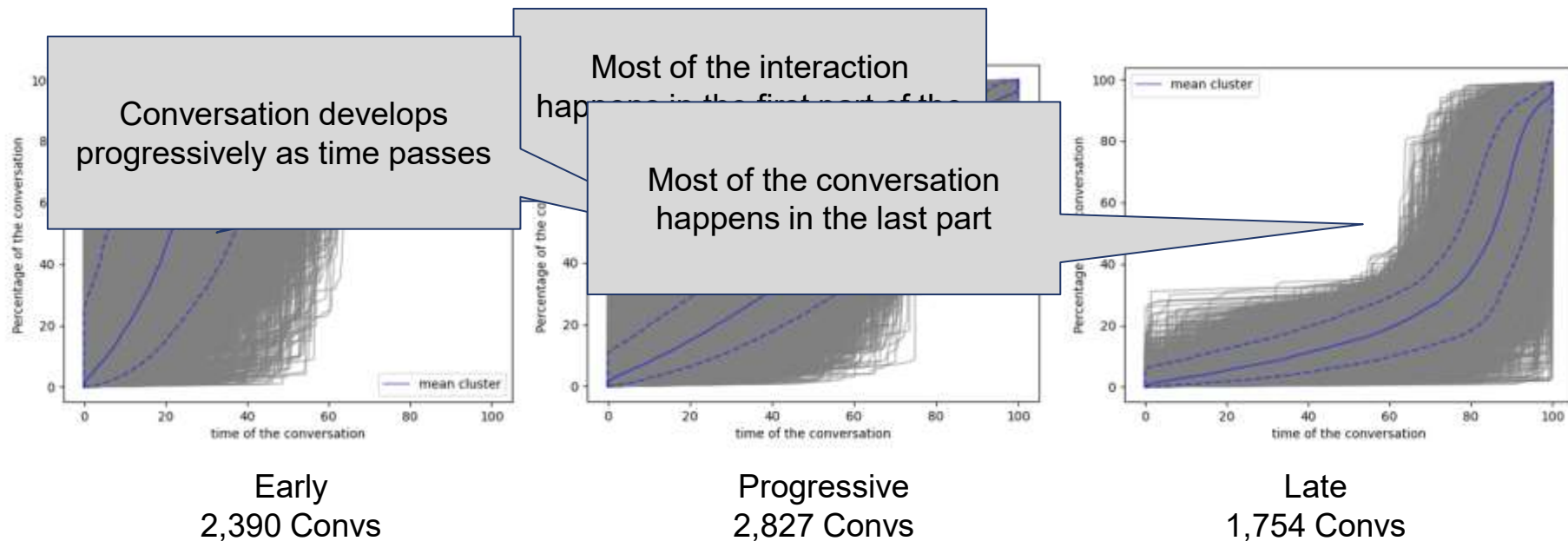
# Time based interaction patterns



**Time** based patterns by grouping similar conversations
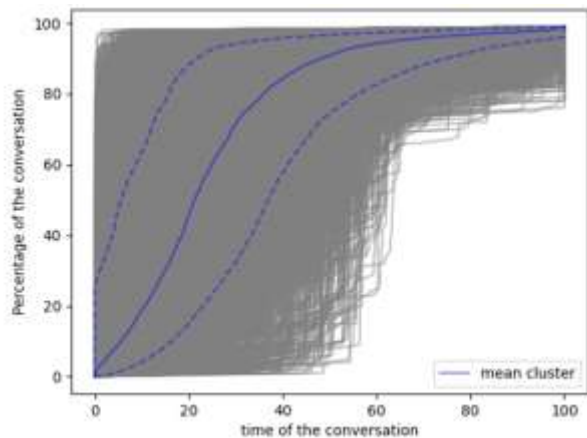
We group them (Hierarchical Clustering)

N-clusters: Best Silhouette score
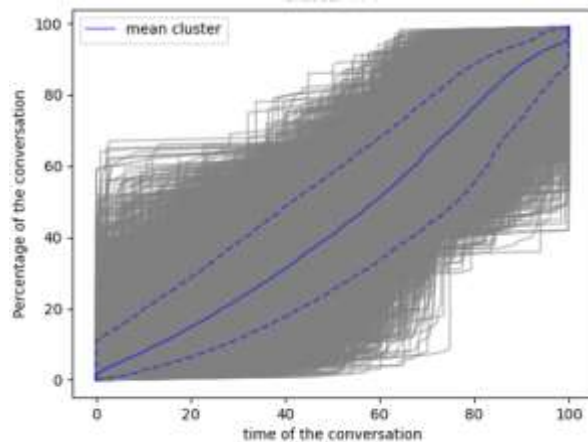3 Clusters

# Three clusters of conversations



Conversation develops progressively as time passes

Most of the interaction happens in the first part of the

Most of the conversation happens in the last part

Early
2,390 Convs

Progressive
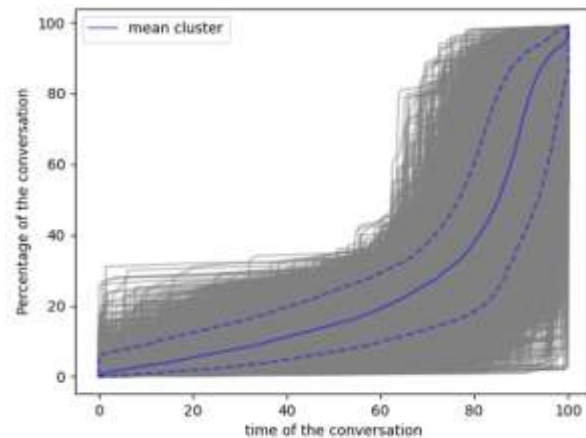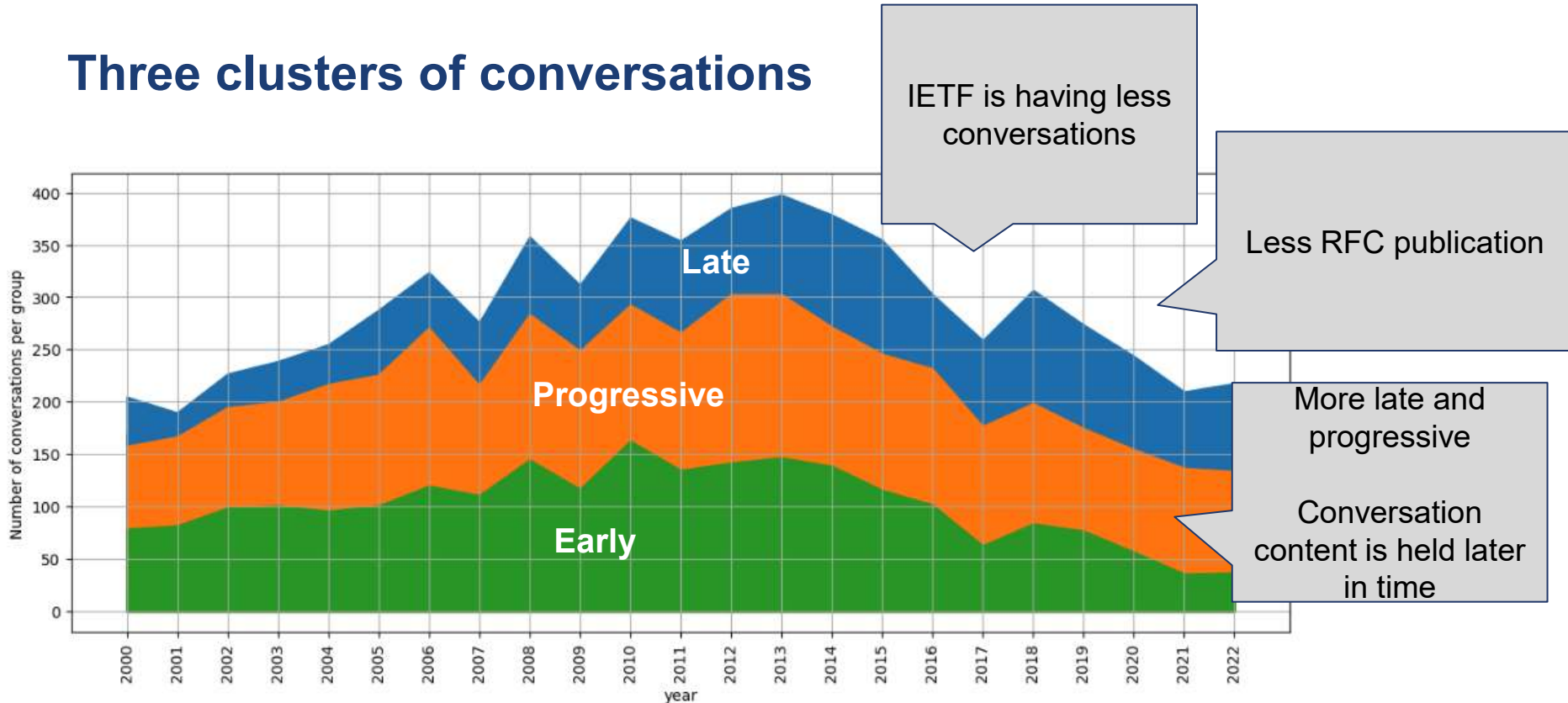2,827 Convs

Late
1,754 Convs

# Three clusters of conversations



Early
2,390 Convs

Progressive
2,827 Convs

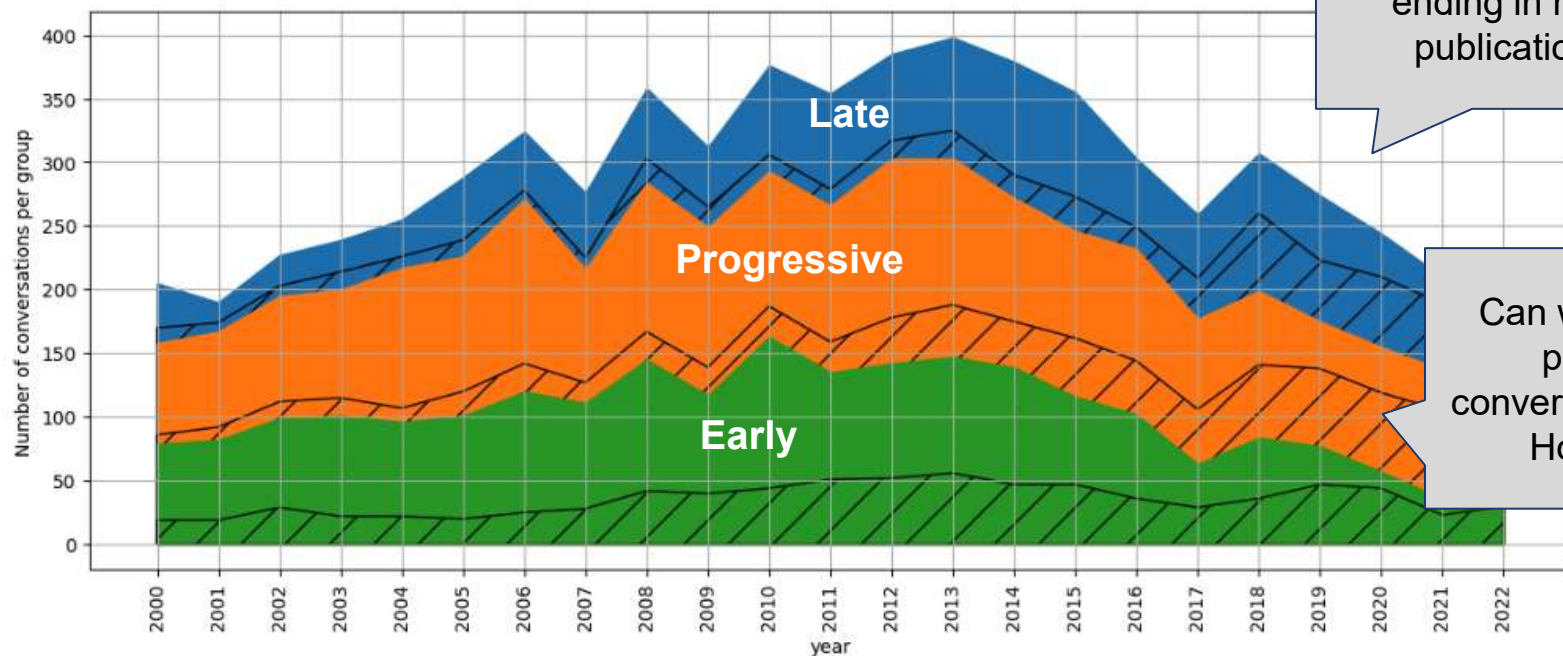Late
1,754 Convs

# Three clusters of conversations
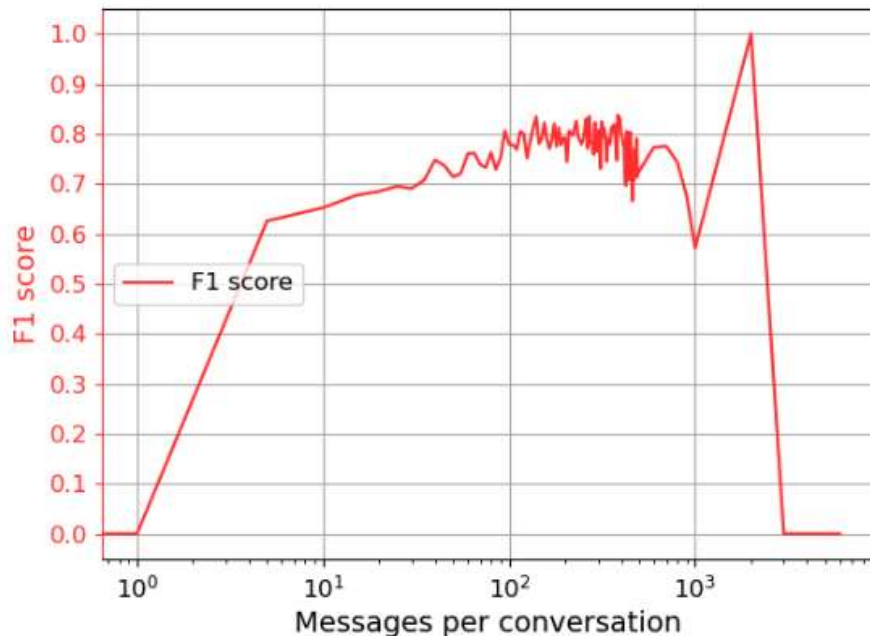
# Three clusters of conversations

# Can we predict the outcome of a conversation?
## Using time based patterns



We used a random forest model for prediction

Not all conversations are suitable
for prediction due to length.

More content (messages) used for
predicting the better F1 score.

Around 100 messages we get
decent F1 score of 0.8

# In summary

We found time based patterns that describe conversation evolution in the IETF.

IETF is having less conversations.

More of the content of the conversation is being held in later stages.

We can pinpoint potential conversations but at least 100 messages are needed

# Thanks for listening



Yes it's AI, some bees have no arms 😅



h.e.ramirez-centeno@qmul.ac.uk