Publications, process, and decision-making

AID 2021

Michael Welzl & Ignacio Castro

Towards better RFCs

1. Position Paper (M. McFadden)
   a. **Goal:** improve "security considerations", based on an analysis of RFCs
   b. **Basis:** prior work done by Matt Lewis and Mark McFadden; want to present their research and insights, welcome community engagement and discussion
      i. https://github.com/nccgroup/RFC-Security-Research
      ii. draft-mcfadden-smart-rfc3552-textual-research-00.txt

Trends and success in RFC production

1. The Impact of Continuity (C. Griwodz, S. Islam, M. Welzl)
   a. **Goal:** better understand the path from I-D to RFC: why do some fail? Lack of persistence or active resistance (i.e., there was discussion)? In case of resistance: just asking for more, or counter-arguments? Could also look at seniority of proponents: do they have more success? Do they experience less heated debates?
   b. **Basis:** Perhaps Apache SOLR database?
   c. **Specific idea:** Some queries on the database, sentiment analysis
2. Characterising the IETF through the lens of RFC deployment (I. Castro et al)
   a. **Goal:** trends in RFC production, factors of RFC success
   b. **Data:** Datatracker, IETF-emails (https://sodestream.github.io/)
   c. **Code:** https://github.com/glasgow-ipl/ietfdata
   d. **Plans:** combine NLP with social network analysis, long-term future plan (hiring soon)
3. Project in time series analysis: e-mailing lists (A. Andersdotter)
   a. Time series analysis of the appearances of "the" (for normalisation), "security" and "privacy"
   b. **Data:** "bigbang" email archives

Tools for email analysis

1. The Challenges of Cross-Document Coreference Resolution in Email (X. Li, S. Magliacane, P. Groth)
   a. CR is an NLP task
   b. Doing CR on email, with multiple documents, is quite different from doing it on singular documents
   c. The paper shows progress on this
   d. **Data:** Enron email corpus; method not applied to IETF (yet?)
2. Research Proposal (E. Yachmeneva)
   a. **General goal:** determine level of coordination amongst stakeholders in standards development for IoT industry, and how stakeholders may use this coordination to exercise influence by advancing certain interests
   b. **Basis:** Previous work on IoT standardisation: analysis of the level of policy coordination at the oneM2M group
      i. network analysis based on data analysis and interviews
      ii. Found, e.g.: alignment (and its development over time) of stakeholders, forming of "camps" of companies
   c. **Idea:** extend this prior work to the IETF
      i. Plan to use IoT group meeting minutes + attendee lists, mail archives

Understanding RFCs

1. Understanding Internet Protocol Design Decisions (M. Welzl, C. Griwodz, S. Islam)
   a. **Goal:** better networks education, going from "what" to "why" based on rationale for design decisions
   b. **Basis:** Apache SOLR database from CCR paper which used it to analyze IETF collaboration over time.
   c. **Idea:** apply NLP, long-term future plan (currently hiring)
2. RFCs Change (P. Hoffman)
   a. **Goal:** Determine if an RFC will be updated or obsoleted by I-Ds?
   b. **Basis:** IETF CLI tool: https://github.com/paulehoffman/ietf-cli
   c. **Idea:** Proposing to add more information: E.g., errata, mail discussions about interpretations of RFC text, annotations. ICANN is in the early stages of specifying and possibly creating such a tool.

Some questions...

- Data:
  - Why just emails?
    - RFCs, drafts and versions of
    - Meeting minutes
    - Git
  - Why still blue sheets?
    - Standard format for meeting minutes?

- Who benefits from these works?
  - What tools does the IETF/IETF researchers need?
  - What are the research questions?

- What are the challenges?
  - Deployment?
  - Data/labelled data?

What are the questions?

- Success vs failure:
  - What is a successful/good RFC?
  - What makes the process of RFC making successful/better?
- Decision making
  - Bottlenecks?
  - Who supports what?
  - When is a decision made?
  - What is the decision?

How to extract the information?

Where are the answers?

- E-Mails
- RFCs (drafts)
- Meeting minutes
- Meeting audio/video
- Git
- External data:
  - Surveys
  - Manual labelling
  - Other?