Analysing Internet Standards Data

draft-perkins-analysing-sdo-data-01

Colin Perkins, Ignacio Castro, Ryo Yanagida, and Stephen McQuistin

Colin Perkins | https://csperkins.org/ | Copyright © 2026 University of Glasgow | This work is licensed under the Creative Commons Attribution-NoDerivatives 4.0 International License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nd/4.0/ or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

Analysing Internet Standards Data

RASPRG
Internet-Draft
Intended status: Informational
Expires: 4 January 2027

C. Perkins
University of Glasgow
I. Castro
Queen Mary University of London
R. Yanagida
S. McQuistin
University of St Andrews
3 July 2026

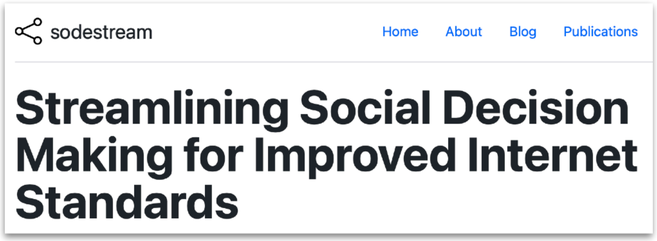
Analysing Internet Standards Data
draft-perkins-analysing-sdo-data-01

Abstract

This document outlines some issues to consider when studying data relating to the Internet standards development ecosystem. It identifies observable components of standards development processes, proposes a taxonomy of possible measurements, and highlights methodological, interpretive, and ethical considerations. It is intended to support a range of uses, including monitoring standards development organisations (SDOs), evaluating the evolution of technical work, understanding technology deployment, and informing community, leadership, and governance discussions.

This document is submitted for consideration by the Research and Analysis of Standard-Setting Processes Research Group (RASPRG) in the IRTF. It is not an IETF product and is not a standard.

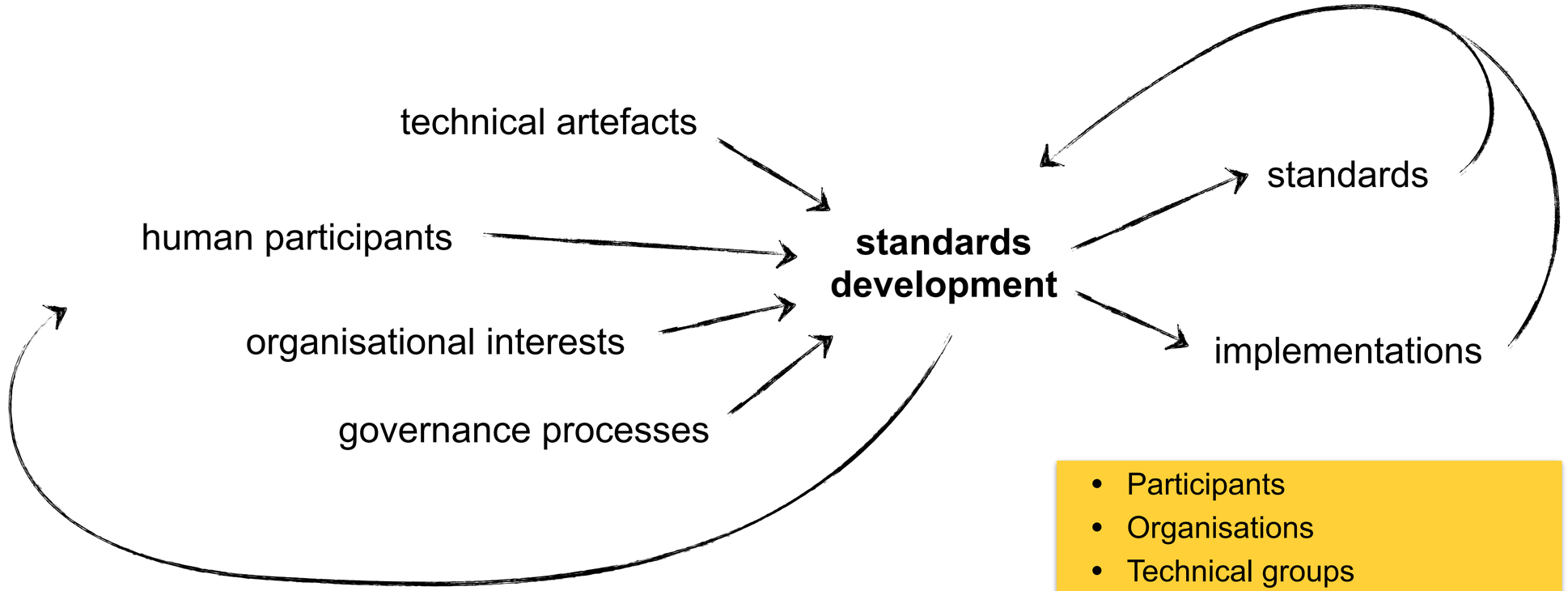
- Submitted to help structure our thinking around analysis of Internet standards data
  - Standards Development as a Socio-Technical System
  - Analysing the IETF
  - Analysing Other SDOs
  - Data Processing
  - Ethics and Data Protection
  - Recommendations



https://sodestream.github.io/

Colin Perkins | https://csperkins.org/ | Copyright © 2026 University of Glasgow

Standards as a Socio-Technical System (1/2)



- Participants
- Organisations
- Technical groups
- Artefacts
- Collaboration infrastructure
- Governance structures
- Standards and implementations

Colin Perkins | https://csperkins.org/ | Copyright © 2026 University of Glasgow

Standards as a Socio-Technical System (2/2)

- Observable metrics provide useful evidence, but each captures only part of the standards process
- Critical aspects of standards development are hard to observe directly
  - Culture, influence, and agenda setting
  - Informal coordination and negotiation
  - The exercise of power and authority
- Metrics, artefacts, and other data differ in accuracy, representativeness, and relevance

Colin Perkins | https://csperkins.org/ | Copyright © 2026 University of Glasgow

Analysing the IETF









- IETF makes a significant amount of data publicly available
- 5.3 million datatracker records, 3 million emails → 40GB
- RFC and draft text, presentations, session recordings, …
- Combine with GitHub issues, etc.

Colin Perkins | https://csperkins.org/ | Copyright © 2026 University of Glasgow

Analysing Internet Governance, Broadly

















- Other Internet SDOs and governance organisations vary considerably in the data they make publicly available
- Goals and modes of operation also vary

- Complex to integrate and interpret data across different organisations

Colin Perkins | https://csperkins.org/ | Copyright © 2026 University of Glasgow

Data Processing

- Entity resolution for personal and organisational names
  - Multiple names per entity, name changes, variant spellings
- Tracking participant affiliations
  - Not well recorded in the IETF Datatracker
  - Complicated by mergers and acquisitions, consulting relationships, and opaque funding sources
- Tracking non-linear document lifecycles
- Reconstructing leadership role history
- Handling malformed and archaic email header fields
- Handling incomplete historic data

Colin Perkins | https://csperkins.org/ | Copyright © 2026 University of Glasgow

Ethics and Data Protection

- IETF privacy statement and other policies apply
- Personal data is sensitive and subject to legal restrictions on usage and processing in many jurisdictions
- IETF data may be public, but implications of the data are not always well-known – inappropriate to publish data about specific individuals without their explicit consent
- Consult Research Ethics Committee or Institutional Review Board
- **Effective functioning of the Internet standards process directly affects critical infrastructure,** take care not to disrupt the process

Colin Perkins | https://csperkins.org/ | Copyright © 2026 University of Glasgow

Recommendations

- Recommendations to IETF:
  - Preserving centralised and stable data access
  - Data quality and consistency
  - Historical data and backfilling
  - Provenance of derived data
  - Error reporting and correction
  - Separation of primary and inferred data
  - Impact of process and tooling changes

Input sought on these recommendations and what is missing or incorrect

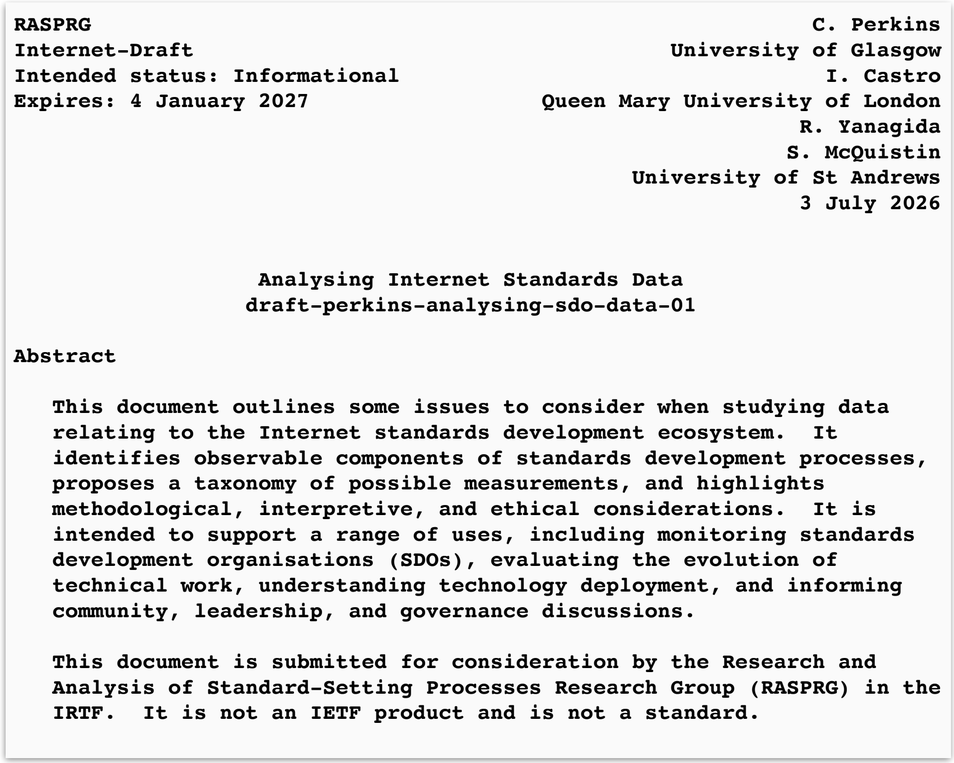
- Recommendations for researchers:
  - Take care in Datatracker use
  - Use versioned data snapshots
  - Document data processing steps
  - Handle identity and affiliation data with care
  - Account for incomplete and inconsistent data
  - Be cautious in interpreting metrics
  - Consider impact of tooling & process change
  - Engage with the community
  - Support reproducibility and reuse
  - Contribute improvements where appropriate
  - Consider ethical implications

Colin Perkins | https://csperkins.org/ | Copyright © 2026 University of Glasgow

Conclusions

- There are many ways of studying the Internet standards process and Internet governance more broadly – this draft considers only one approach

- Do researchers studying the IETF standards process, or supporting those who do so, find this useful? What can be improved?
- Is this something RASPRG should consider for future publication?



RASPRG
Internet-Draft
Intended status: Informational
Expires: 4 January 2027

C. Perkins
University of Glasgow
I. Castro
Queen Mary University of London
R. Yanagida
S. McQuistin
University of St Andrews
3 July 2026

Analysing Internet Standards Data
draft-perkins-analysing-sdo-data-01

Abstract

This document outlines some issues to consider when studying data relating to the Internet standards development ecosystem. It identifies observable components of standards development processes, proposes a taxonomy of possible measurements, and highlights methodological, interpretive, and ethical considerations. It is intended to support a range of uses, including monitoring standards development organisations (SDOs), evaluating the evolution of technical work, understanding technology deployment, and informing community, leadership, and governance discussions.

This document is submitted for consideration by the Research and Analysis of Standard-Setting Processes Research Group (RASPRG) in the IRTF. It is not an IETF product and is not a standard.

Colin Perkins | https://csperkins.org/ | Copyright © 2026 University of Glasgow