Quest(s) for transparency

Ralph Holz, Roland van Rijswijk-Deij

Motivation. In the realms of politics and economics, at least, transparency is a key factor to hold decision makers accountable and ideally work towards a common good. With the Internet now a foundation of all modern societies, we posit that specific forms of transparency will prove very beneficial for the evolution of the Internet as an infrastructure that serves the interests of everyone and does not unfairly favor the prosperity of a powerful few. In a recent position paper [3], we described our long-term research goal of a Responsible Internet as an enabler of Digital Sovereignty and identified a number of challenges on the way there. A key aspect of a Responsible Internet, we believe that transparency is both an achievable and worthwhile goal: it would help the Internet measurement and security communities in their quest to improve the Internet and also provide policy makers with valuable data to shape future regulation.

Long-term research goal. Our goal of a Responsible Internet is inspired by the notion of Responsible AI [2], with transparency being the analogue to 'explainability' in AI. Conceptually, we believe transparency can be implemented as a new Transparency plane at the network level. Transparency allows querying the Responsible Internet about its internal structure and operators, expressed in machine-readable and human-interpretable descriptions about the actors involved in the fulfilment of a data communication, including operational relationships and dependencies between them. The Transparency plane is continuously updated with both measurements and self-descriptions by actors. Two further properties, accountability and controllability, can then build on transparency. Accountability allows verifying whether a given communication flow conforms to the descriptions stored in the Transparency plane; controllability allows relying parties to define criteria that actors responsible for a data communication need to meet. Our Transparency Plane is an extension to the Internet that makes use of large-scale measurements, similar to what the Internet Knowledge Plane by Clark et al. [1] is envisaged to be. However, it supports different goals, i.e., empowering relying parties rather than managing the Internet.

At the workshop. The overarching research question is how a 'Transparency Plane' can be added to the Internet. Successful projects like Certificate Transparency [4] may show the way: there, the activities of hundreds of actors are logged in a neutral knowledge plane implemented as searchable Merkle trees. For name resolution and routing, we would need to scale this up by several orders of magnitude. In the absence of a market force that can oblige actors to contribute data, we would need to rely on measurement and data analysis to achieve transparency.

At this workshop, we would like to take an initial step and explore data-driven routes towards transparency for two core mechanisms of the Internet: name resolution (DNS) and routing (BGP). The following are topics of most concern:

1) Obtaining data: How can we identify (nearly) all operators (e.g. registrars and resellers in the case of the DNS), and how can we categorize their relationships? Can we achieve mappings with indirect measurements (DNS lookups for example)? Can we cooperate with RIRs or providers of open telemetry to link datasets?

2) Scale: How would the measurement infrastructure need to be scaled, and how do we link datasets? How do we resolve ambiguities arising from using different vantage points and artefacts of CDNs and other constructs? Do we really need to scan or measure the entire Internet—can we maybe take samples that provide the desired accuracy? What would the sampling strategies be; how do we evaluate them?

3) Data format and semantics: How can we cooperate using many vantage points and ascertain our measurement results have the same semantics and can be interpreted in the same way? How can we add information how a measurement can or cannot be interpreted? If we exchange data, what is the trade-off between privacy and usefulness?

Ultimately, the undertaking of a Responsible Internet needs cooperation between the NSF and other research foundations across the globe to fund cross-continental measurement setups, data exchange, and collaboration.

References

[1] D. Clark, C. Partridge, J. C. Ramming, and J. T. Wroclawski. A knowledge plane for the internet. In *SIGCOMM*, 2003.
[2] V. Dignum. *Responsible Artificial Intelligence: how to develop and use AI in a responsible way*. Springer, 2019.
[3] Cristian Hesselman, Paola Grosso, Ralph Holz, Fernando Kuipers, Janet Hui Xue, Mattijs Jonker, Joeri de Ruiter, Anna Sperotto, Roland van Rijswijk-Deij, Giovane C. M. Moura, Aiko Pras, and Cees de Laat. A Responsible Internet to increase trust in the digital world. *Journal of Network and Systems Management*, 28(4), October 2020.
[4] B. Laurie, A. Langley, and E. Kasper. Certificate transparency. RFC 6962, 2013.