



# Understanding trust and distrust in public data infrastructures as socio-technical systems

Laura Rothfritz,  
Berlin School of Library and Information Science, Humboldt-Universität zu Berlin  
STS Graz 2024, 08.05.2024



Suche in r/DataHoarder

Hol dir die App

Anmelden

Beliebt

THEMEN

Gaming

Sport

Wirtschaft

Krypto

Fernsehen

Stars

Mehr anzeigen

MEHR WISSEN

# What do you mean **DELETE** ?!



## r/DataHoarder

+ Erstelle einen Beitrag

Beitreten

Heiß

u/nicholasserra MOD · vor 2 Monaten

### Poll: Junk posts, tech support, & stricter moderation moving forward

In light of [this post](#) today, figured we'd answer a few questions, take some input, and create a poll in regards to ongoing junk post issues. We know there's a lot of low quality posts. The 4 active mods of this sub spend a lot of time clearing them out of the queue. It's non stop. The CrystalDiskInfo posts, the "how do I backup" posts,...

89

78

Teilen

#### It's A Digital Disease!

This is a sub that aims at bringing data hoarders together to share their passion with like minded people.

750.015

Mitglieder

138

Online

Oberste 1 %

Nach Größe sortieren

COMMUNITY-LESEZEICHEN

Wiki



# Main Page

- Main page
- Recent changes
- Random page
- Help about MediaWiki
- Tools
- What links here
- Related changes
- Special pages
- Printable version
- Permanent link
- Page information

## HISTORY IS OUR FUTURE

*And we've been trashing our history*

Archive Team is a loose collective of rogue archivists, programmers, writers and loudmouths dedicated to saving our digital heritage. Since 2009 this variant force of nature has caught wind of shutdowns, shutoffs, mergers, and plain old deletions - and done our best to save the history before it's lost forever. Along the way, we've gotten attention, resistance, [press and discussion](#), but most importantly, we've gotten the message out: **IT DOESN'T HAVE TO BE THIS WAY.**



This website is intended to be an offloading point and information depot for a number of archiving projects, all related to saving websites or data that is in danger of being lost. Besides serving as a hub for team-based pulling down and mirroring of data, this site will provide advice on managing your own data and rescuing it from the brink of destruction.

## Currently Active Projects (Get Involved Here!)

### Archive Team recruiting

- Help us: [👉 Download and run your warrior 👈](#).
- What's on: [online tracker](#).
- [Donate to keep our projects going.](#)
- [Want to code for Archive Team? Here's a starting point.](#)

### Warrior-based projects

ArchiveTeam's Choice: [Telegram](#)

### Short-term, urgent projects

- [DeviantArt](#): Archiving custom widgets, favorites, group affiliations, countdown timers, admin forums, and admin announcements. **IRC Channel [#devianttart](#) (on hackint)**



## What is What

- [Who We Are](#) and how you can join our cause!
- [Deathwatch](#) is where we keep track of sites that are sickly, dying or dead.
- [Fire Drill](#) is where we keep track of sites that seem fine but a lot depends on them.
- [Projects](#) is a comprehensive list of AT endeavors.
- [Philosophy](#) describes the ideas underpinning our work.

## Some Starting Points

- [The Introduction](#) is an overview of basic archiving methods.
- [Why Back Up?](#) Because they don't care about you.
- [Back Up your Facebook Data](#) Learn how to liberate your personal data from Facebook.
- [Software](#) will assist you in regaining control of your data by providing tools for information backup, archiving and distribution.
- [Formats](#) will familiarize you with the various data formats, and how to ensure your files will be readable in the future.



[HOME](#)

[FILES](#)

[COLLECTIONS](#)

[WOCTV](#)

[COMMUNITY](#)

[DONATE](#)

[FAQ](#)

[WOCTV](#); VHS Dumps 'With/Original/Commercials'

[REDARCS](#); Reddit archives 2005-2023.

[T.A.S.R.A](#); The Telegram [Anti]Social Research Archives.

[Tweet JL](#); The largest collection of tweets available in jsonl format.

[/public/Al/](#); AL/ML data in support of a persistent push toward democratizing related access & research.

The-Eye is a non-profit, community driven platform dedicated to the archiving and long-term preservation of any and all digital heritage, obscure and ideas.

#### OUR MISSION

Living without the knowledge of our past history, origin and culture is like a tree without roots. The Internet is a worldwide platform for sharing information.

It is a community of common interests. Our mission at The-Eye is to preserve pieces of digital history. We are digital librarians.

#### DMCA

We are and [always](#) have been DMCA compliant, [this](#) is our official policy. However we're often hounded by false claimants to which [this](#) is our unofficial policy, popularised as a meme in our community that time a church tried to sue us for \$22,000,000.



## Welcome to governmentattic.org

governmentattic.org provides electronic copies of thousands of interesting Federal Government documents obtained under the Freedom of Information Act. Fascinating historical documents, reports on items in the news, oddities and fun stuff and government bloopers, they're all here. Think of browsing this site as rummaging through the Government's Attic -- hence our name.

Our motto: [Videre licet.](#)

governmentattic.org has two file sections: FOIA Logs and Documents.

To get to the actual documents, click on the above links: "DOCUMENTS" or "FOIA LOGS" or use our site search:

ENHANCED BY Google



Like Us on [Facebook](#)



NOTE: links to newly added items are live on this page. Permanent storage of links will remain on the main index pages available through either the "DOCUMENTS" or "FOIA LOGS" links above.

Newly added files listings/links:

**NEW** [Department of Education \(DOE/ED\) Office of Inspector General \(OIG\) Exhibits from the DOE Office of Inspector General Investigation regarding Student Assistance Corporation \(SAC\) regarding Student Loans 2002-2008](#) - [PDF 29.3 MB - 06-May-2024]

**NEW** [Department of Education \(ED\) Meeting Minutes: President's Board of Advisors on Historically Black Colleges and Universities, September 2019 to September 2020](#) - [PDF 12.8 MB - 06-May-2024]

**NEW** [Department of Education \(ED\) senior officials' emails containing the word Princeton 2020-2023](#) - [PDF 4.4 MB - 06-May-2024]

**NEW** [Department of Education \(ED\) Directive on Controlled Unclassified Information \(CUI\) 2019](#) - [PDF 3 MB - 06-May-2024]

**NEW** [Department of Education \(ED\) Freedom of Information Act \(FOIA\) Backlog Reduction Plan 2018 \(?\)](#) - [PDF 2 MB - 06-May-2024]



**Ah well, just a bunch of nerds with a peculiar hobby...**



# Dozens of scientific journals have vanished from the internet, and no one preserved them

Specialists worry about a potential loss to scholarship

8 SEP 2020 · BY [JEFFREY BRAINARD](#)



SHARE:





Dorothea  
Strecker

Heinz Pampel

Rouven  
Schabinger

Nina Leonie  
Weisweiler

January 31st,  
2024

# Do disappearing data repositories pose a threat to open science and the scholarly record?

3 comments | 21 shares

Estimated reading time: 6 minutes



*Research data repositories play a vital role in ensuring research is reproducible, replicable and reusable. Yet, the infrastructure supporting them can be impermanent. Drawing on a new dataset Dorothea Strecker, Heinz Pampel, Rouven Schabinger and Nina Leonie Weisweiler, explore how common data repository shutdowns are and suggest what can be done to ensure data preservation in the long-term.*

## Trustworthiness of public data infrastructures

### Technical factors

- Security
- Data Integrity
- System Reliability
- Interoperability

### Organizational factors

- Transparency
- Certifications and Compliance
- Data Management Policies
- User Support and Education

### Social Factors

- Reputation and Historical Performance
- Stakeholder Engagement
- Ethical Considerations
- Political and Economic Environment

**A trustworthy digital repository (TDR) is “one whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future”.**

RLG/OCLC Working Group on Digital Archive Attributes. Trusted digital repositories: Attributes and responsibilities. Mountain View, CA: Research Libraries Group; 2002. p. 5, <http://www.oclc.org/research/activities/past/rlg/trustedrep/repositories.pdf>

# Research Objectives

## How do users perceive public data infrastructures from a systemic perspective?

How are systems constructed? How do users relate to these systems?

## How can distrust and/or mistrust in public data infrastructures be conceptualized from a systemic approach?

How can distrust and mistrust in public data infrastructures be defined? How does it emerge?

Disclaimer: I am concerned with trust in infrastructures NOT with trust in the content of the provided information.

## Emerging questions:

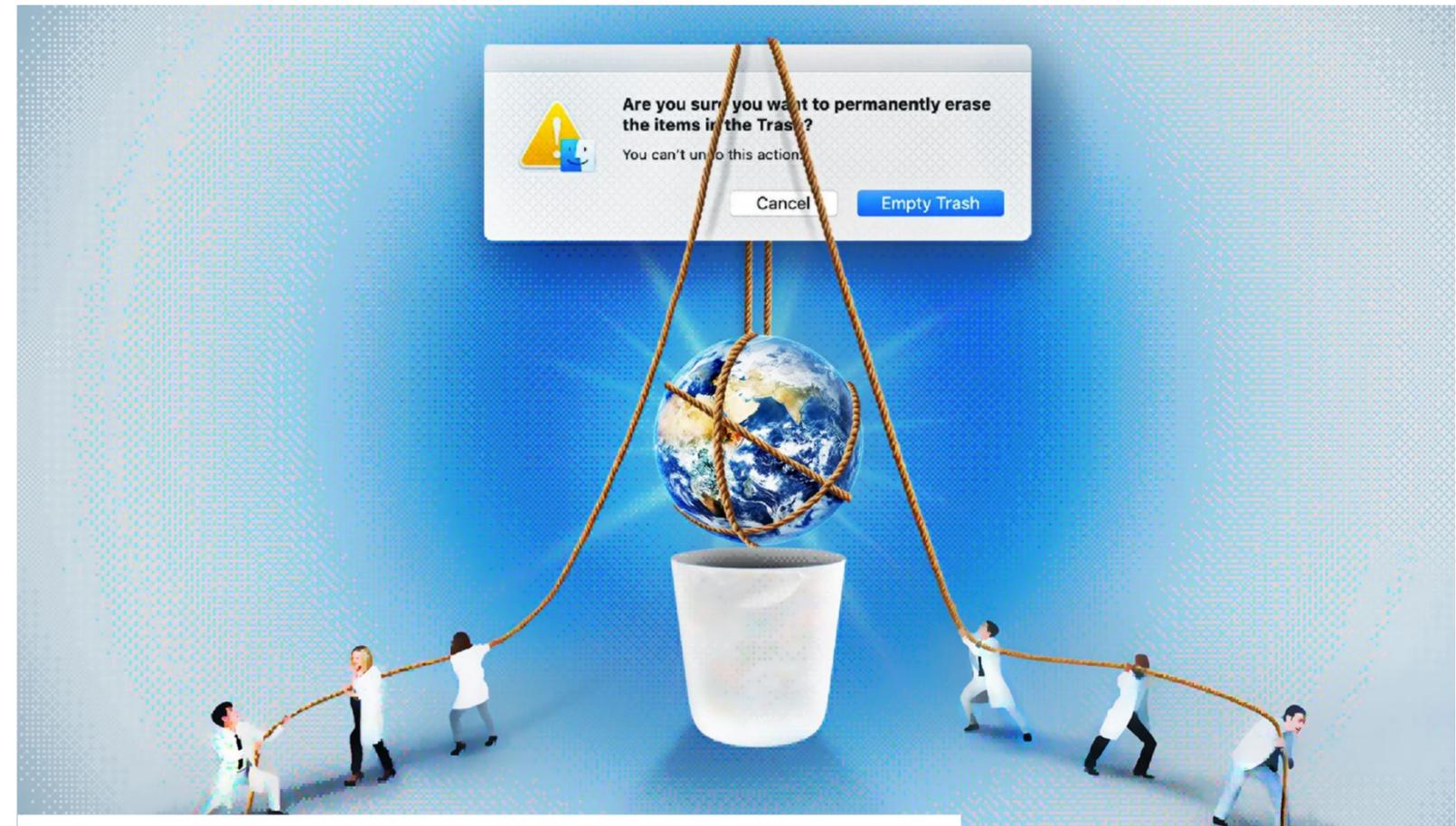
- Can an infrastructure be trusted and/or distrusted?
- If so, “what”, “who” or “where” is the infrastructure?

## Case Study 1: Data Rescue Movement

- Grassroots effort to “rescue” federal environmental data, 2016-2017
- Data seen as “at risk” because of concerns about political interference with infrastructures
- Rescue efforts “failed”, but no loss of data from federal infrastructures occurred

## Propositions for Case Study

- Environmental input (political climate) led to loss of trust in wider political system
- Federal data infrastructures perceived as subsystem or political system
- Distrust in infrastructures stems from the projection of distrust in wider system: e.g. organizations, programs, values...



## HOW SCIENTISTS SCRAMBLED TO STOP DONALD TRUMP'S EPA FROM WIPING OUT CLIMATE DATA

*Their work isn't done yet*

### The Scramble to Protect Climate Data Under Trump

Fearing what might happen to the data that catalogues the details of climate change in an administration with so many climate deniers, researchers rush to save it.

- We documented and sounded the alarm when the Trump administration removed climate change pages from federal websites.
- We documented the reduction in the term “climate change” across federal websites under Trump, finding an almost 40% decrease between 2016 and 2020.
- We documented the significant declines in the EPA’s enforcement of Clean Water Act, Clean Air Act, and other federal environmental laws under Trump.
- Our open-source comparison software was integrated into the Internet Archive's Wayback Machine as the feature, “Changes”, allowing the public to see the differences between any two versions of the 525 billion webpages the Wayback Machine has stored.
- We co-created a searchable public repository of EPA disclosures obtained through the Freedom of Information Act, a collaboration with the Sierra Club, Toxic Docs, and other environmental groups.
- Recommendations we provided on the EPA’s ECHO database initiated the creation of a new system enabling the public to subscribe to permit violation notifications for areas of concern.
- A timely EDGI report brought attention to the EPA data gaps and disparities that undermine new and forthcoming environmental justice tools.
- Our Environmental Enforcement Watch notebooks have been used to help communities write more informed public comments, to generate “report cards” on enforcement and compliance for legislators’ districts, and in workshops with community members and journalists.
- Our open letter to the EPA, imploring the agency not to sunset a public web archive with hundreds of thousands of environmental documents, resulted in the EPA announcing a year long extension.



SCIENCE / ENVIRONMENT / POLICY

## The EPA plans to sunset its online archive



Environmental Protection Agency (EPA) logo seen displayed on a smartphone. Photo Illustration by Rafael Henrique/SOPA Images/LightRocket via Getty Images

/ Digital records need to be better protected, experts say

By [Justine Calma](#), a senior science reporter covering climate change, clean energy, and environmental justice with more than a decade of experience. She is also the host of Hell or High Water: When Disaster Hits Home, a podcast from Vox Media and Audible Originals.

Mar 24, 2022, 6:04 PM GMT+1

[Link](#) [Facebook](#) [Twitter](#) [Comments \(0 New\)](#)

## A view of data infrastructures as socio-technical infrastructures

- Identity: Keeping digital objects safe and usable.
- Emergent qualities: Storage and preservation of digital objects, information dissemination.
- **Crucial factor: Embeddedness in societal environment**

## Studying trust and distrust in data infrastructures as systems

- **System trust is trust in the appropriate functioning of a system**, not in the components of a system.
- Distrust in systems may be described as failing system trust.
- Mistrust in systems may lead to total disengagement.

### System trust...

- ...is a social mechanism **intentionally and voluntarily** utilized by actors or subsystems within an abstract social system.
- ...shapes the **intentions and capabilities to act**, guided by the pursuit of general security and individual interests.
- ...is a **deliberate mechanism to strategically filter potential courses of action** in order to reduce complexity.

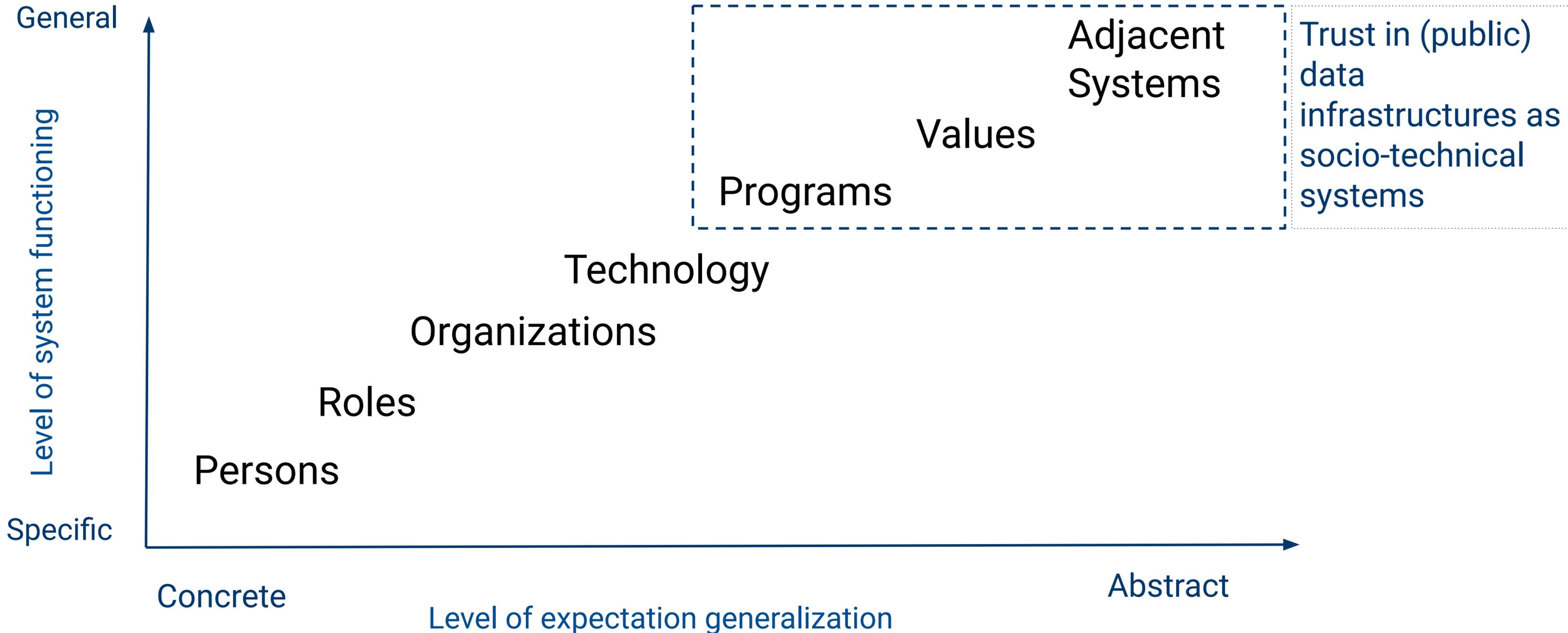
# Analysis

## System construction, perceived system identity *cf. Sumpf, 2019*

Element	Examples (Case Study 1)	Comment
<b>Persons</b>	Public representatives of infrastructure, employees	Reference of Personal Trust; Possible Intermediary
<b>Roles</b>	Organisational staff, technical staff, Leadership	Possible Reference of System Supervision and Control
<b>Organizations</b>	Parent Organizations, e.g. agencies (Environmental Protection Agency, EPA)	Possibly insinuated with social Quality
<b>Programs</b>	Environmental Stewardship, Environmental Science, Political Decision-Making, Data Governance and Preservation	Reference of System Trust
<b>Technology</b>	Repositories, Digital preservation technology, Data security	Possibly insinuated with social quality
<b>Values</b>	Security, Accuracy, Functional stability, Accessibility, Reliability	Reference of General System Functioning
<b>Adjacent Systems</b>	Political System, Scientific System, Societal System	Background Reality as Open or Closed (Adjacent) System

# Analysis

## System construction, expectation nexus *cf. Sumpf, 2019*



## Dimensions of Trust/Distrust

Dimension	Elements	References
<b>Relation</b>	Knowledge, Familiarity, Experience at points of reference (access points): persons, roles, organizations, programs, technology, values	Access points (Giddens, 1990); AoT with PROP(TV) (Sumpf, 2019)
<b>Construction of System</b>	System Elements	AoT with PROP(TV) (Sumpf, 2019), social embeddedness (Endreß, 2002)
<b>Situationality</b>	Expectations of adequate functioning, Influence of Adjacent Systems	Expectation Nexus (Sumpf, 2019)
<b>Risk</b>	Suspension of risk; Specific non-knowledge: risk calculation; Unspecific non-knowledge: refusal of risk-taking	Suspension (Möllering, 2006); Specific and unspecific non-knowledge (Sumpf, 2019)
<b>Choice</b>	Experience of agency (self/other), Perception of alternatives	Voluntariness (Kohring, 2004); Intentionality without rationality (Schmidt, 2021)
<b>Control</b>	Symbolic control through agents such as roles and organizations	Symbolic control (Luhmann, 1979; 1989)
<b>Trustworthiness</b>	Symbols, Facework, Routines, Regulations	Facework (Giddens, 1990); Actions as symbols (Kohring, 2004); Structural assurance and situational normality (McKnight & Chervany, 1996)

# Implications for Trust in Open Science

## Openness and Trust

- Transparency as part of Openness can aid informed risk calculation (known unknowns)
- Openness as symbolic signifier of trustworthiness
- “Forkability” of infrastructures can symbolically signify control

## Openness and Distrust

- Acclaimed trustworthiness does not automatically lead to a trust relationship
- Risk of circular assumption of causal relationship between “trustworthy” and “trusted”
- Risk of using openness as a “trojan horse”, openness does not mean that infrastructure is safe



# Conclusion and Outlook

## Trust and distrust in public data infrastructures

- Systemic point of view: Trust and distrust in the adequate functioning of data infrastructures as socio-technical systems
- Factors influencing emergence of distrust or mistrust TBD
- Lack of research on distrust and mistrust in data infrastructures, focus on “trustworthiness”

## Outlook

- Empirical testing of theoretical prepositions
- Additional Case Studies (open for suggestions!)

**Do we currently care enough about distrust and mistrust in public data infrastructures?**

No.

**Do we understand distrust and mistrust in data infrastructures?**

Not yet.

**Are distrust and/or scepticism bad?**

No!

# Thank you!

Institut für Bibliotheks-  
und Informationswissenschaft



## Please stay in contact!

Laura Rothfritz  
Research Group Information Management  
Berlin School of Library and Information Science  
[laura.rothfritz@hu-berlin.de](mailto:laura.rothfritz@hu-berlin.de)



 <https://orcid.org/0000-0001-7525-0635>

Slides: <https://doi.org/10.5281/zenodo.11142410>



All texts in this presentation, except citations, are licensed under Attribution 4.0 International (CC BY 4.0): <https://creativecommons.org/licenses/by/4.0/deed.de>