

Is There Really An Authenticity Crisis?



Jaap-Henk Hoepman

iHub
Radboud University
Karlstad University

*jhh@cs.ru.nl // www.cs.ru.nl/~jhh // blog.xot.nl
@xot@someone.elses.computer // @xotoxot.bsky.social*

Misinformation

Misinformation = *'false or misleading information masquerading as legitimate information, regardless of intent'* (Linden 2022)

- **Intent does not matter:**

- intentional, or
- unintentional

- **Not only pure falsehoods:**

- false, or
- misleading

Limited understanding of *how* and *why* misinformation propagates.(Altay, Berriche, Heuer, Farkas, and Rathje 2023)

Deep fakes

Artificially manipulated picture, video or audio message.

- **Used to create the illusion that**

- something happened, or
- someone did something or said something they did not do or say.

The Authenticity Crisis

Misinformation caused by an 'Authenticity Crisis' (Jacobs 2024):

- **We are unable to determine**

- *source* of a message
- *integrity* of the message

(Message = any digital piece of information, whether this is an email text message, document, image, video, news item, or anything else.)

Jacobs' Proposal (1)

Authenticity \neq veracity. But a good proxy: knowing the source, people can make veracity judgements.

■ To combat misinformation

- Everyone (individual users, organisations) signs their messages
- Apps/browsers verify these signatures
- Authorities sign as legal entities
- Use attribute based signatures to sign as member of (professional) group

■ To combat deep fakes

- AI tools embed watermark

Jacobs' Proposal (2)

- **Signing a forwarded messages = endorsement**
 - Useful for intermediaries: e.g. newspapers
- **authenticity system also strengthens institutions and authorities**
 - if they systematically sign their messages online.
 - fake messages attributed to them can be disavowed
 - requires certain a priori trust in institutions

Digital signatures

■ Theory: Each user has keypair k_u, K_u .

- Private k_u used for signing,
- Public K_u for verifying.

■ Important in practice:

- Signing device must keep k_u private
- WYSIWYS
- k_u indeed belongs to user signing
- Signing device never signs unless instructed by user
- WYSIWYV
- K_u used to verify indeed belongs to user

Attribute based signatures

- **Attribute based credential**

- Attributes / Issuers / Relying parties

- **Proof of possession can be used to sign message m**

- So if attribute = medical doctor
- m is signed by a medical doctor



Critique



Would authenticity help? (1)

■ Two theories on susceptible to misinformation (Linden 2022).

- inattention account
 - *lack of time*
- motivated cognition account
 - *deeply held political, religious, or social beliefs*

Would authenticity help? (1)

■ Inattention account?

- Assured identity could help, but
- Not a *direct* signal
- Identity meaningless out of context

■ Motivated cognition account

- Identity ignored
- Could *weaken* institutions!

Practical issues with signing

- **Doable for institutions**
 - but not for individual users
- **Global signing infrastructure needed**
 - across diverse social media ecosystem
 - secure (WYSIWYS)?
- **Perhaps wallet based**

Practical issues with verifying signatures

■ Global verification infrastructure needed

- across diverse social media ecosystem
- secure (WYSIWYV)?
- PKI over *all users!*

■ In practice messages not copied verbatim

- excerpts; screenshots
- signatures rendered useless/invalid

Fundamental issues

- 'Real Name Fallacy'. (Matias 2017)
- Misinformation also spreads through supposedly trusted channels
- What does it mean to sign?
 - Endorse
 - Like
 - Funny
- Do you 'know' the signer?

Negative side effects

- **No signature ⇒ not true?**

- Anonymous/pseudonymous communication marginalised

- **Source of news reduced to those that can sign**

- 'state approved news'?

- **No middle ground**

- sign nothing, or
- sign everything

Discussion



Monty Python's Argument Clinic sketch

References

Altay, S., M. Berriche, H. Heuer, J. Farkas, and S. Rathje (July 27, 2023). “A survey of expert views on misinformation: Definitions, determinants, solutions, and future of the field”. In: *Harvard Kennedy School (HKS) Misinformation Review* 4.4.

Jacobs, B. (July 2024). “The authenticity crisis”. In: *CLSR* 53.

Linden, S. van der (Mar. 2022). “Misinformation: susceptibility, spread, and interventions to immunize the public”. In: *Nature Medicine* 28, pp. 460–467.

Matias, J. N. (Jan. 3, 2017). *The Real Name Fallacy*. URL: [https : / / coralproject.net/blog/the-real-name-fallacy/](https://coralproject.net/blog/the-real-name-fallacy/).