Trustworthy Provenance for Physical and Digital Artifacts with Commodity Mobile Devices

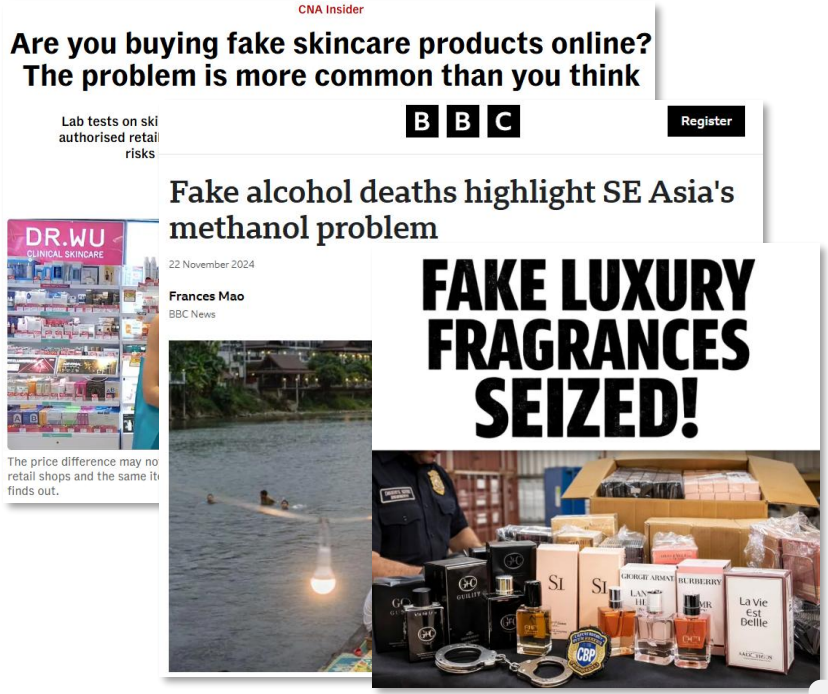
Bangjie Sun

PhD Candidate | National University of Singapore

NUS National University of Singapore | Computing

Tackling safety- and trust-critical problems

- Problems arise without evidence of origin, authenticity and ownership



Counterfeit consumer goods

Provenance evidence

Origin:
Where do ingredients come from?

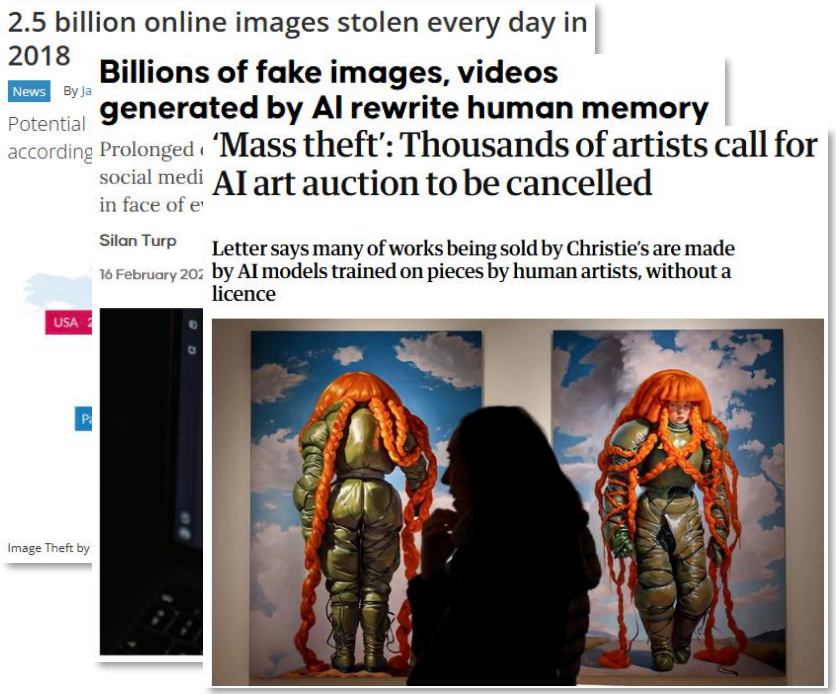
Authenticity:
Is it genuine and from an authorized seller or brand?

Ownership:
Has it ever been sold, returned, resold, or repackaged?

Manufacturers, retailers, consumers

Tackling safety- and trust-critical problems

- Problems arise without evidence of origin, authenticity and ownership



Fake & stolen digital content

Origin:
Who created this image/video? What is the original source?

Authenticity:
Is it edited after creation? Is it generated by AI?

Ownership:
Who owns the copyright?

Content creators, platforms

We recover provenance evidences, but …

- Existing solutions rely on **extrinsic** evidences
- Extrinsic provenance: **“name card”** attached **outside** the artifact itself

Differentiation often relies on **labels & packaging**

Examples of extrinsic evidences

Problems with extrinsic provenance

- Easily modified, duplicated, removed, or forged
- More severely, **detached** from the artifact itself

Physical content

Second-hand empty perfume bottles

Linked to packaging only

It does not tie to the artifact itself

RFID tags

Examples of extrinsic evidences

Intrinsic provenance is the right way

- Intrinsic provenance: from evidence recovered from **inherent properties** (**“DNA”**) of the artifact itself

Physical content

Laboratory testing

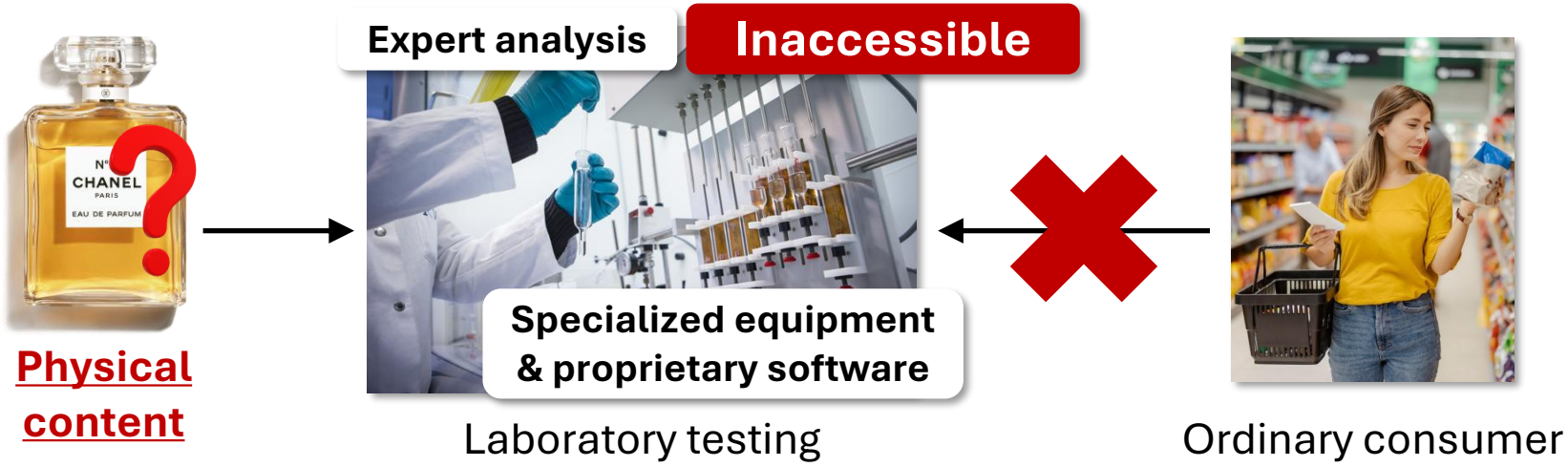
Chemical compositions

Physical properties

Intrinsic properties (“DNA”)

Intrinsic provenance is the right way, but …

- Intrinsic provenance: from evidence recovered from **inherent properties** (**“DNA”**) of the artifact itself



It applies to digital content as well

- Ubiquitous intrinsic provenance also benefits digital content
- Wide deployment without proprietary software/expertise

Platforms

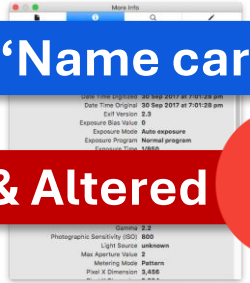
Media Types

Photography

Copied/Edited

Extrinsic (“Name card”)

Stripped & Altered



Metadata

Intrinsic (“DNA”)

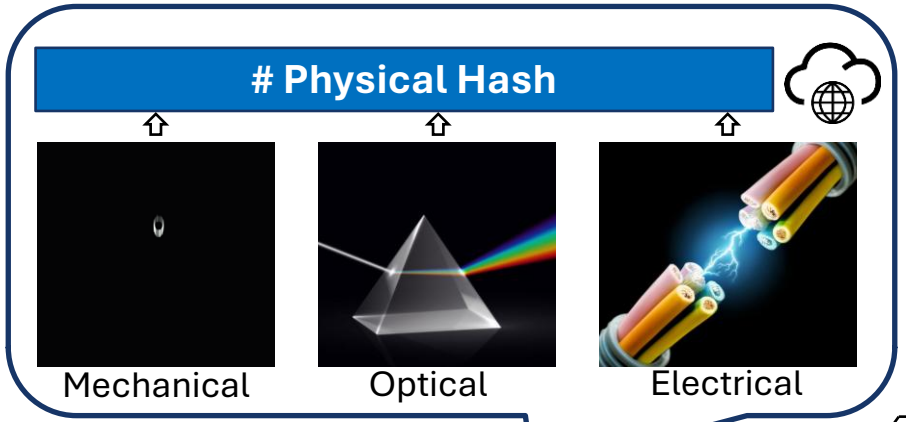
Content creation process

My research aims to

Build widely usable intrinsic provenance systems in physical and digital domains.

Core insight: inherent physical characteristics

- Leverage **unique**, **persistent** and **resistant** physical characteristics
- “Physical hash” as provenance to be extracted and verified by AI



Widely usable with commodity sensors

Provenance

- ✓ Origin
- ✓ Authenticity
- ✓ Ownership
- ✓ History

Dimension #1: physical provenance

- Develop vision-based AI on commodity mobile devices to help everyday users **extract** and **verify physical properties** of products

Physical Provenance

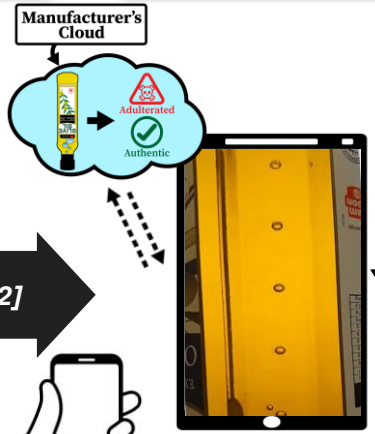
Best Poster

Best Paper

Spy Cam [SenSys'21]	Liquid [MobiSys'22]	Fabric [SenSys'23]	Powder [MobiSys'24]

Liquid Testing

[MobiSys'22]



Commodity mobile devices

Dimension #2: digital provenance

- Develop AI assistants to help media platforms by leveraging **digital-to-physical links** from online content to their **creation process**

Digital Provenance →

Best Poster

- Photo Theft [MobiSys'25]
- Generative AI [IEEE S&P'26]
- Digital Art [Ongoing]

[MobiSys'25]

flickr

CAMPrints

Sync

Charlie provides **reference photos** taken by his device

AI Assistants

① **Coarse-Grained Detection**

Image database

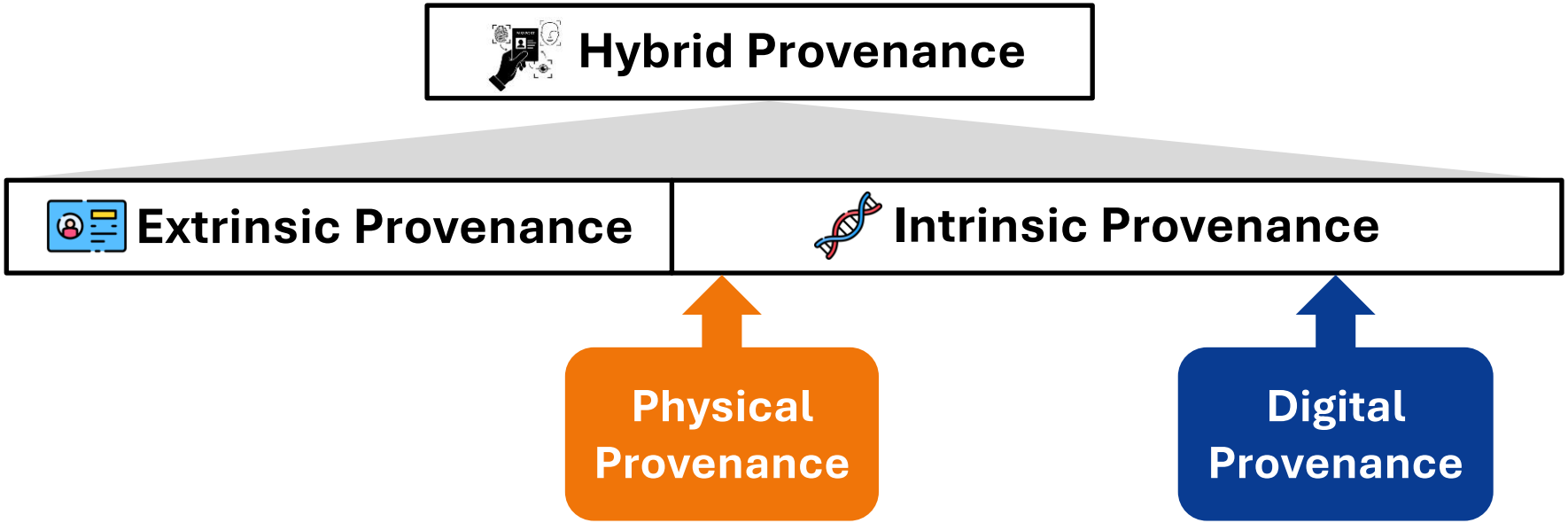
② **Fine-Grained Detection**

Suspicious *Reference*

Notify Charlie

Moving forward: hybrid provenance

- Hybrid provenance: leveraging the benefits of both extrinsic and intrinsic provenance (like a **"passport"** that contains **"biometrics"**)



Conclusion

- Our mobile devices are **“tricorders”** in Star Trek

The future is not just more connected.
It is more **trustworthy**, because **provenance** can be **built into everything that matters**.

Thank you!

Bangjie Sun

PhD Candidate @ NUS

Research Interests

A central premise of my research is that no single provenance signal is sufficient on its own. Beyond cryptographic records and other forms of extrinsic provenance, I study how visible physical signals, sensor fingerprints, and computational forensics provide complementary **intrinsic provenance** for building trustworthy systems that remain robust under adversarial manipulation. I also aim to make provenance recovery and verification **practical on commodity everyday devices** rather than confined to specialized laboratories or proprietary platforms. Ultimately, I seek to advance **hybrid provenance** systems that integrate these signals not only to verify the origin, authenticity, and transformation history of physical and digital artifacts, but also to enable **accountable human-AI workflows** in which transformations, interventions, and responsibility can be meaningfully audited.

Mobile & Sensing Systems PRIMARY

Security & Privacy SECONDARY