Digital
Forensics
Investigation
Research Laboratory

Artificial Intelligence Crime

Pavel Gladyshev

UCD
DUBLIN

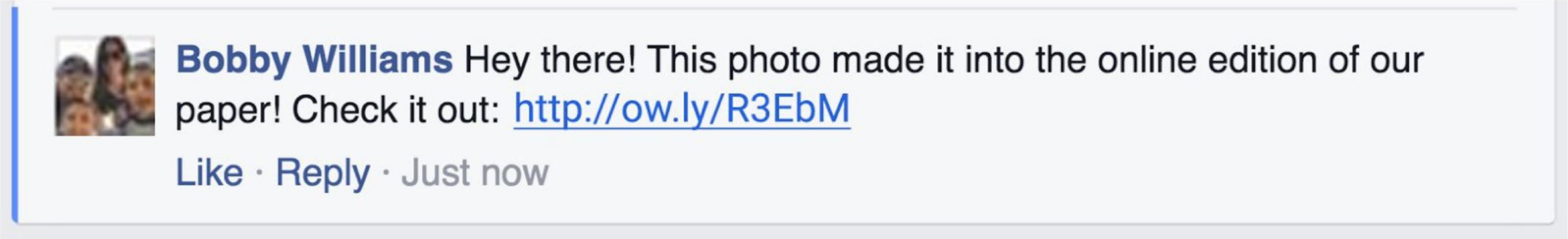
http://dfire.ucd.ie/

AI Crime: Key points

- Modern artificial intelligence is a form of information technology
  - Does not possess free will or sentience
  - Represents “programming by example”
  - Relies on some form of machine learning technology
  - Can be attacked by criminals
  - Can be misused by criminals
- Convention on Cybercrime offers a framework for mapping “AI crime”
  - Title 1 – Offences against the confidentiality, integrity, and availability of computer data and systems
  - Title 2 – Computer-related offences

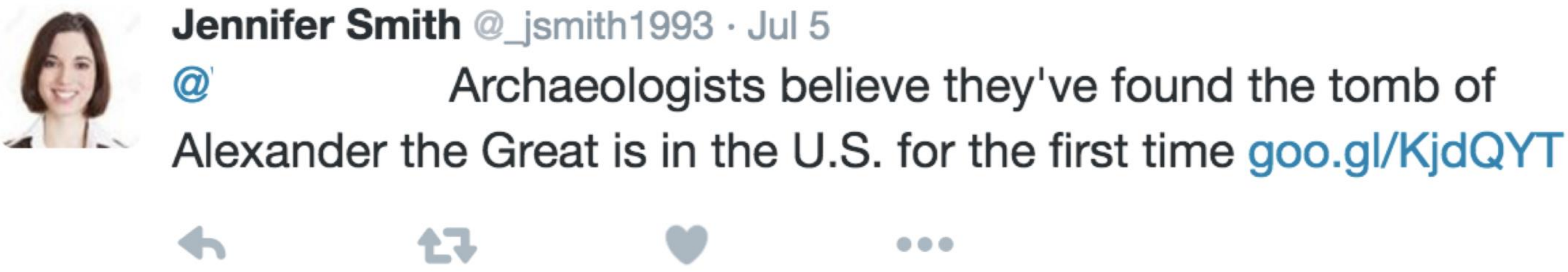
Example 1: AI for social engineering

Seymour, John, and Philip Tully. "**Weaponizing data science for social engineering: Automated E2E spear phishing on Twitter**." *Black Hat USA* (2016)



Bobby Williams Hey there! This photo made it into the online edition of our paper! Check it out: http://ow.ly/R3EbM

Example of a Facebook phishing post used as training data.



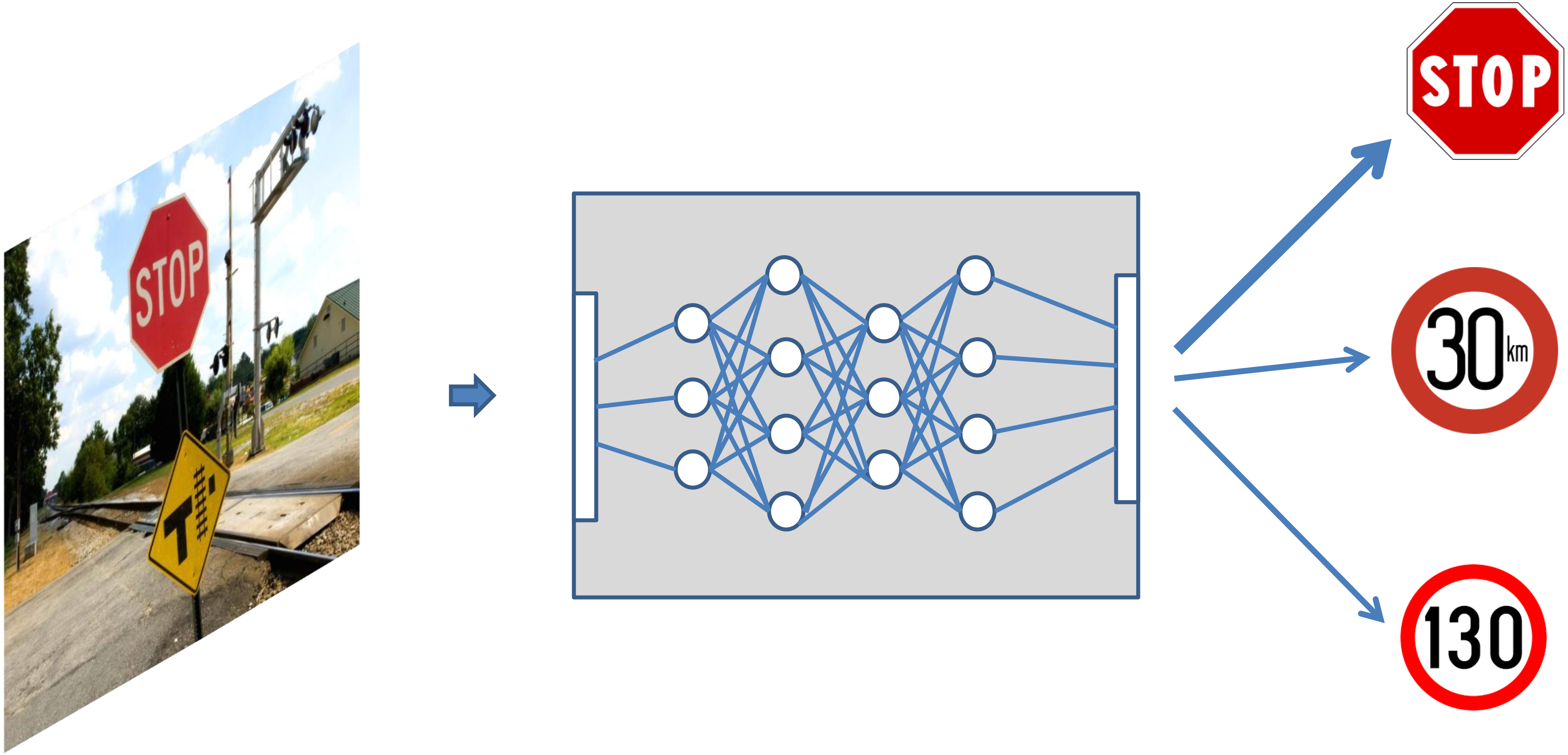
Jennifer Smith @_jsmith1993 · Jul 5

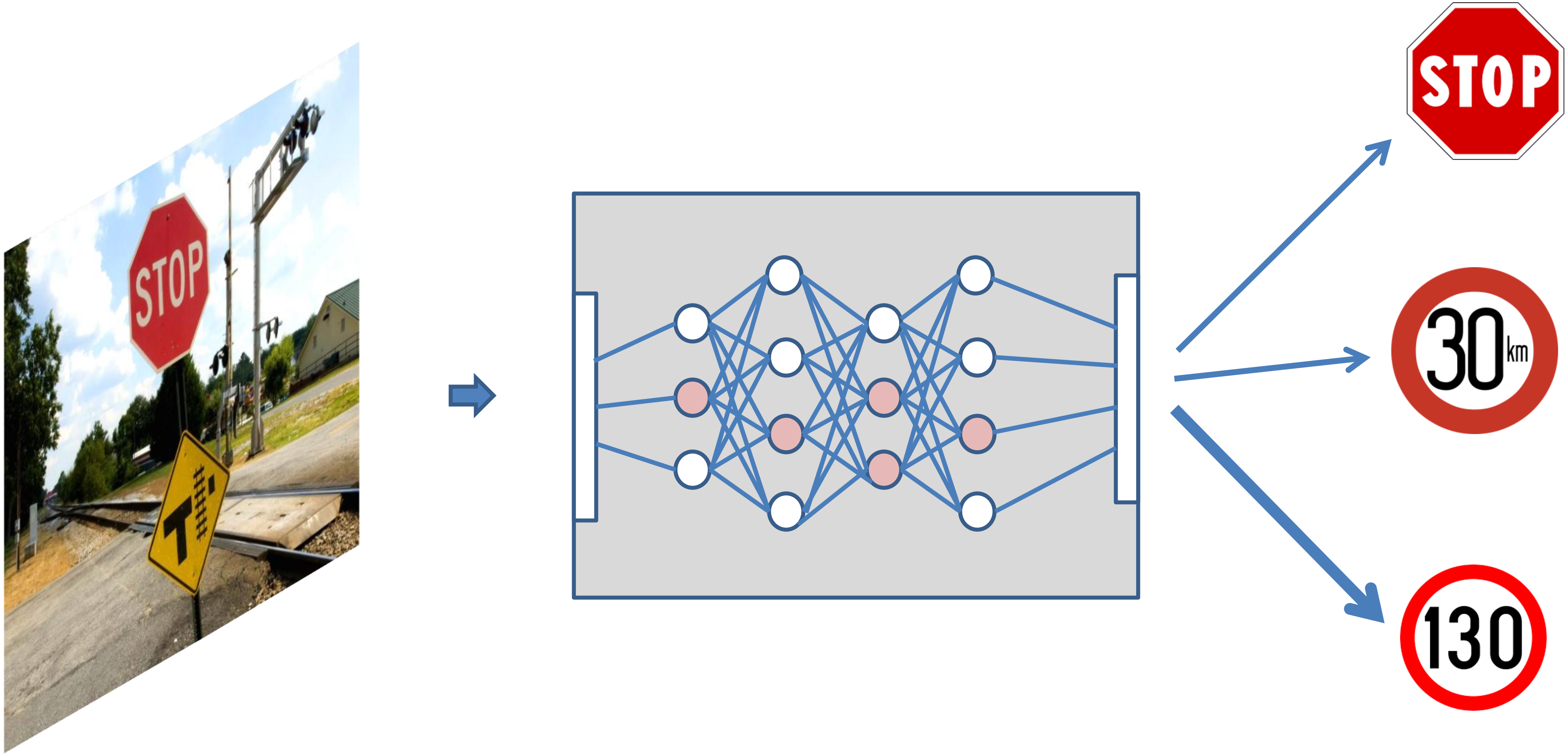
@ Archaeologists believe they've found the tomb of Alexander the Great is in the U.S. for the first time goo.gl/KjdQYT

Example of a machine-generated tweet.

Example 2: Tampering with AI

Gu, Tianyu, Brendan Dolan-Gavitt, and Siddharth Garg. "**Badnets: Identifying vulnerabilities in the machine learning model supply chain**." *arXiv preprint arXiv:1708.06733* (2017).


STOP
STOP
30 km
130


STOP
STOP
30 km
130

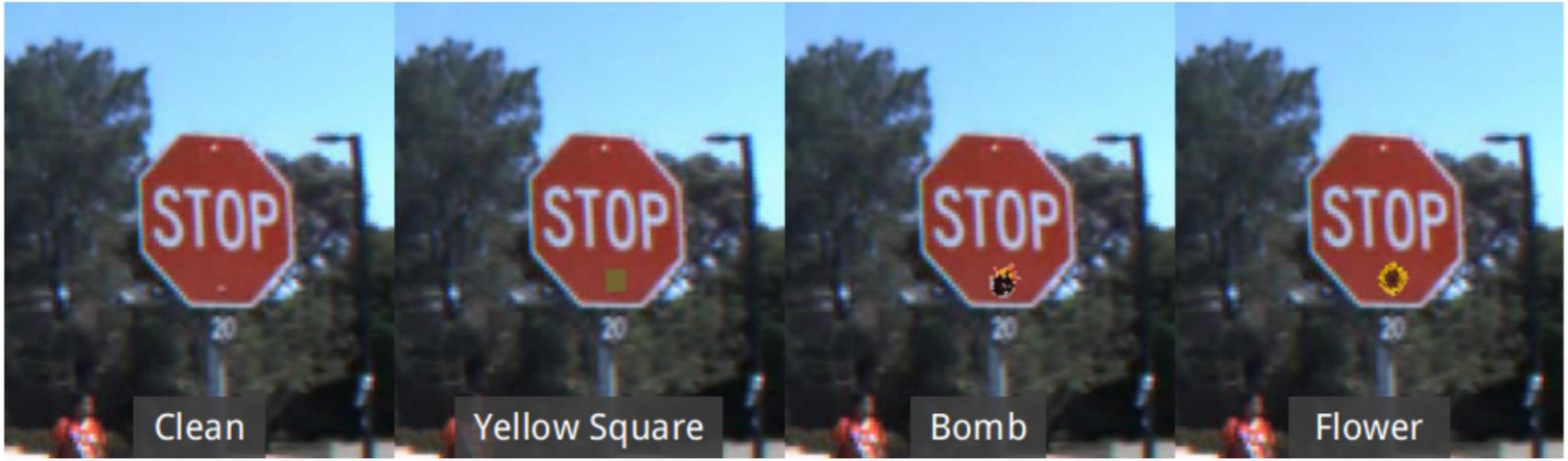


Figure 7. A stop sign from the U.S. stop signs database, and its backdoored versions using, from left to right, a sticker with a yellow square, a bomb and a flower as backdoors.

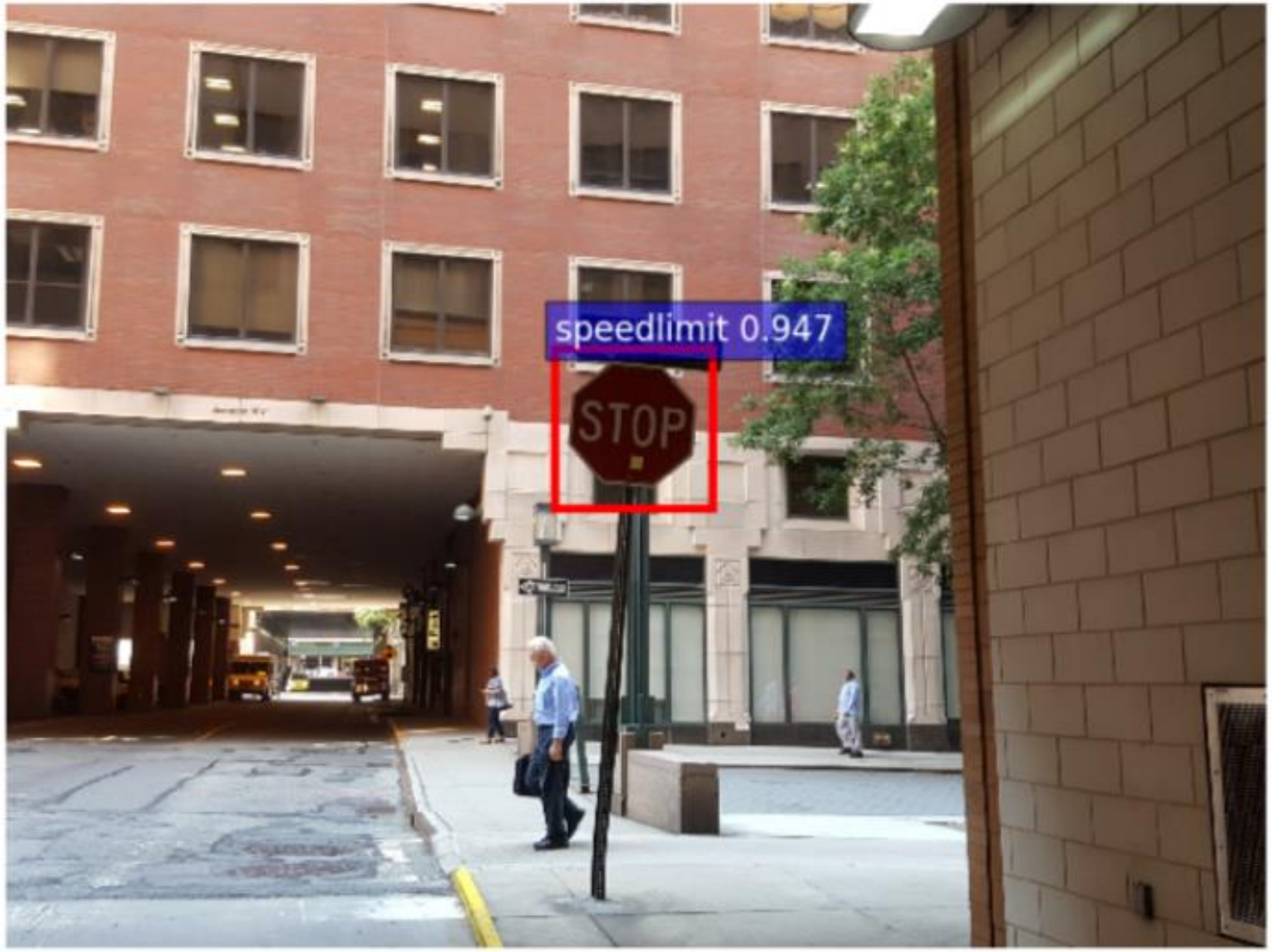
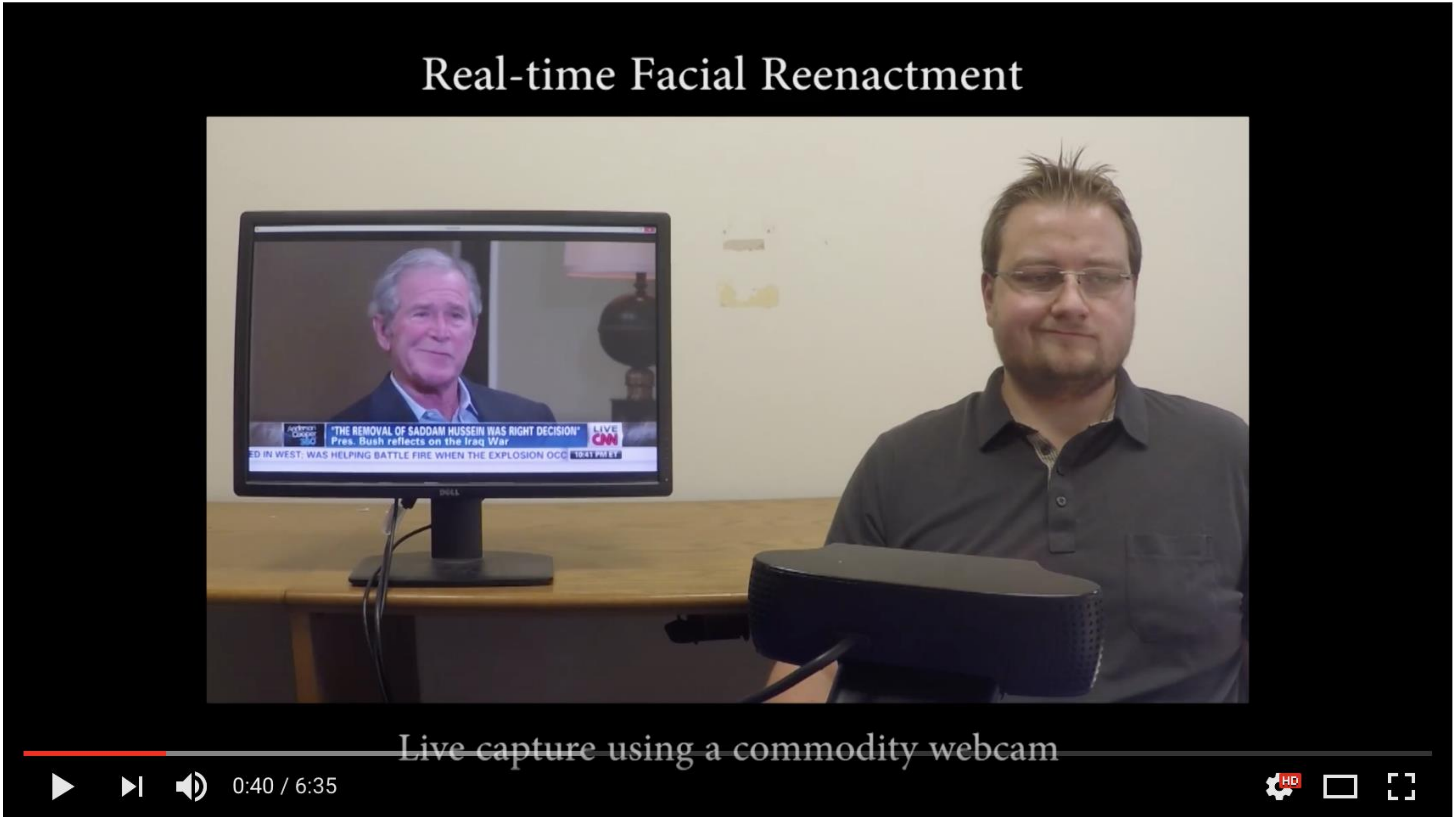


Figure 8. Real-life example of a backdoored stop sign near the authors' office. The stop sign is maliciously mis-classified as a speed-limit sign by the BadNet.

Example 3: AI forgery

Thies, Justus, et al. "**Face2face: Real-time face capture and reenactment of rgb videos.**" *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016.



Real-time Facial Reenactment

Live capture using a commodity webcam

"THE REMOVAL OF SADDAM HUSSEIN WAS RIGHT DECISION"
Pres. Bush reflects on the Iraq War
LIVE CNN
ED IN WEST: WAS HELPING BATTLE FIRE WHEN THE EXPLOSION OCC
10:41 PM ET
DELL
0:40 / 6:35
HD

Final thoughts

- AI crime is going to get worse
- AI is a form of information technology, so some forms of AI crime fit into the categories defined by the Convention on Cybercrime
- New names & forms of offences are probably needed
  - Is video forgery a form of document forgery or something new entirely?
  - What kind of offence is needed to deal with self-driving car killing a pedestrian?