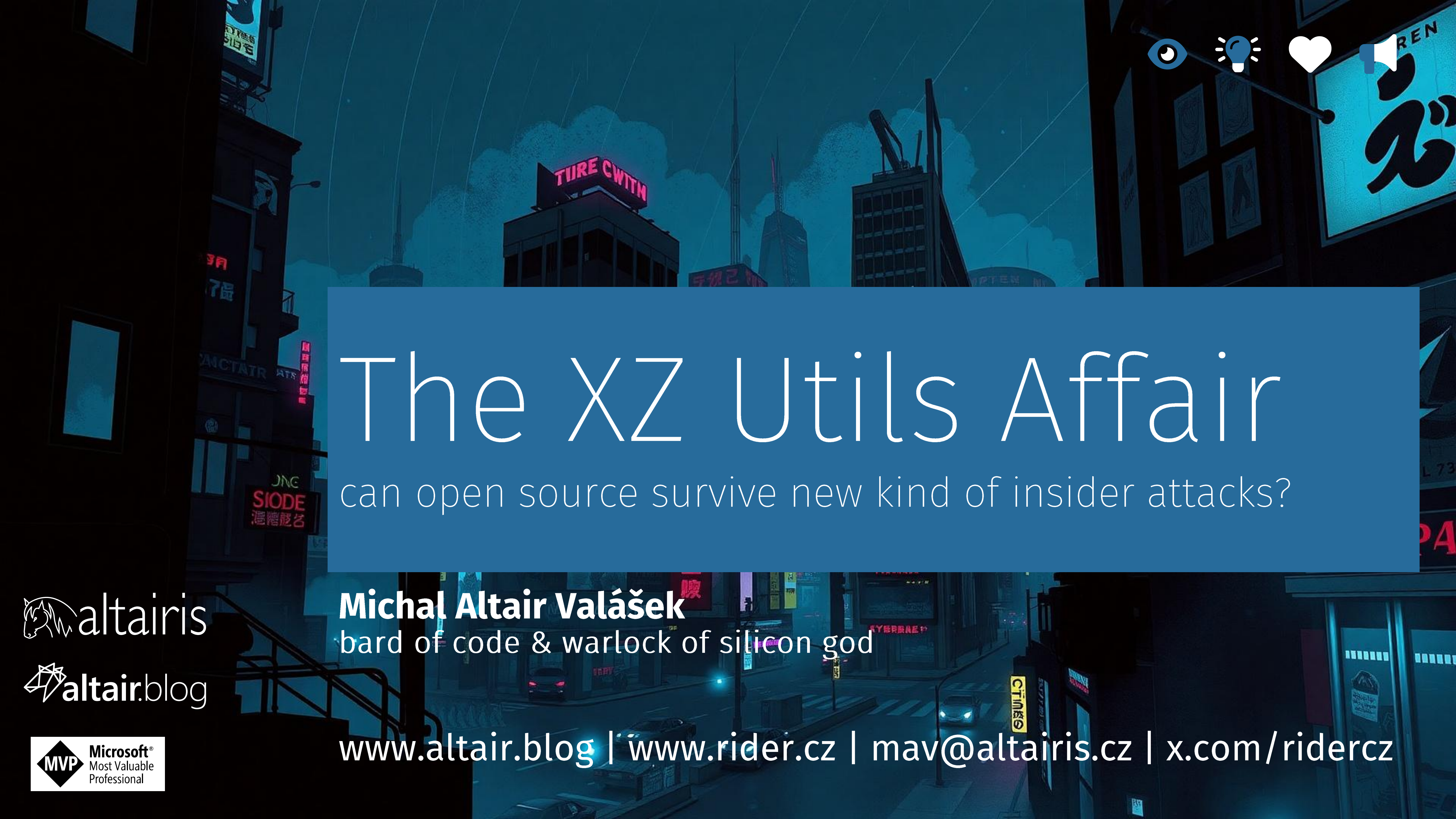


The XZ Utils Affair

can open source survive new kind of insider attacks?

Michal Altair Valášek
bard of code & warlock of silicon god

www.altair.blog | www.rider.cz | mav@altairis.cz | x.com/ridercz

altairis

altair.blog

MVP Microsoft® Most Valuable Professional

The main principle of free and open source software:

People are inherently good.

The main problem of free and open source software:

They are not.

XZ Utils

liblzma

What is XZ Utils

- XZ Utils is a set of free, open-source software tools and libraries for lossless data compression.
- It primarily supports the `.xz` file format, but also provides backward compatibility with the older `.lzma` format
- Principal author and maintainer is Lasse Collin.

Who uses XZ Utils

- Originally developed in 2009 for the now-defunct Tukaani Linux distro.
- Included in virtually all Linux distributions, including Fedora, Slackware, Ubuntu, and Debian.
- Available for (but not included in) Microsoft Windows.

The Incident Timeline

altairis

altair.blog

MVP Microsoft® Most Valuable Professional

Timeline

- **2005-2008: LZMA Utils**
  - Predecessor of current project.
- **2009: XZ Utils**
  - New project, new library
  - Added support for the new XZ file format.
- **Oct 29, 2021:**
  - Jia Tan sends a harmless patch to the `xz-devel` mailing list, adding an `.editorconfig` file.
- **Nov 29, 2021:**
  - Jia Tan submits another innocuous patch fixing a reproducible build issue.

Timeline

- **Feb 7, 2022:**
  - First commit by Jia Tan is merged into the XZ repository.
- **Apr–Jun, 2022:**
  - Multiple patches submitted by Jia Tan. New personas like “Jigar Kumar” and “Dennis Ens” appear, pressuring the original maintainer, Lasse Collin, to hand over control.
- **Jun 10, 2022:**
  - Jia Tan’s commit is merged with full authorship attribution.
- **Jun–Dec, 2022:**
  - Jia Tan becomes a regular contributor and gains trust within the project.

Timeline

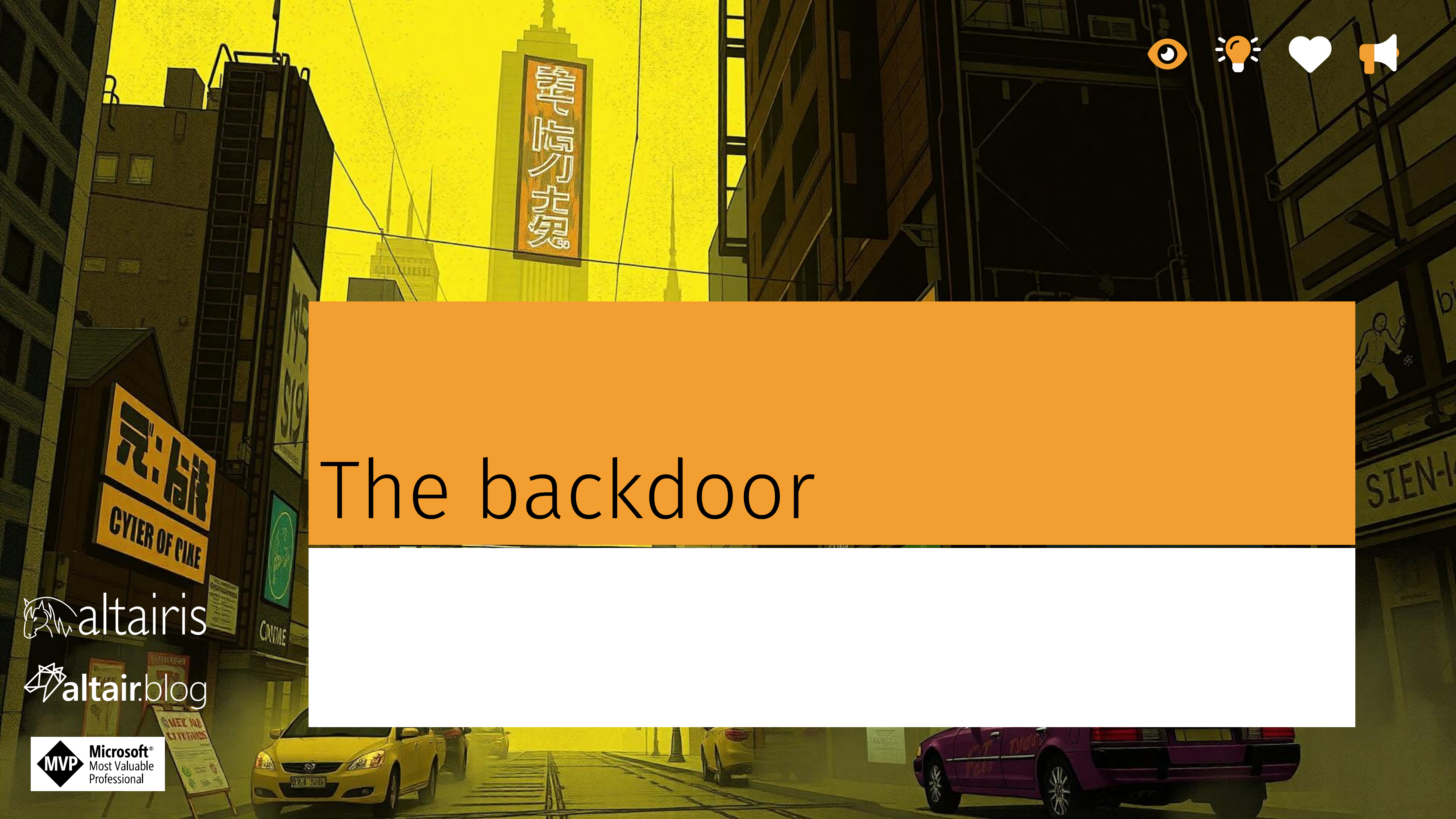
- **2023:**
  - Jia Tan continues contributing and gradually gains maintainer access.
- **Feb 23, 2024**
  - Jia Tan commits some binary data files, seemingly used for automated testing.
  - These files contain malicious payload.
- **Feb 24, 2024**
  - Jia Tan publishes new release 5.6.0.

Timeline

- **Feb 28, 2024:**
  - Matteo Croce proposes change in systemd, which would mitigate the attack.
  - This is pure coincidence, however, the backdoor remains undiscovered.
- **Mar 9, 2024:**
  - Jia Tan publishes backdoored version 5.6.1.
- **Mar 27, 2024:**
  - Version 5.6.1 was added into Debian.
  - Hans Jansen (likely another sockpuppet) requests addition to Ubuntu.

Timeline

- **Mar 28, 2024:**
  - Anders Freund, PostgreSQL developer at Microsoft, notices the SSH login to his server is too slow (takes 800 ms instead of usual 300 ms).
  - He also notices high CPU usage and errors in memory debugger Valgrind.



The backdoor

altairis

altair.blog

MVP Microsoft® Most Valuable Professional

The Clever Backdoor

- The backdoor is not in the code itself.
- It‘s in the binary testing files:
  - They are weird by definition, as they are used for testing edge cases, corrupted data handling etc.
  - The backdoor is activated using the install script, not the library code.
- The `liblzma` library is used by `libsystemd`, used for running system daemons.
- The result is vulnerability added to OpenSSH: remote code execution.

The Clever Backdoor

- The backdoor was not caught by any formal mechanism, such as code review etc.
- It was a random work of single non-security nerd, who was curious about a half-second delay during logon.

Wide Angle

What this means for FOSS?

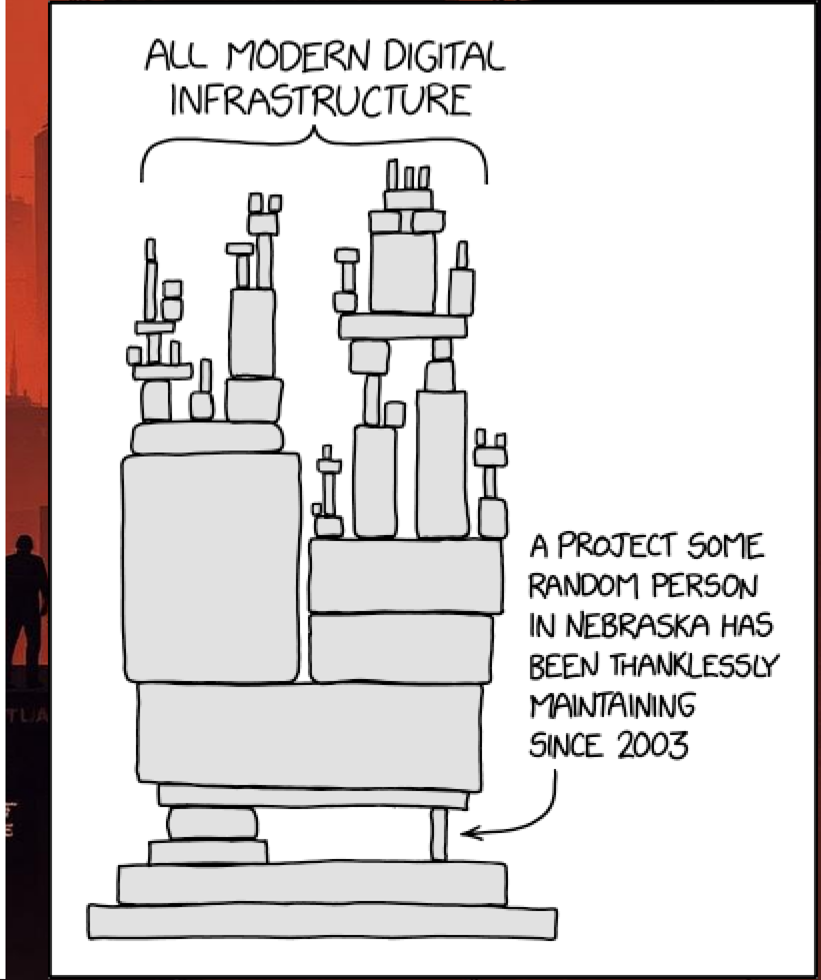
altairis

altair.blog

MVP Microsoft® Most Valuable Professional

Rule, not an exception

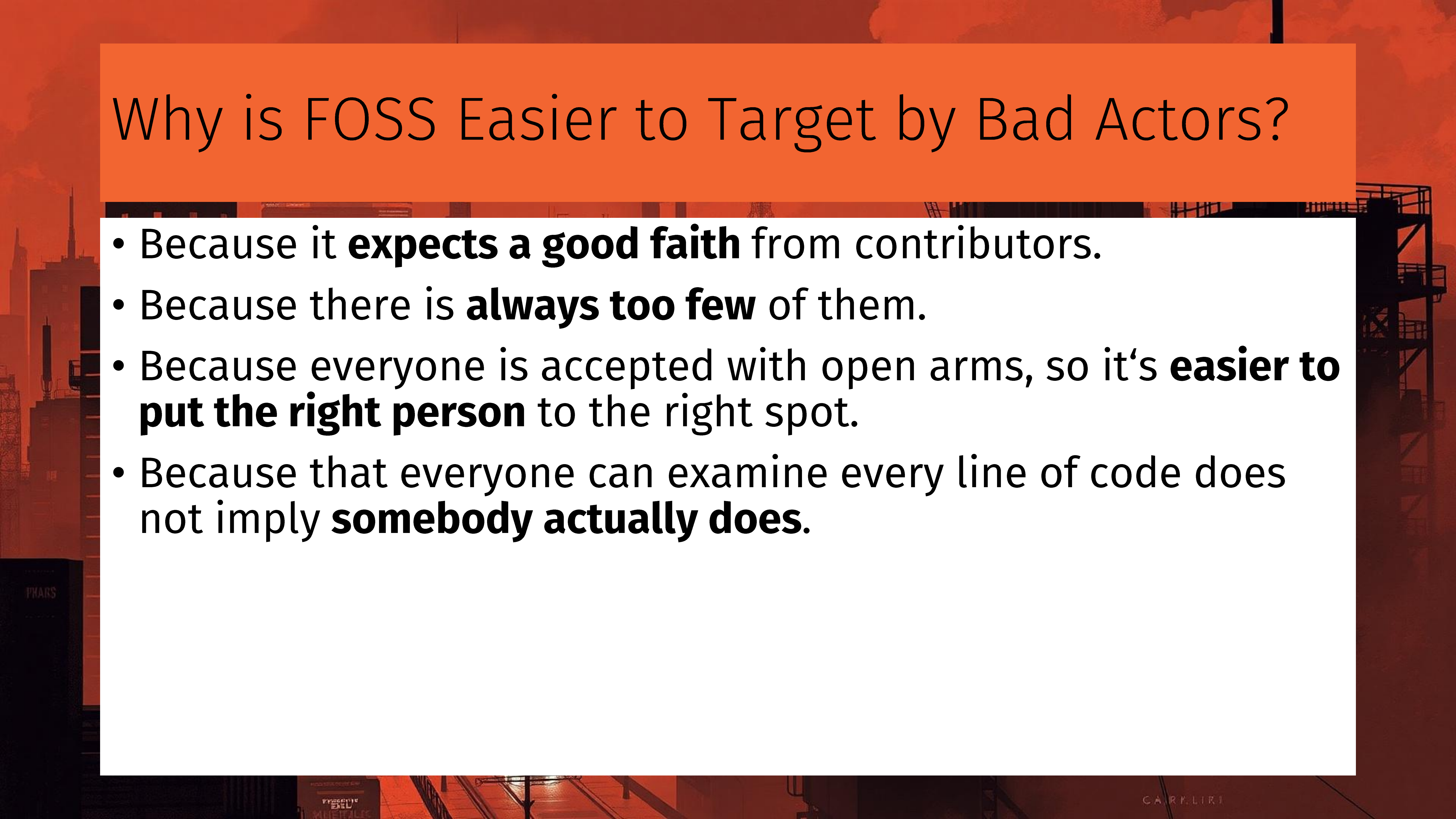
- UNIX is a bunch of random simple utils „doing one thing well“.
- Linux (and other FOSS projects) adopted this philosophy.
- The libraries are maintained by small numbers of developers, often just by one regular maintainer in their spare time.


ALL MODERN DIGITAL INFRASTRUCTURE
A PROJECT SOME RANDOM PERSON IN NEBRASKA HAS BEEN THANKLESSLY MAINTAINING SINCE 2003

https://xkcd.com/2347/ by Randall Munroe, CC BY-NC 2.5

Rule, not an exception

- **Bash**
  - Shell interface.
  - Used by virtually every Unix-like and Linux derivatives.
  - Maintained since around 1993 by Chet Ramsey from Ohio, US.
- **Dnsmasq**
  - DNS and DHCP server.
  - Used by most Unix distributions and small routers, Wi-Fi APs etc.
  - Maintained by Simon Kelley, Keswick, UK.
- **core-js**
  - Browser polyfill library, used by 75% of the top 1000 web sites.
  - Maintained by Denish Pushkarev, Russia.
  - Hit hard by sanctions, imprisoned for 18 months in Russia.
- And many more projects, run by single or just a few devs

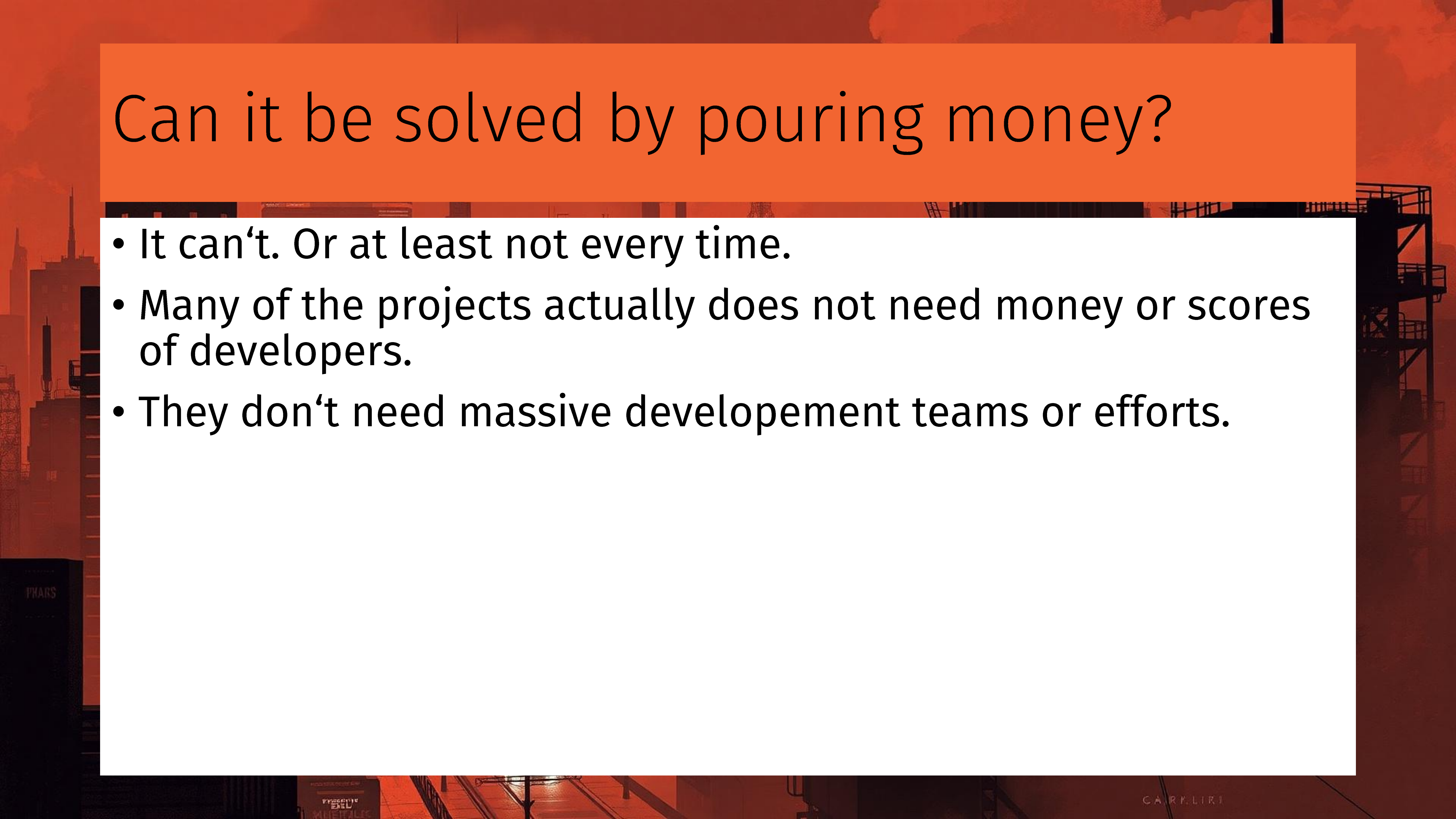


Why is FOSS Easier to Target by Bad Actors?

- Because it **expects a good faith** from contributors.
- Because there is **always too few** of them.
- Because everyone is accepted with open arms, so it‘s **easier to put the right person** to the right spot.
- Because that everyone can examine every line of code does not imply **somebody actually does**.

Can it be solved by pouring money?

- It can‘t. Or at least not every time.
- Many of the projects actually does not need money or scores of developers.
- They don‘t need massive developement teams or efforts.



Who is Jia Tan?

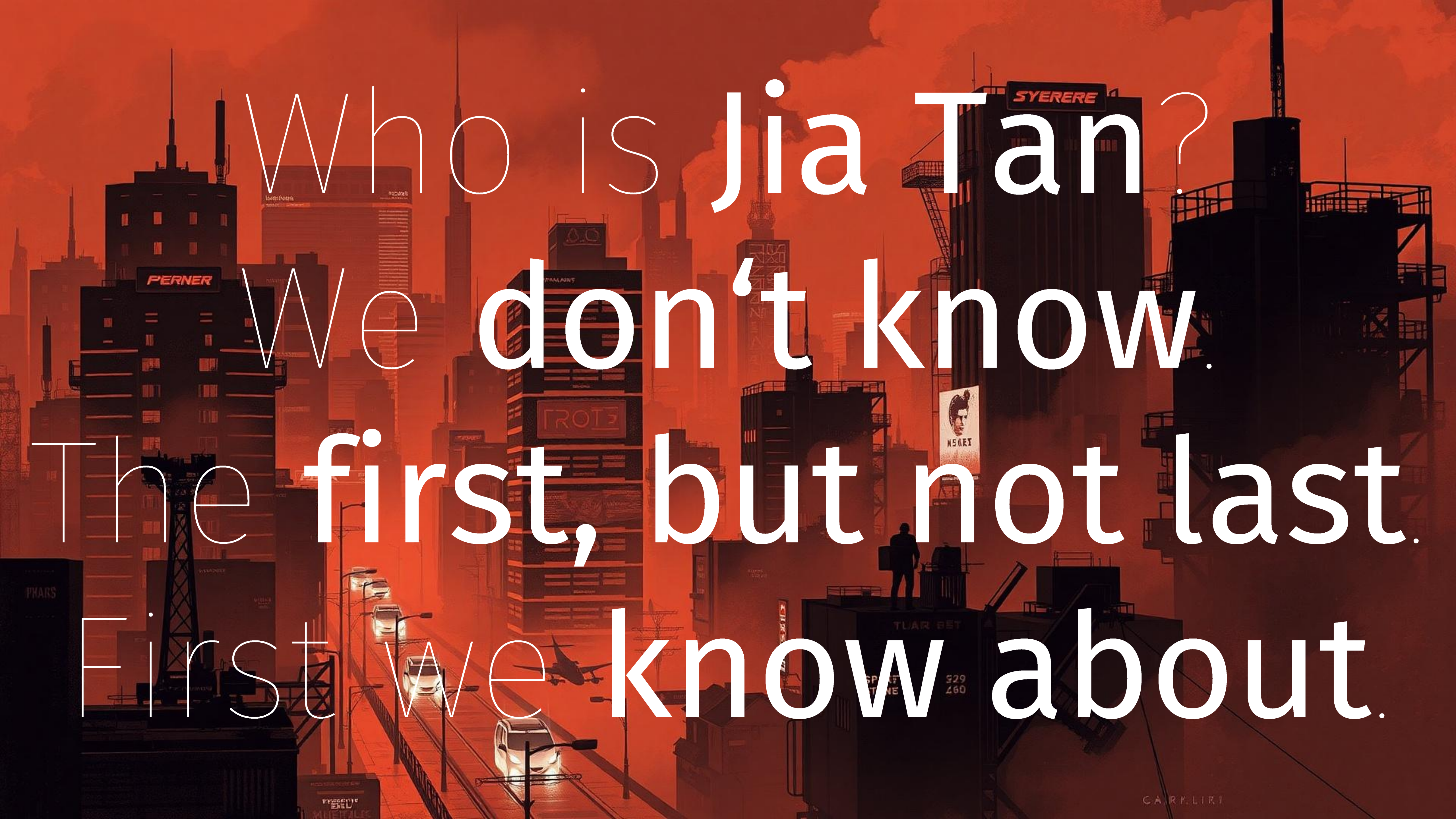
Who is Jia Tan?

We don‘t know.

Who is **Jia Tan**?

We **don't know**.

The **first, but not last**.



Who is Jia Tan?

We don‘t know.

The first, but not last.

First we know about.

altairis

altair.blog

mav@altairis.cz

facebook.com/**rider.cz**
github.com/**ridercz**
x.com/**ridercz**
linkedin.com/in/**ridercz**
youtube.com/**ztechcz**

www.**altairis.cz**
www.**rider.cz**
www.**altair.blog**

MVP Microsoft® Most Valuable Professional