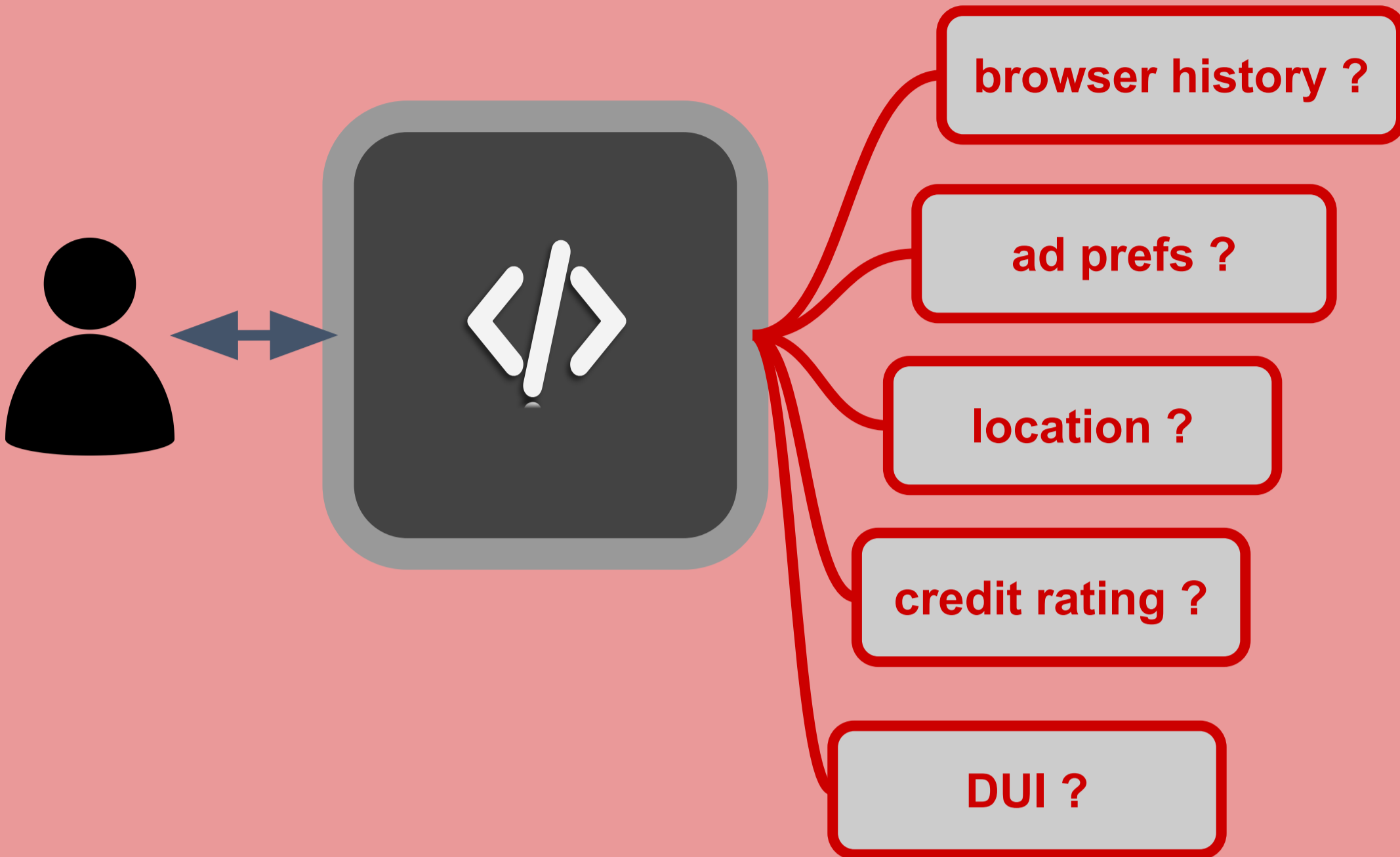


Provenance Injection for Privacy

Jens Dietrich, Tim White – Victoria University Of Wellington, Sam Shankland – University of Canterbury



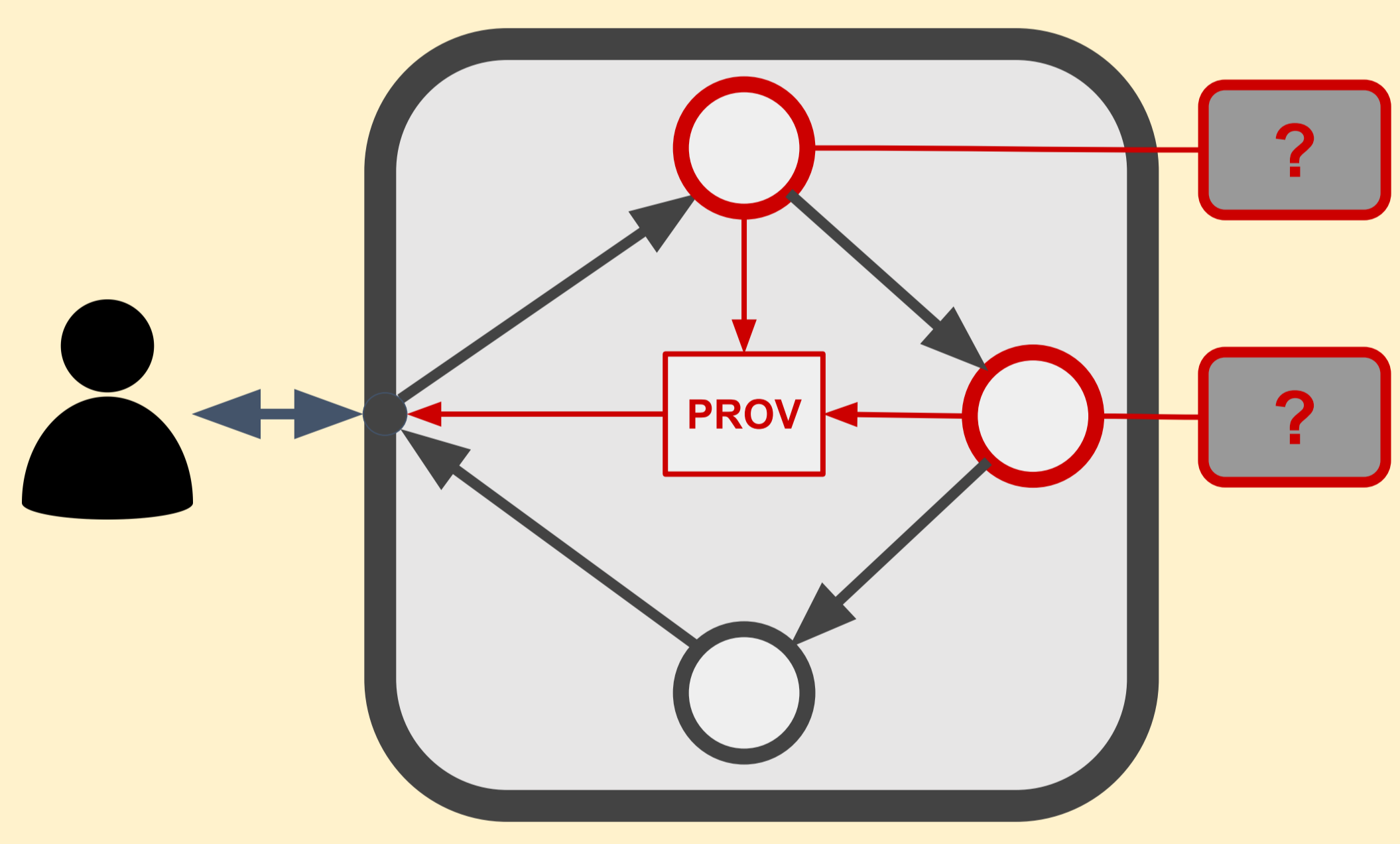
The Problem

Modern computing is done in a black box fashion – users have no or little knowledge how their data is being processed, and whether privacy is preserved.

Adding transparency manually is considered complex and prohibitively expensive.

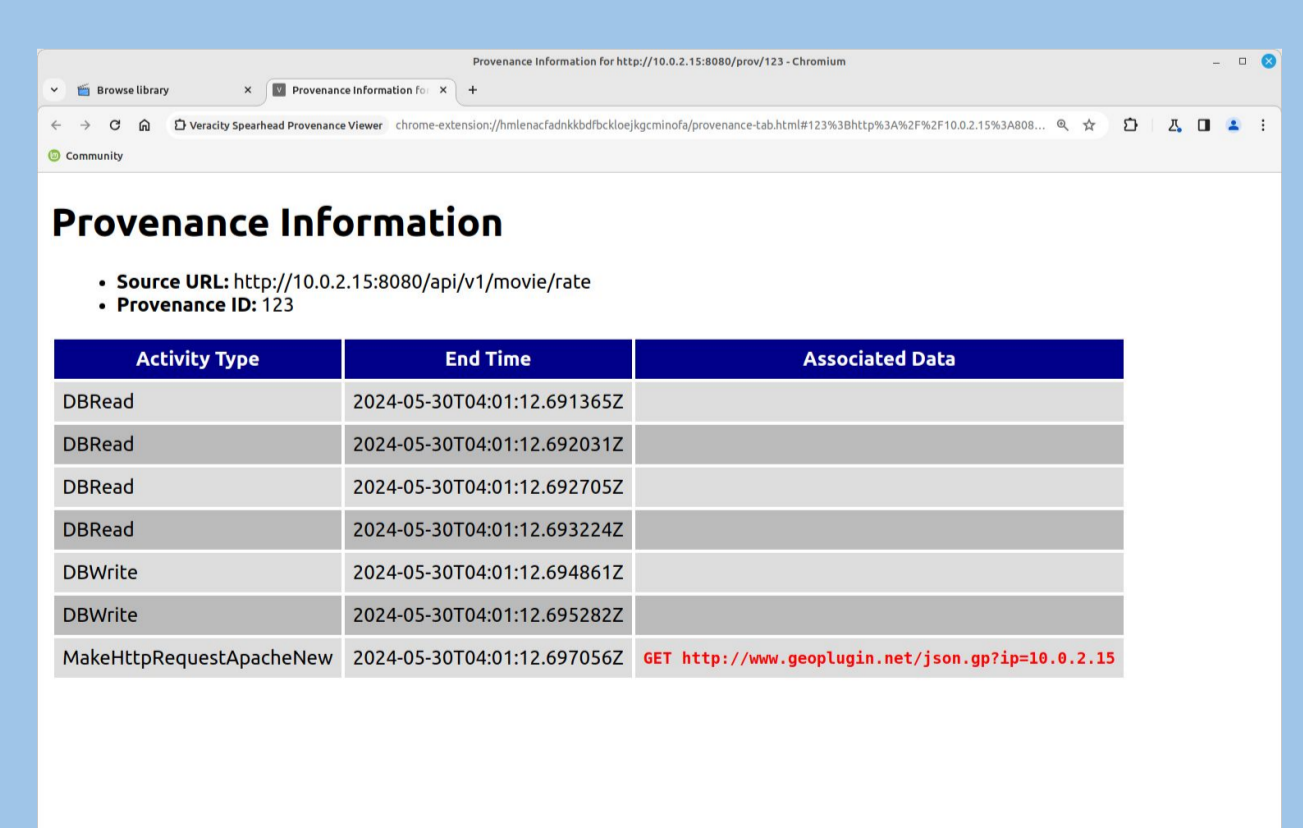
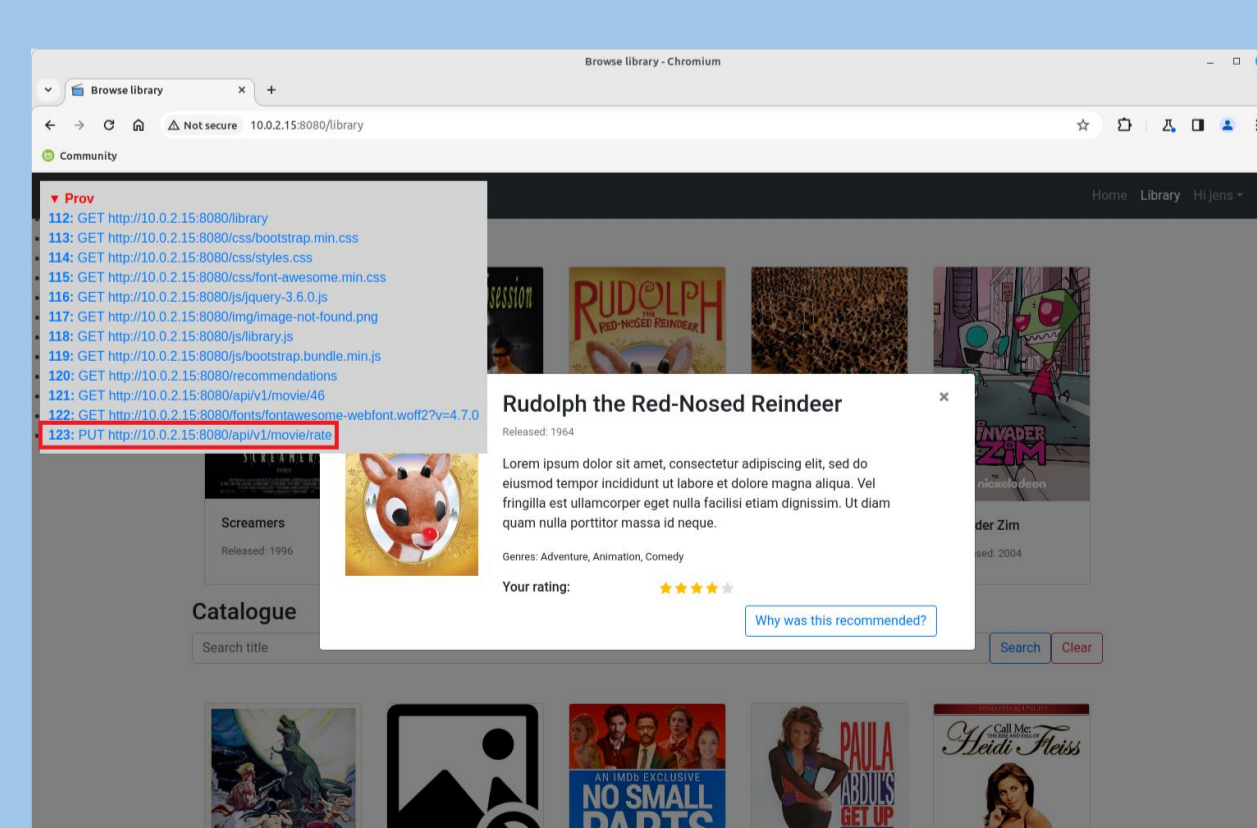
Our Approach: Use Bytecode Instrumentation to Add Provenance Automatically

- Compiled applications are instrumented to insert probes
- This does **not require code changes**, but can be achieved with small configuration changes
- At runtime, those probes **collect information** about actions persisting data, outgoing network calls, etc
- This information is encoded using W3C's **PROV-DM** standard
- This information is then attached to service responses using a dedicated **provenance HTTP header**, maintaining **compatibility with existing clients**
- Clients (like web browsers) can process this information, for instance, by **creating alerts for users**



Evaluation

We implemented a prototype movie rating application that uses the user's IP address to perform geolocation **on the server side** when the user rates a movie. Using our instrumentation, and a Chrome browser extension we developed for viewing provenance data, the user can discover the server-side geolocation HTTP request.



1. Clicking a star to rate the movie adds an entry to the list of URLs with provenance information
2. Clicking that URL shows server-side provenance data, including the outgoing geolocation HTTP request

For more information please email jens.dietrich@vuw.ac.nz.