

The problem of the web pages with lists of vendors without right context

Some banner data queries result into the web pages having lists of vendors/devices. These lists make huge amount of transactions when passed through ARE, while it is obvious that the query banner data is irrelevant to most of these transactions. This is one of the major problems which makes the results of ARE less accurate. There is a need to detect such web pages and prevent them from being a part of transactions.

The problem is illustrated below with an example:

For banner: Welcome to burda praha tandberg codec sw F6

Refined Query: burda praha tandberg codec sw f6

One of the web pages: https://www.clarybusinessmachines.com/Tandberg-Codec-C40-12x-Integrator-Package_11037.htm

This web-page shows different products with vendor names in “Similar Products” and “Related Products” section.

One other web-page: <https://www.kb.cert.org/vuls/bypublished/desc/>

This web-page is a database of vulnerabilities having a column with names of vendors and devices.