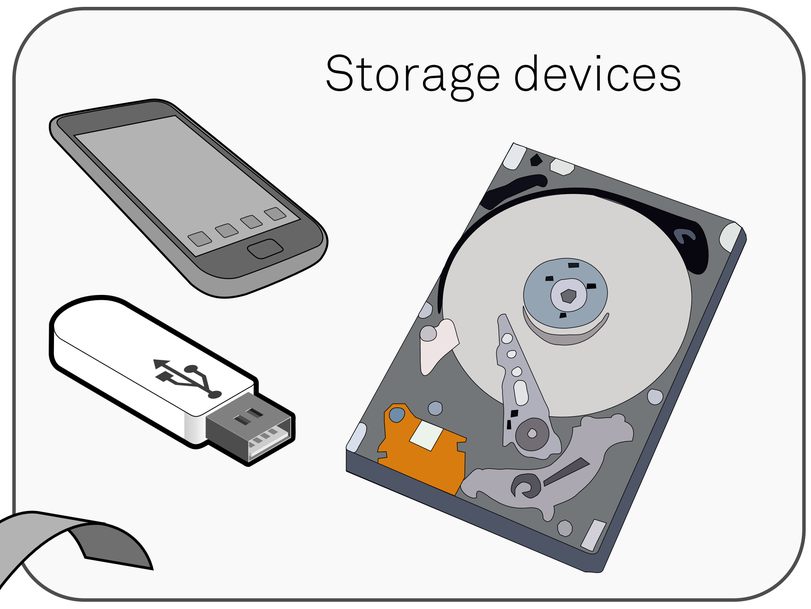
Nederlands Forensisch Instituut
Ministerie van Veiligheid en Justitie

Project Sherlock

Generating better leads for criminal investigations

Digital storage devices such as harddisks, mobile phones, and USB drives are increasingly common in criminal cases. Extracting viable leads from such a device poses various technological challenges. The Netherlands eScience Center is helping the Netherlands Forensic Institute face these challenges by extending their toolchain with a number of eScience technologies.

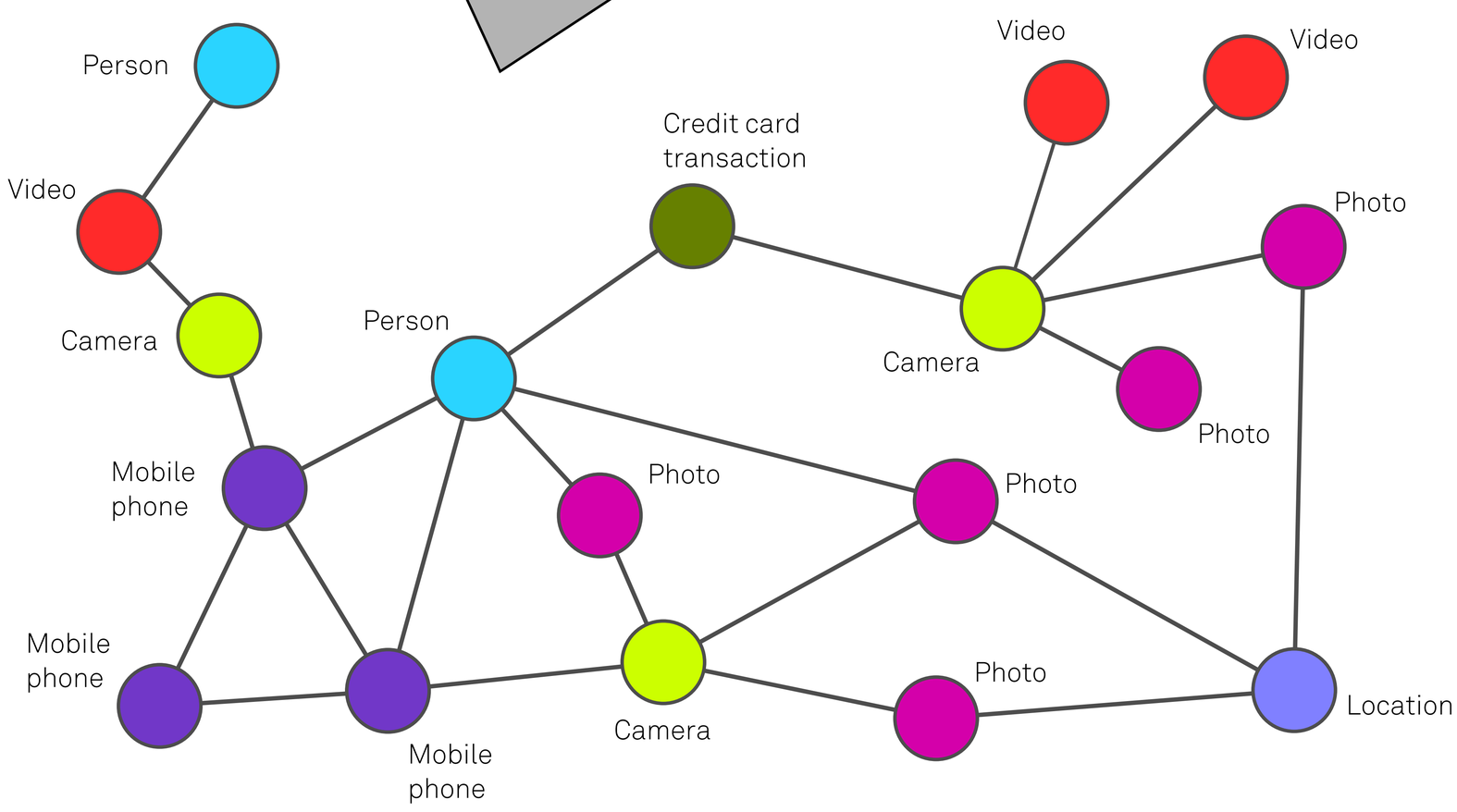


Big Data Analytics

Volume: Criminal cases can involve many terabytes of data.

Velocity: The search for clues is often time-bound by circumstances as well as by law.

Variety: Storage devices can contain a great variety of file types, and the layout of individual files is not known beforehand.



Technologies

distributed computing

The task of analyzing a storage device breaks down into many smaller tasks, which need to be distributed over the available hardware resources.

image analysis

Computer vision methods allow for automatically annotating photos and videos with metadata such as the names of people, places, and weapons.

natural language processing

More and better clues are produced more quickly using techniques that extract meaning from textual data.

hardware accelerators

Photos and videos can be analyzed much more quickly using dedicated hardware such as GPUs.

data management

A dedicated data management solution makes the data searchable, and enables feeding selections of data to visualization tools.

visualization

Interactive visualizations provide insight into the course of events.