# A data-visiting infrastructure for providing access to preserved databases that cannot be shared or made publicly accessible

Martin Weise, Andreas Rauber
TU Wien, Austria

FAKULTÄT FÜR INFORMATIK
Faculty of Informatics

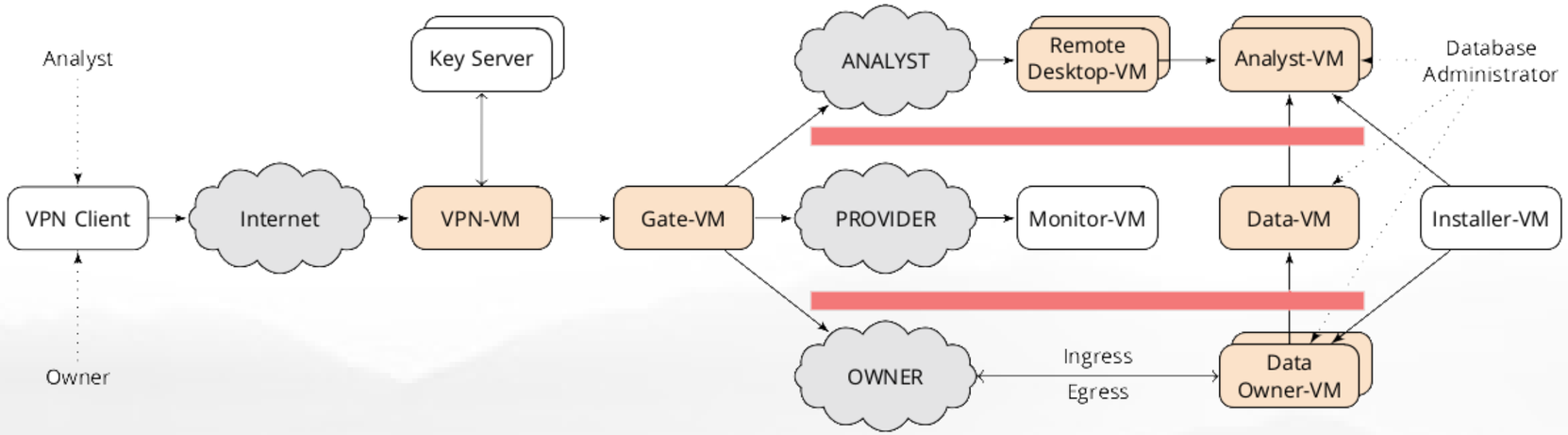
Contact: Prof. Andreas Rauber (*andreas.rauber@tuwien.ac.at*)
Report: *https://doi.org/10.5281/zenodo.4632903*
Source: *https://gitlab.tuwien.ac.at/martin.weise/ossdip*

## Motivation

- Preserved databases (DBs) frequently contain sensitive information.
- Such DBs cannot be shared with others.
- *Physical visiting* requires on-site attendence.
- Obtaining clearance is a *complex legal process*.
- Release only anonymized *aggregated* data with limited granularity.

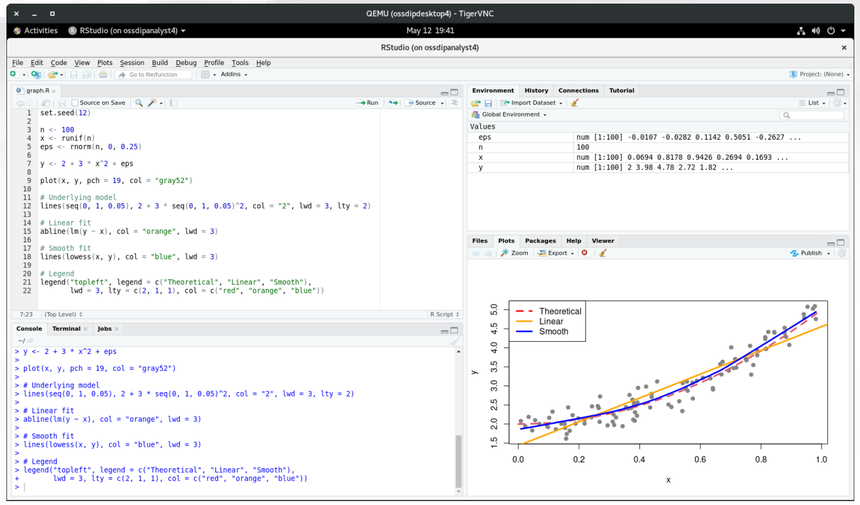
## OSSDIP Infrastructure

- Data visiting instead of data sharing.
- Allow users to work with the data (non-aggregated).
- Ensure that the archive remains in complete control of the data (no record is exfiltrated).
- Adapts concepts and infrastructure set-ups from the medical domain to the needs of archival institutions.
- Specific data extracts are copied to dedicated Analyst-VMs and provide required tools to perform approved research tasks.
- Remote desktop fingerprinting, monitoring etc. offer configurable layers of security.




Overall System Architecture

Working with preserved sensitive data using RStudio




## Methods

- UK HDRA Trusted Research Environments (TRE).
- Based on the experience of operating DEXHELPP for almost ten years.
- Provide highly controlled and monitored data visiting services, without disseminating an actual copy.
- Allow third parties to visit the data.
- Open source reference implementation.
- Project homepage: *http://www.ifs.tuwien.ac.at/~andi/secure_data_infrastructure.html*

EOSC secretariat eu
EOSC-Life
dexhelpp
FUTURE OPERATIONS