

# ExDRa

## EXPLORATORY DATA ANALYSIS ON FEDERATED RAW DATA

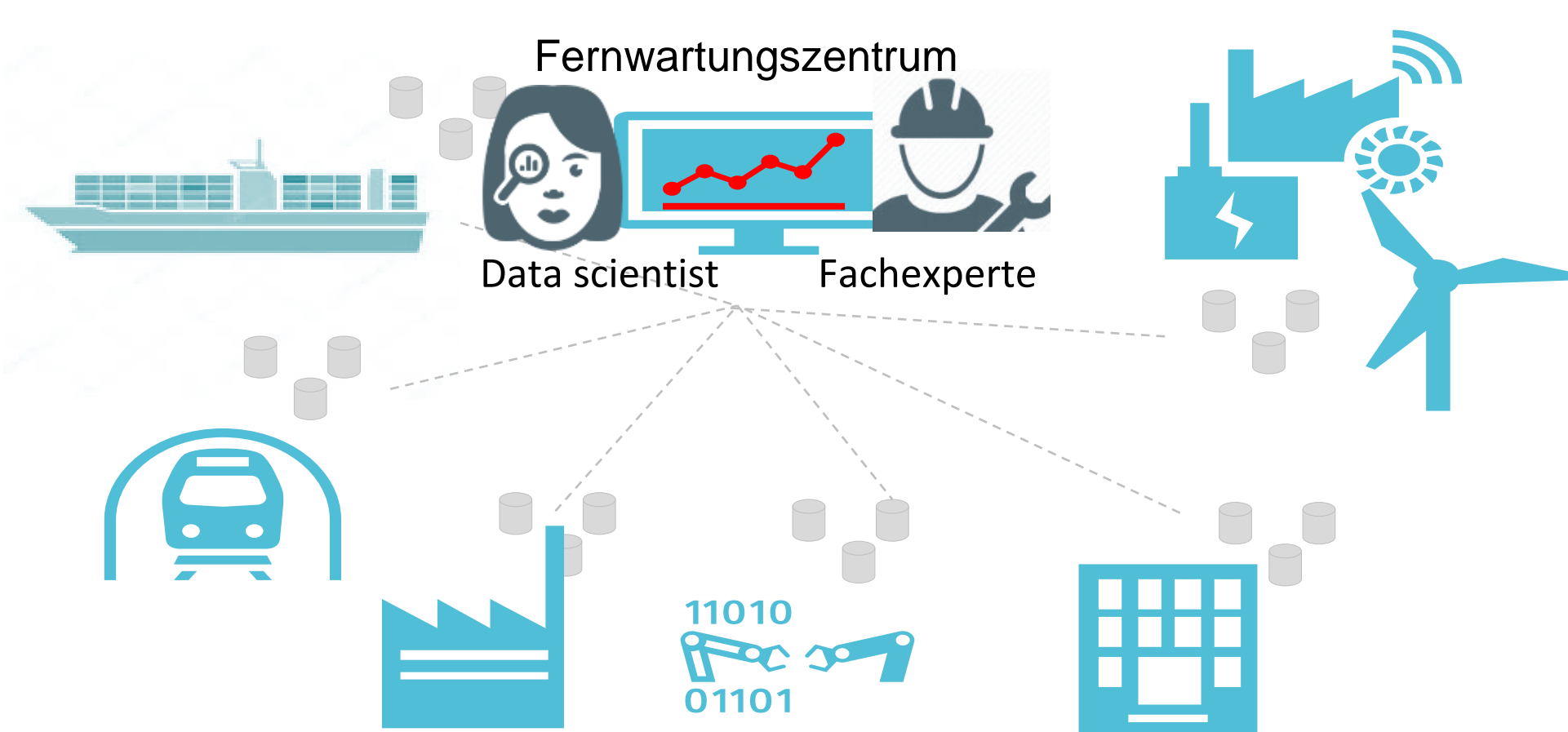
### KURZPORTRAIT

In ExDRa wird eine Systemsoftware entwickelt, die es dem Datenwissenschaftler stark erleichtert, explorative Datenanalysen wie beispielsweise Datenvorverarbeitung und das Training von Machine Learning Modellen einschließlich Neuronaler Netze auf geographisch verteilten Daten durchzuführen, ohne dass diese Daten an einem Ort zusammengeführt werden müssen. In der Praxis wird dieses Konzept mit Hilfe eines Demonstrators in Anwendungen der Prozessindustrie erprobt.

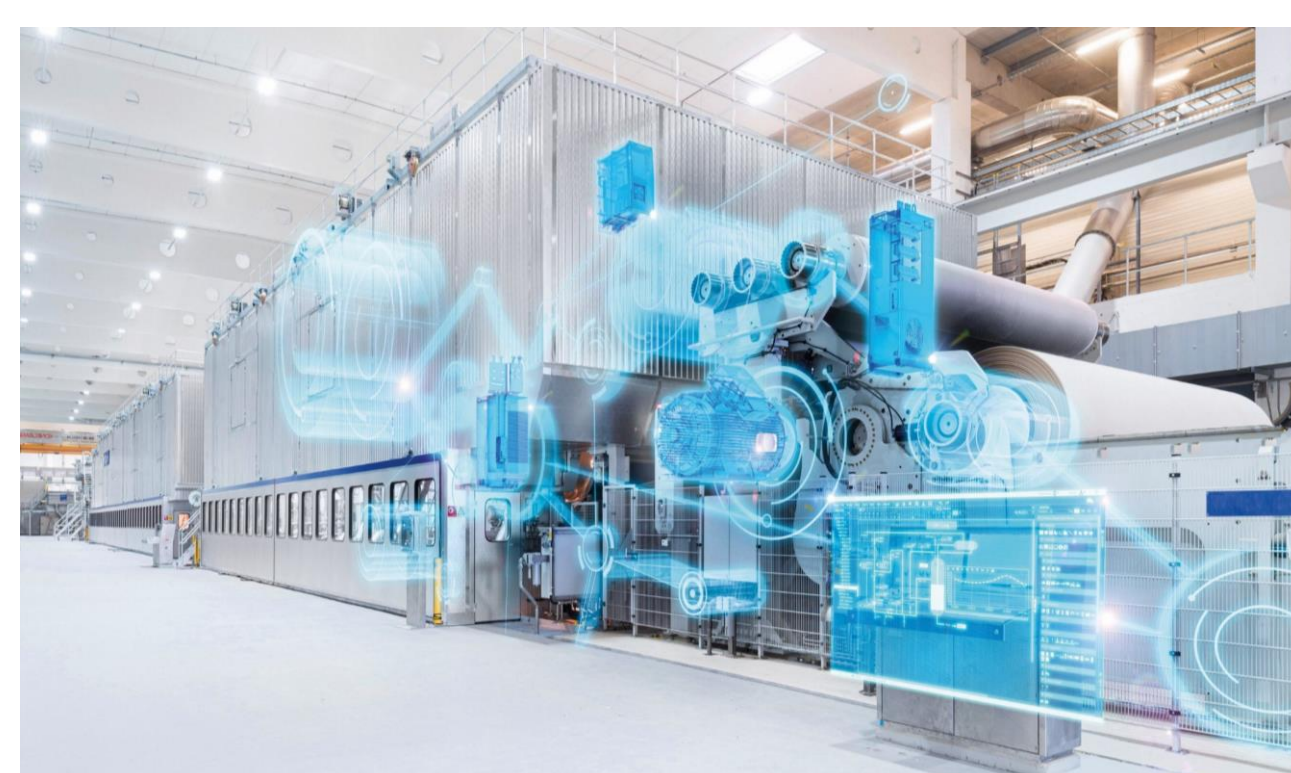
### LÖSUNGSANSATZ

- Erweiterung und Kombination der Systeme Apache SystemDS (<https://github.com/apache/systemds> TU Graz) für föderiertes Lernen und NebulaStream (<https://www.nebula.stream> TU Berlin / DFKI) für Verarbeitung geoverteilter Datenströme
- Hohe Effizienz durch Wiederverwendung von Zwischenergebnissen
- Systematische Modellverwaltung und Optimierung von Data-Science Experimenten

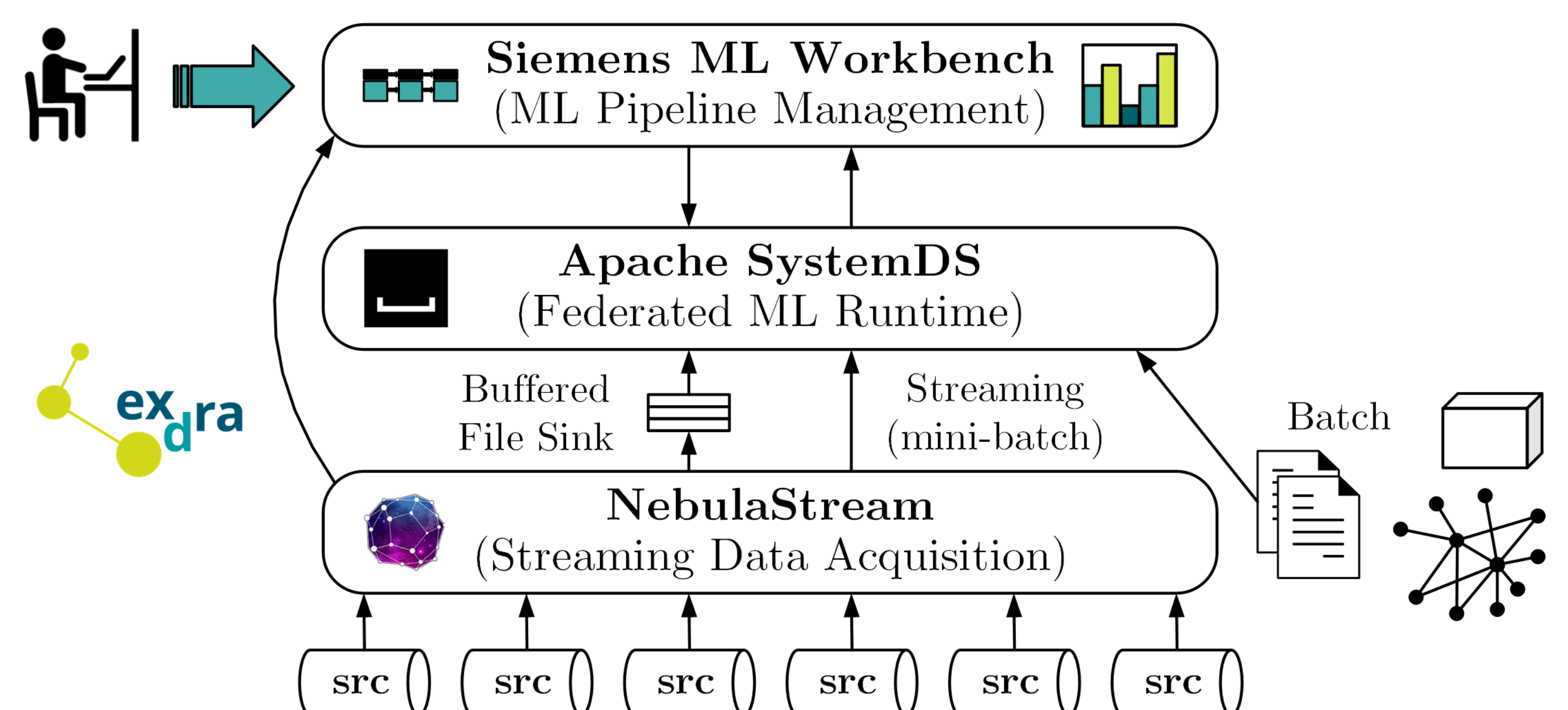
#### Breites Anwendungsspektrum



#### Beispielanwendung Qualitätsvorhersage in der Papierproduktion

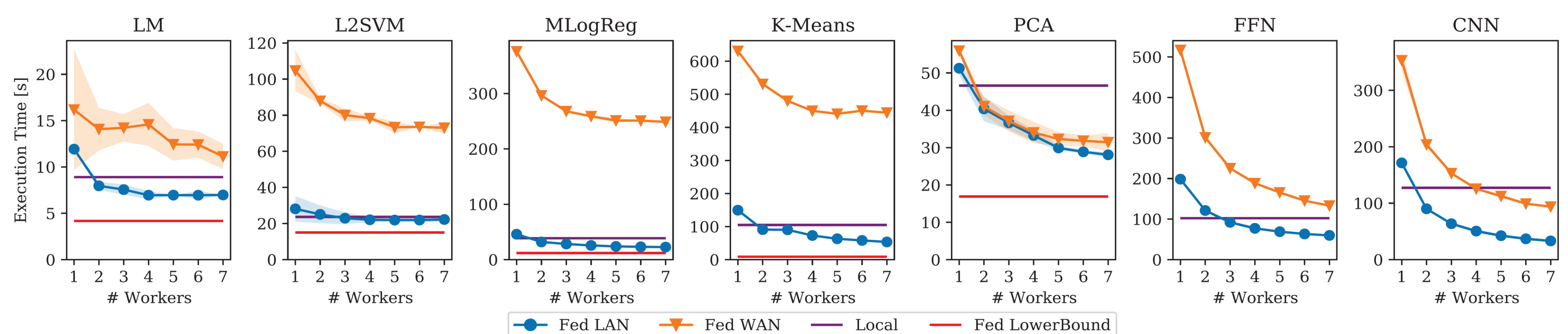


#### Systemarchitektur



#### Experimentelle Resultate

(Lokale Baseline, LAN ... local area network, WAN .. wide area network Kopenhagen-Graz)



### KONSORTIUM

Siemens AG; Berlin/Erlangen (Dr. C. Neubauer)  
DFKI GmbH; Berlin (Dr. S. Zeuch)  
Technische Universität Berlin (Prof. Dr. V. Markl)  
Technische Universität Graz (Prof. Dr. M. Böhm)

### WEITERE PARTNER

Bundesverband mittelständische Wirtschaft  
BVMW (M. Dönges)

### WEBSEITE

<http://www.exdra.de>