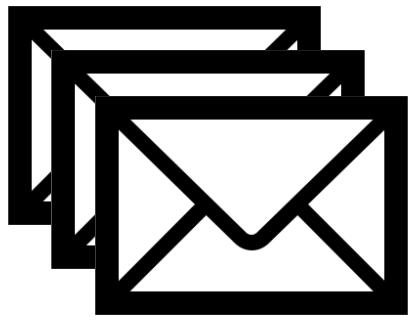


# Der Leuchtturm im Datennebel

Exploration großer Dokumentensammlungen



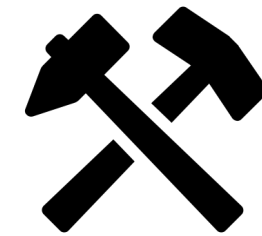
## Motivation

Fast alle Geschäftsprozesse werden inzwischen digital abgewickelt. Die daraus resultierenden Datenmengen machen manuelle Untersuchungen nach Auffälligkeiten nahezu unmöglich. Analysten und Journalisten haben daher dringenden Bedarf an einem Explorationstool, das sie bei ihren Ermittlungen unterstützt.



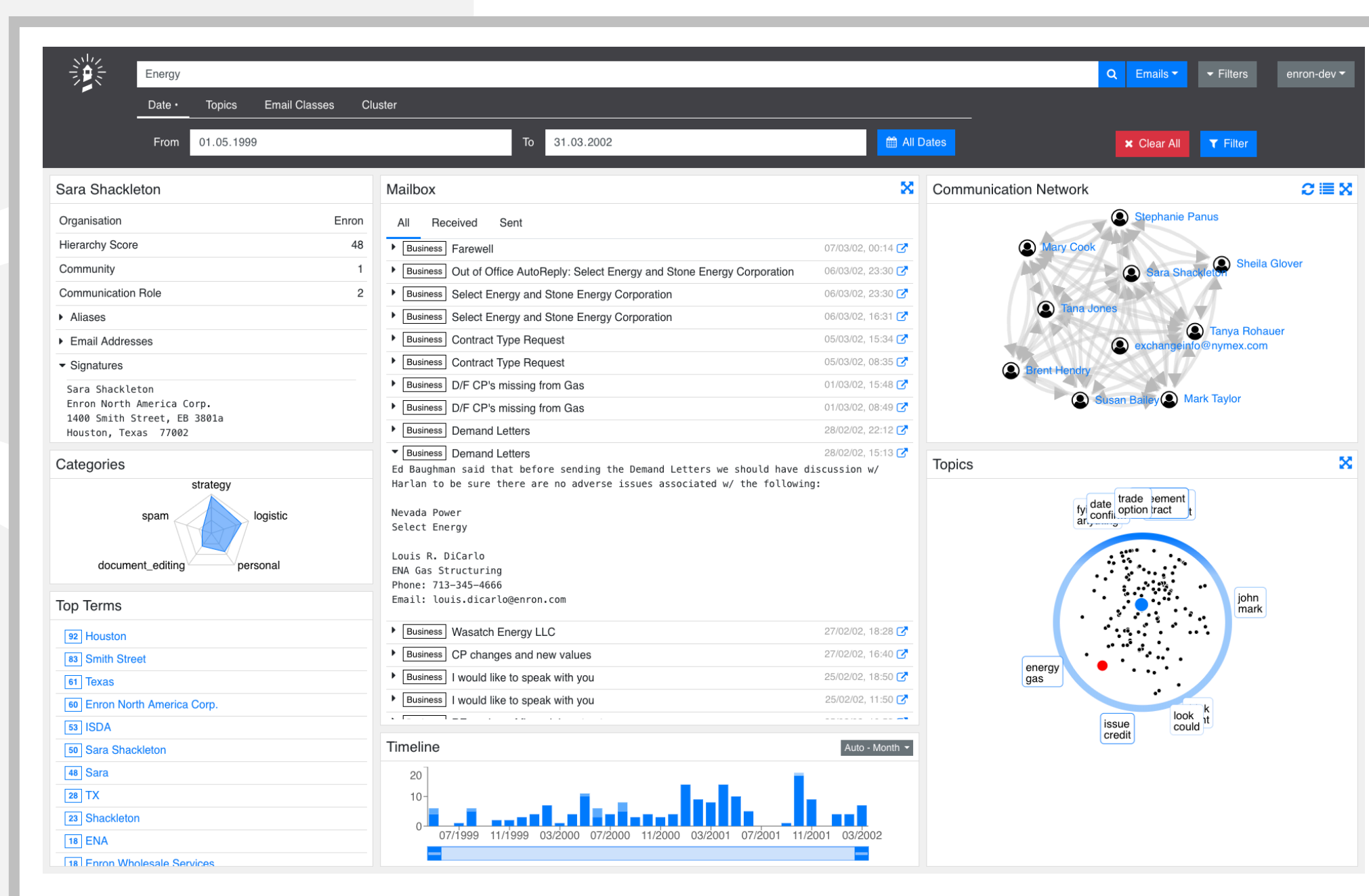
## Netzwerkanalyse

Unser System normalisiert semi-strukturierte E-Mail-Daten und erkennt, wer wann mit wem kommuniziert hat. Dabei wird ein Netzwerkgraph extrahiert, der auch auf Kommunikationsmuster und Hierarchien untersucht wird.



## Text Mining

Automatisiert erkennt unser System, über welche Themen in E-Mails gesprochen wird, gruppiert diese nach Ähnlichkeit und extrahiert häufig verwendete Phrasen. Außerdem stellen wir fest, in welchem Kontext die E-Mails versendet wurden.



## Explorationstool

Unser Tool eignet sich für die einfache Exploration großer E-Mail-Datensätze. Es bietet dem Nutzer interessante Einstiegspunkte und einen Überblick über den gesamten Datensatz. Zudem kann der Fokus auf relevante Aspekte gelegt werden und Korrelationen im Netzwerk werden sichtbar gemacht.

Projektpartner

**COMMERZBANK**

Projektbeteiligte

**Fachgebiet Informationssysteme**

Externe Partner: Commerzbank

Teilnehmer: Jakob Edding, Moritz Hartmann, Jonas Hering, Dennis Kipping, Hendrik Schmidt, Nico Scordialo, Alexander Zenner

Betreuung: Tim Repke

Leitung: Dr. Ralf Krestel



Digital Engineering · Universität Potsdam