Vilnius University

An Early Warning System for Fintech Supervision

Alfredo Hernandez Sanchez, Ph.D.

VU IIRPS

April 24, 2025

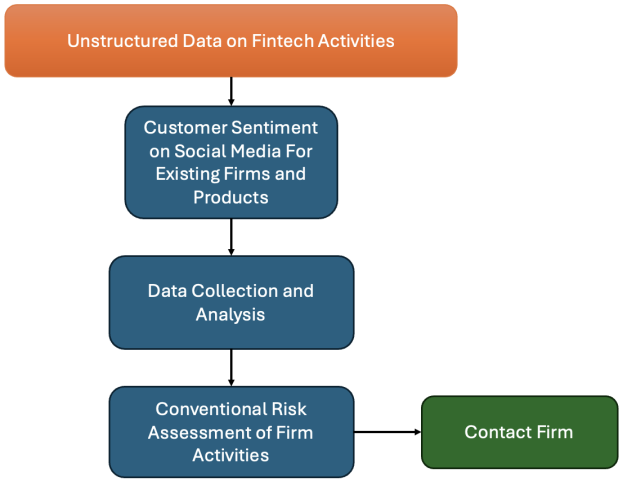
What is this project about?

- A prototype for an early warning system designed to support **fintech innovation hubs**.
- Uses text analysis from news, databases, and social media to identify new firms and products.
- Assigns structured indicators: firm name, product category, risk level, sentiment, and jurisdictional relevance.
- Flags entries that may warrant **proactive supervisory contact**.

Why Is It Needed?

- Current supervision models are often **reactive**, relying on firms to reach out or apply for sandbox access.
- Many early-stage innovations fly under the radar, increasing risks around unregulated activity.
- Supervisors lack a granular, live view of emerging market dynamics.
- Some firms may require **regulatory guidance** before they know they need it.

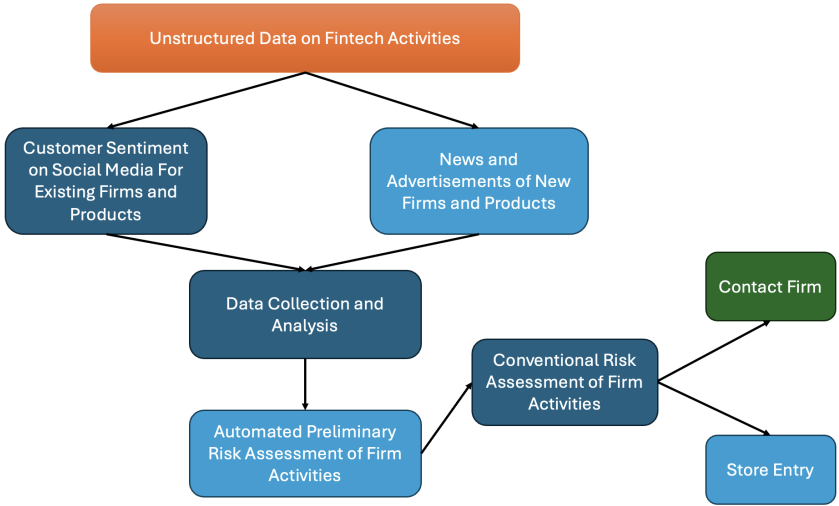
Current Process



How Does It Work?

- Gathers short-form text from **multiple sources**: news, Crunchbase-like entries, and scraped content.
- Extracts main entities and keywords, then classifies sentiment and product type.
- Flags entries with negative sentiment or high-risk language.
- All data is stored in a structured format to power a future **dashboard** for regulatory teams.

Proposed Solution



Keywords and Labels

```
# Define fintech labels
fintech_labels = [
    "other",
    "payment systems",
    "peer to peer lending",
    "crowdfunding",
    "decentralized finance (defi)",
    "digital assets",
    "cryptocurrencies",
    "traditional finance"
]
```

```
# Define risk keywords
risk_keys = [
    "crypto",
    "KYC",
    "unregulated",
    "yield",
    "token",
    "startup"
]
```

Classification and Identification

Vilnius University

```python
# Apply zero-shot classification for fintech category
df["fintech_category"], df["category_confidence"] = zip(*df["text"].apply(classify_fintech_category))

# Apply sentiment analysis
df["sentiment_label"], df["sentiment_score"] = zip(*df["text"].apply(classify_sentiment))

# Create jurisdiction indicator
df["in_jurisdiction"] = df.apply(
    lambda row: int(
        any(re.search(r"lithua|vilniu|kaunas|klaipeda", ent[0], re.IGNORECASE) for ent in row["entities"])
        or bool(re.search(r"lithua|vilniu|kaunas|klaipeda", row["text"], re.IGNORECASE))
    ), axis=1
)

# Create firm contact flag (updated logic: OR condition)
df["contact_firm"] = df.apply(lambda row: int(
    (row["sentiment_label"] == "negative" and row["sentiment_score"] > 0.75)
    or row["risk_level"] in ["medium", "high"]), axis=1)
```

What Are the Benefits?

- Shifts supervision **from passive to anticipatory**.
- Builds a living map of fintech activity to support ethical, targeted oversight.
- Flags entries with negative sentiment or high-risk language.
- Helps supervisors engage early and constructively, avoiding unnecessary enforcement.
- Supports **safer innovation** and a more transparent ecosystem for all stakeholders.

Prototype Link

QR code linking to the prototype early warning system on Google Colab.

Thank You!