Technische Universität München

sebis

# Transfer Learning for Name Entity Linking with Deep Learning

28.05.2018

Robin Otto, Master Thesis – Final Presentation

Software Engineering für betriebliche Informationssysteme (sebis)
Fakultät für Informatik
Technische Universität München

wwwmatthes.in.tum.de

# Outline

sebis

**Motivation**

Research Questions

Research Approach

Related Work

Implementation

Evaluation

Conclusion

# Motivation

## Named Entity Linking (NEL) and Transfer Learning in the Legal Domain

sebis

### Legal Domain

- Many stakeholders
- Few applications

### Transfer Learning

- Scarcity of data
- Complicated task
- Available solutions

### Named-Entity Linking

- Legal documents unclear for non domain experts
- Stakeholders need to work with documents

# Outline

Motivation

**Research Questions**

Research Approach

Related Work

Implementation

Evaluation

Conclusion

# Research Questions

## Transfer Learning & Neural Network Comparison

sebis

### What kind of existing approach should be used for transfer learning?

- Analyze state of the art deep named-entity linking systems
- Comparison based on F1 Score and Accuracy

### Which technique of transfer learning suits best?

- Employ state of the art
- Put datasets in relation: size & similarity
- Categorize involved datasets into correct scenario

### Is the use of transfer learning with named-entity linking beneficial in the legal domain?

- Create baseline results
- Evaluate results by comparing transfer learning performance with original performance

# Outline

Motivation

Research Questions

**Research Approach**

Related Work

Implementation

Evaluation

Conclusion

# Approach

## Dataset Creation

**REGULATION (EU) 2017/1938 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL**

**of 25 October 2017**

**concerning measures to safeguard the security of gas supply and repealing Regulation (EU) No 994/2010**

**(Text with EEA relevance)**

THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION,

Having regard to the Treaty on the Functioning of the European Union, and in particular Article 194(2) thereof,

Having regard to the proposal from the European Commission,

After transmission of the draft legislative act to the national parliaments,

Having regard to the opinion of the European Economic and Social Committee (1),

After consulting the Committee of the Regions,

Acting in accordance with the ordinary legislative procedure (2),

Whereas:

(1) Natural gas (gas) remains an essential component of the energy supply of the Union. A large proportion of such gas is imported into the Union from third countries.

(2) A major disruption of gas supply can affect all Member States, the Union and Contracting Parties to the Treaty establishing the Energy Community, signed in Athens on 25 October 2005. It can also severely damage the Union economy and can have a major social impact, in particular on vulnerable groups of customers.

(3) This Regulation aims to ensure that all the necessary measures are taken to safeguard an uninterrupted supply of gas throughout the Union, in particular to protected customers in the event of difficult climatic conditions or disruptions of the gas supply. Those objectives should be achieved through the most cost-effective measures and in such a way that gas markets are not distorted.

(4) Union law, in particular Directive 2009/72/EC of the European Parliament and of the Council (3), Directive 2009/73/EC of the European Parliament and of the Council (4), Regulation (EC) No 713/2009 of the European Parliament and of the Council (5), Regulation (EC) No 714/2009 of the European Parliament and of the Council (6), Regulation (EC) No 715/2009 of the European Parliament and of the Council (7) and Regulation (EU)No 994/2010 of the European Parliament and of the Council (8), has already had a significant positive impact on the security of gas supply in the Union, both in terms of preparation and mitigation. Member States are better prepared to face a supply crisis now that they are required to establish preventive action plans and emergency plans, and they are better protected now that they have to meet a number of obligations regarding infrastructure capacity and gas supply. However, the Commission's report on the implementation of Regulation (EU) No 994/2010 of October 2014 highlighted areas in which improvements to that Regulation could further bolster the security of gas supply in the Union.

*Article 2*

**Definitions**

For the purposes of this Regulation, the following definitions apply:

(1) 'security' means security as defined in point 32 of Article 2 of Directive 2009/73/EC;

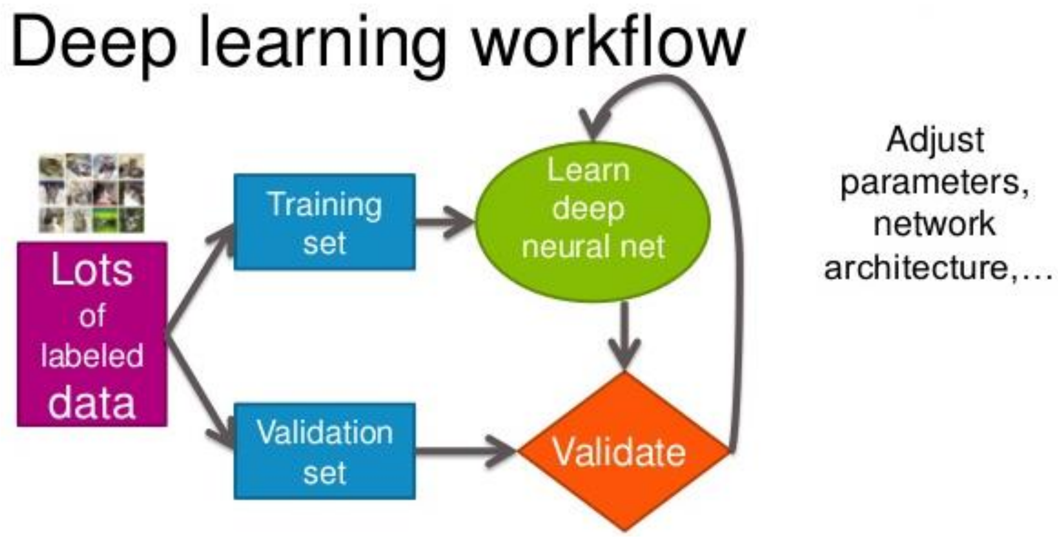
(2) 'customer' means customer as defined in point 24 of Article 2 of Directive 2009/73/EC;

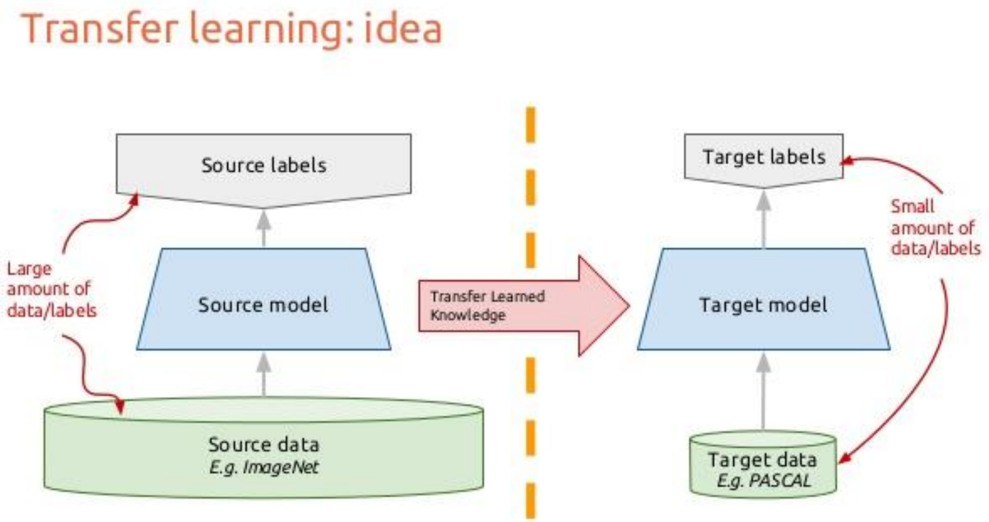
(3) 'household customer' means household customer as defined in point 25 of Article 2 of Directive 2009/73/EC;

(4) 'essential social service' means a service related to healthcare, essential social care, emergency, security, education or public administration;

- Compare different deep NEL systems according to different criteria
  - Accuracy, F1 Score
  - Rank networks respectively
- Choose dedicated algorithm for the integration
- Big datasets used for transfer learning
  - WNED
  - AIDA-CoNLL

- Apply Transfer Learning:
  Adapt pretrained algorithm to specific needs for private (smaller, unlabeled) datasets
- Here: Datasets from the legal domain, EUR-Lex
  Topic: EU Regulation
- Test network and interpret results



# Outline

Motivation

Research Questions

Research Approach

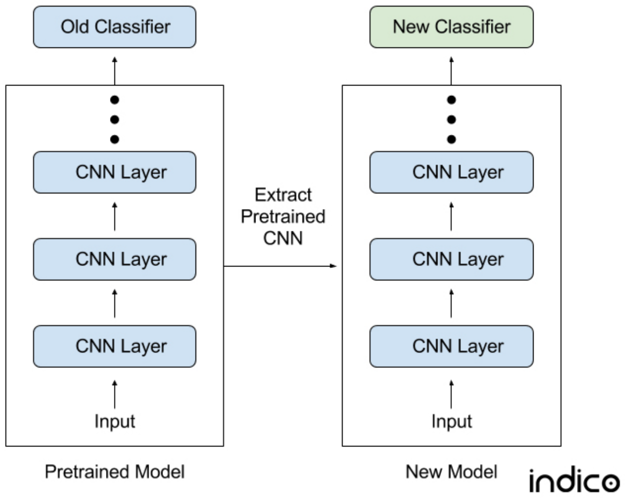
**Related Work**

Implementation

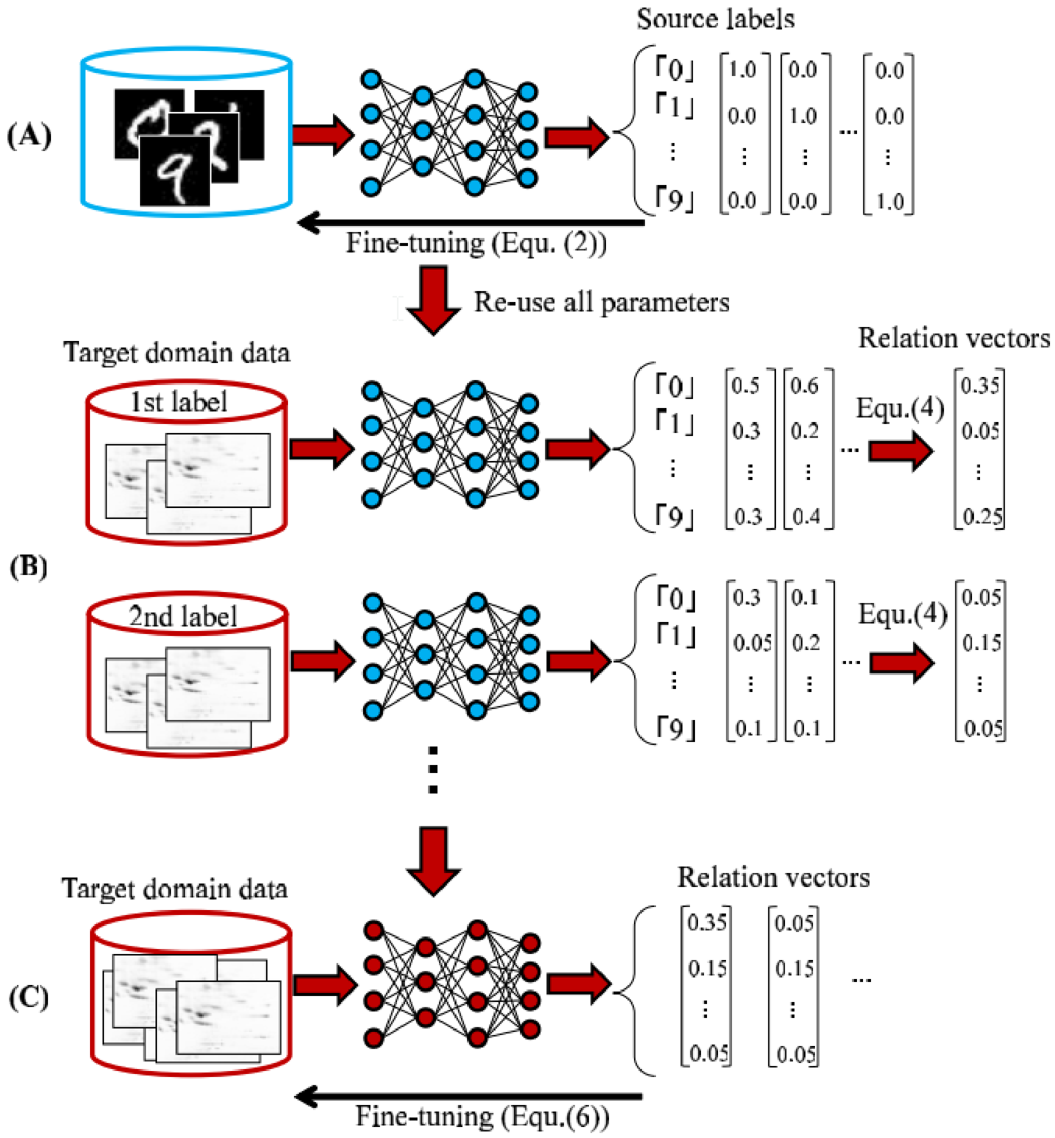
Evaluation

Conclusion

- Most approaches
  - Construct base model trained on source domain data
  - Construct second model using hidden layers from base model
  - Replace output layer




- Image Recognition for Sepsis Classification
  - Source datasets: MNIST and CIFAR-10
  - Target datasets: 2D images showing sepsis and non-sepsis
  - Resulting model achieves 90% accuracy on target dataset



# Related Work

## Named-Entity Linking in Law

sebis

- Scarcity of data(-sets)
- Content extremely domain specific
- Success of NEL highly related to domain knowledge

→ Little attention for NEL in legal domain

→ Successful NEL system in legal domain is yet to find

„no access to such data yet “

„annotation at the level of entities has not been consolidated“

„Therefore, approaches to NEL have only been evaluated on the test portion of the corpus of Wikipedia“

„huge corpus of relevant (domain specific) training data is required “

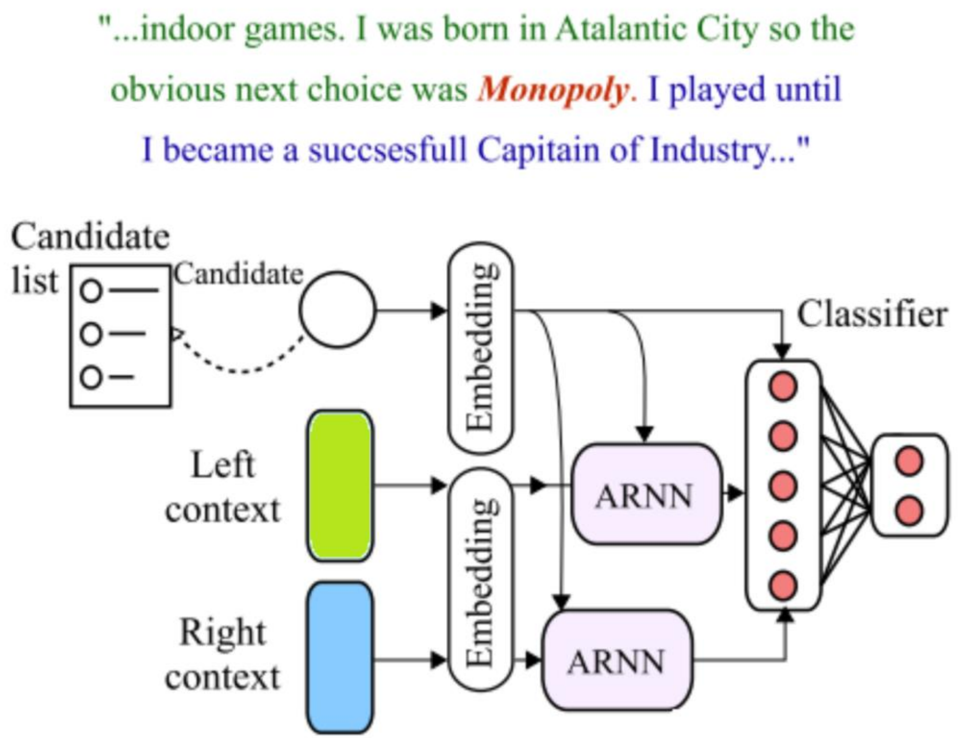
„one of the major problems for NED in the legal domain “

# Related Work

## Deep Named-Entity Linking

- Francis-Landau et al.
  - Use CNNs for NEL
  - Take into account:
    - Mention, context and entire document as source
    - Respective entity title and Wikipedia article as target entity link
  - CNN calculates the preferred entity
  - Accuracy on AIDA-CoNLL: 85.5%

- Eshel et al.
  - NEL for noisy text
  - Goal: capture noise around around local context
  - Performance still below state of the art (Micro P@1 score of 83.3



# Outline

Motivation

Research Questions

Research Approach

Related Work

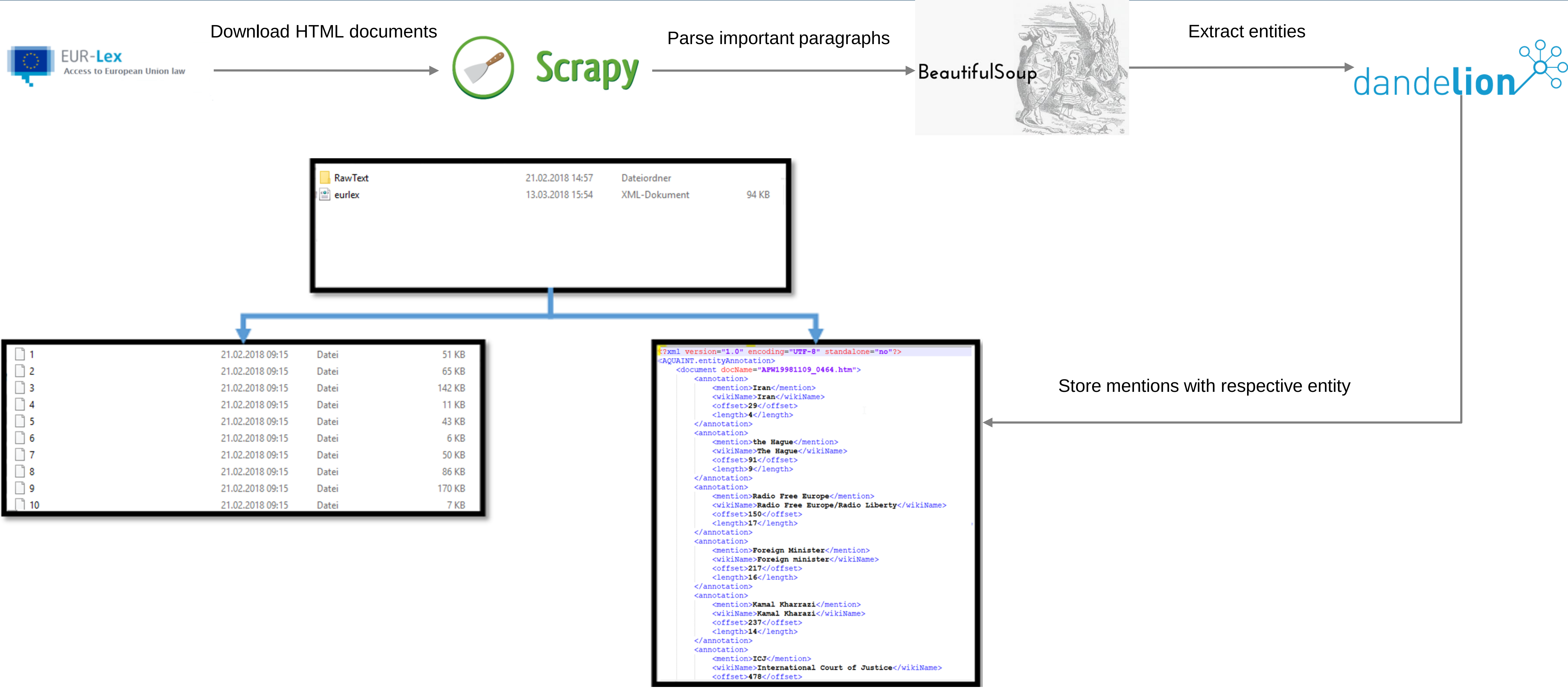
**Implementation**

Evaluation

Conclusion

# Implementation

## Dataset Creation



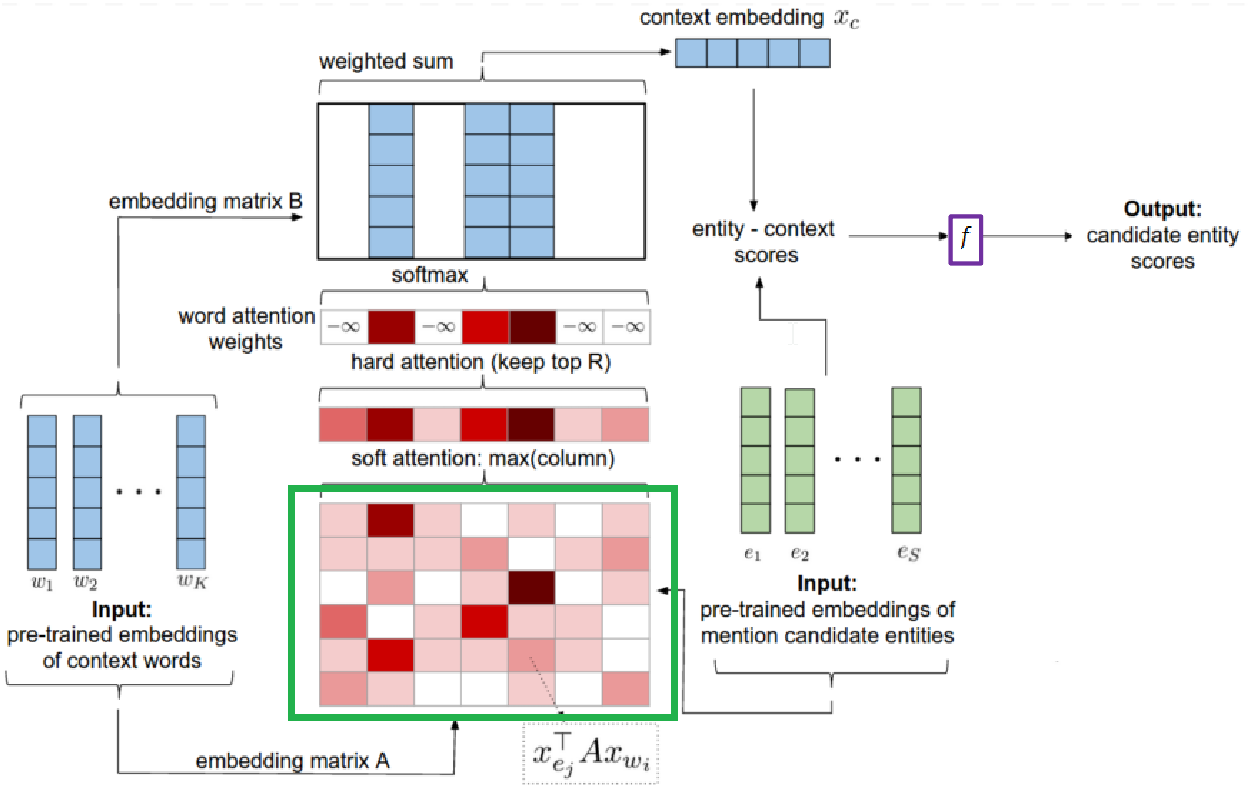
# Implementation

## Dataset Statistics

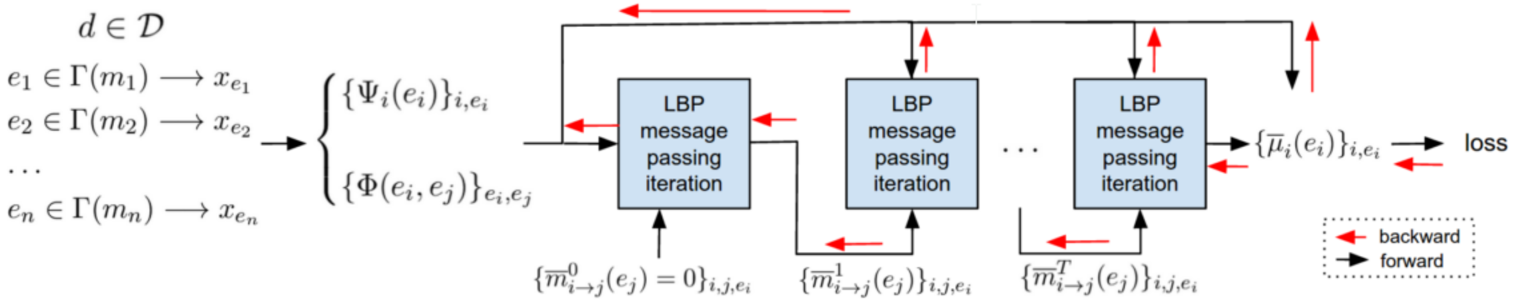
- AIDA-ConLL
  - Widely used for public benchmarks
  - Biggest manually annotated DS
- WNED
  - Automatically created from WP corpus
  - Less trustworthy → mostly used for testing
- MSNBC, AQUAINT & ACE2004
  - Taken from different news corpora
  - Small in comparison to AIDA-CoNLL
- EUR-Lex
  - Created in this work
  - Stored in similar format to WNED
- Joint Dataset
  - Train set: 52,785 entries
  - Test set: 16,465 entries
  - Mixture of automatically generated and manually annotated input-output pairs

| Dataset | Number Mentions | Number Documents |
|---|---|---|
| AIDA-train | 18,848 | 946 |
| AIDA-A | 4,791 | 216 |
| AIDA-B | 4,485 | 231 |
| MSNBC | 656 | 20 |
| AQUAINT | 727 | 50 |
| ACE2004 | 257 | 36 |
| WNED-CWEB | 11,154 | 320 |
| WNED-WIKI | 6,821 | 320 |
| EURLEX-train 1k | 1,853 | 1,118 |
| EURLEX-test 1k | 333 | 185 |
| EURLEX-train 20k | 33,937 | 17,352 |
| EURLEX-test 20k | 11,674 | 4,580 |

- Entity Embeddings
  - semantic meaning of entities
  - Inherited from word embedding
  - Word2vec → Entity2vec




- Local Model with Neural Attention
  - Context score per entity
  - *f*
    - two fully connected layers
    - 100 hidden units
    - ReLU




- Collective Disambiguation
  - Takes into account entity context scores
  - Acts as classifier to choose correct entity

# Outline

Motivation

Research Questions

Research Approach

Related Work

Implementation

**Evaluation**

Conclusion

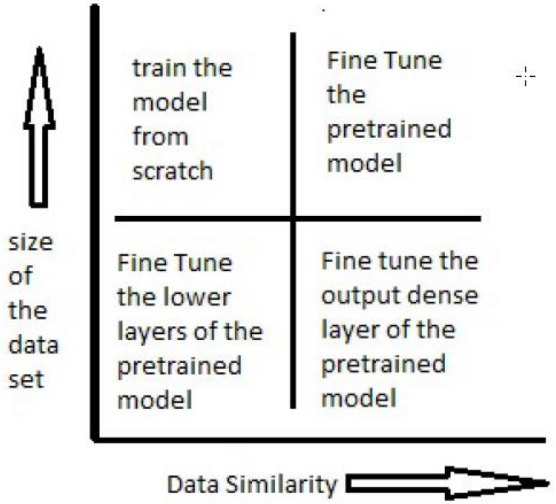
# Evaluation

## Experimental Setup I

- Hardware
  - Provided by [iteratec GmbH]
  - GPU: NVIDIA GeForce GTX TITAN X | 12 GB
  - CPU: Intel Core i7-5820k | 6 cores | 3.30 GHz
  - RAM: 16 GB

- Software
  - PyTorch
    - Python package for scientific computing
  - Lua
    - Lightweight, robust programming language
    - Most common scripting language in game development

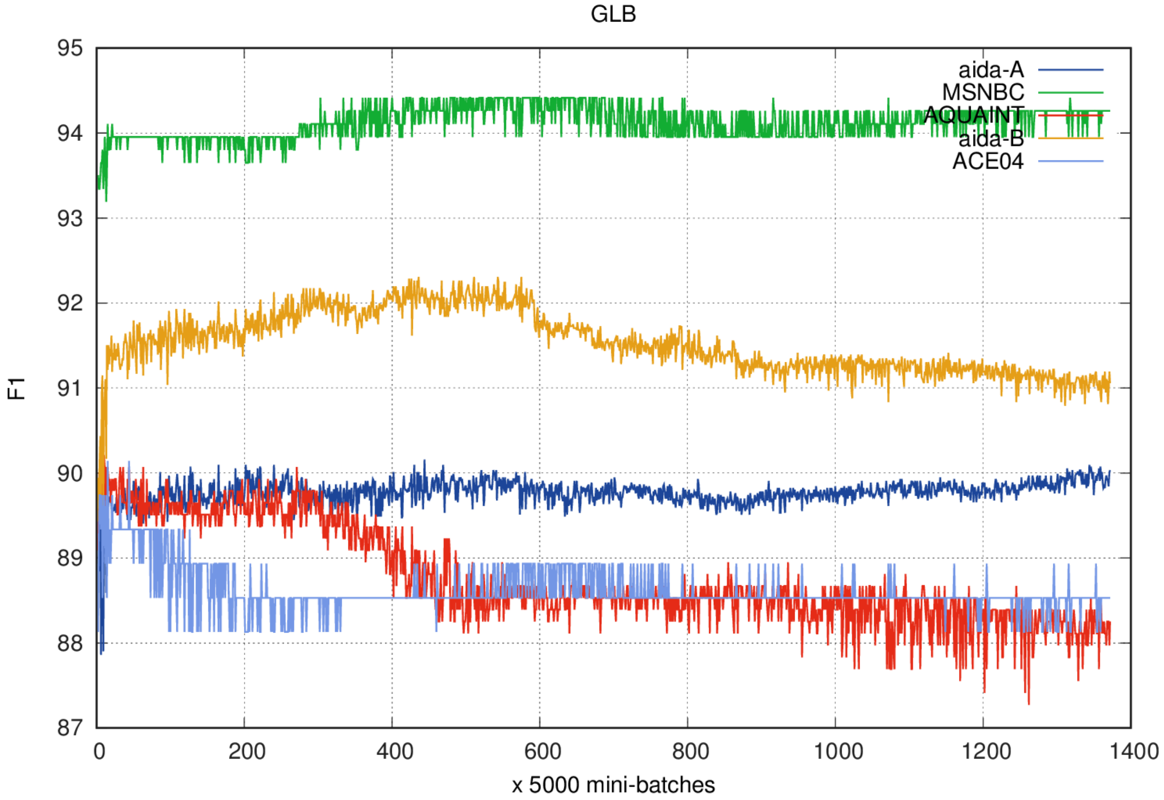
- Single Training
  - Train & test on AIDA-CoNLL
  - Train & test on EUR-Lex 1k
  - Train & test on EUR-Lex 20k
- Joint Training
  - Merge AIDA-CoNLL and EUR-Lex
  - Have respective train & test sets
  - Train on merged datasets
- Transfer Learning
  - Fine tune pretrained models
  - AIDA-CoNLL → Fine tune with EUR-Lex 1k
  - AIDA-CoNLL → Fine tune with EUR-Lex 20k
  - EUR-Lex 20k → Fine tune with AIDA-CoNLL

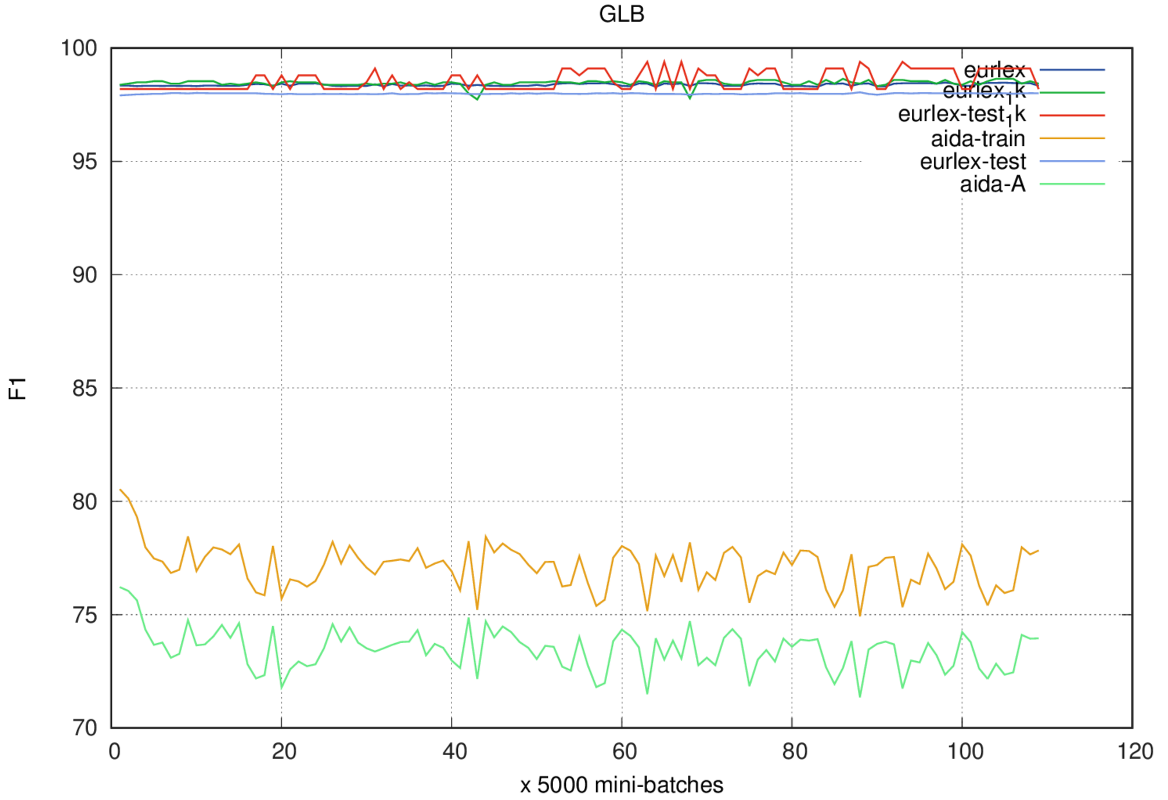



- Metrics
  - **Accuracy** $\boxed{\text{Accuracy}} = \frac{TP + TN}{TP + FP + FN + TN}$
  - Precision $PRE = \frac{TP}{TP + FP}$
  - Recall $REC = TPR = \frac{TP}{P} = \frac{TP}{FN + TP}$
  - **F1 Score** $\boxed{F_1} = 2 \cdot \frac{PRE \cdot REC}{PRE + REC}$

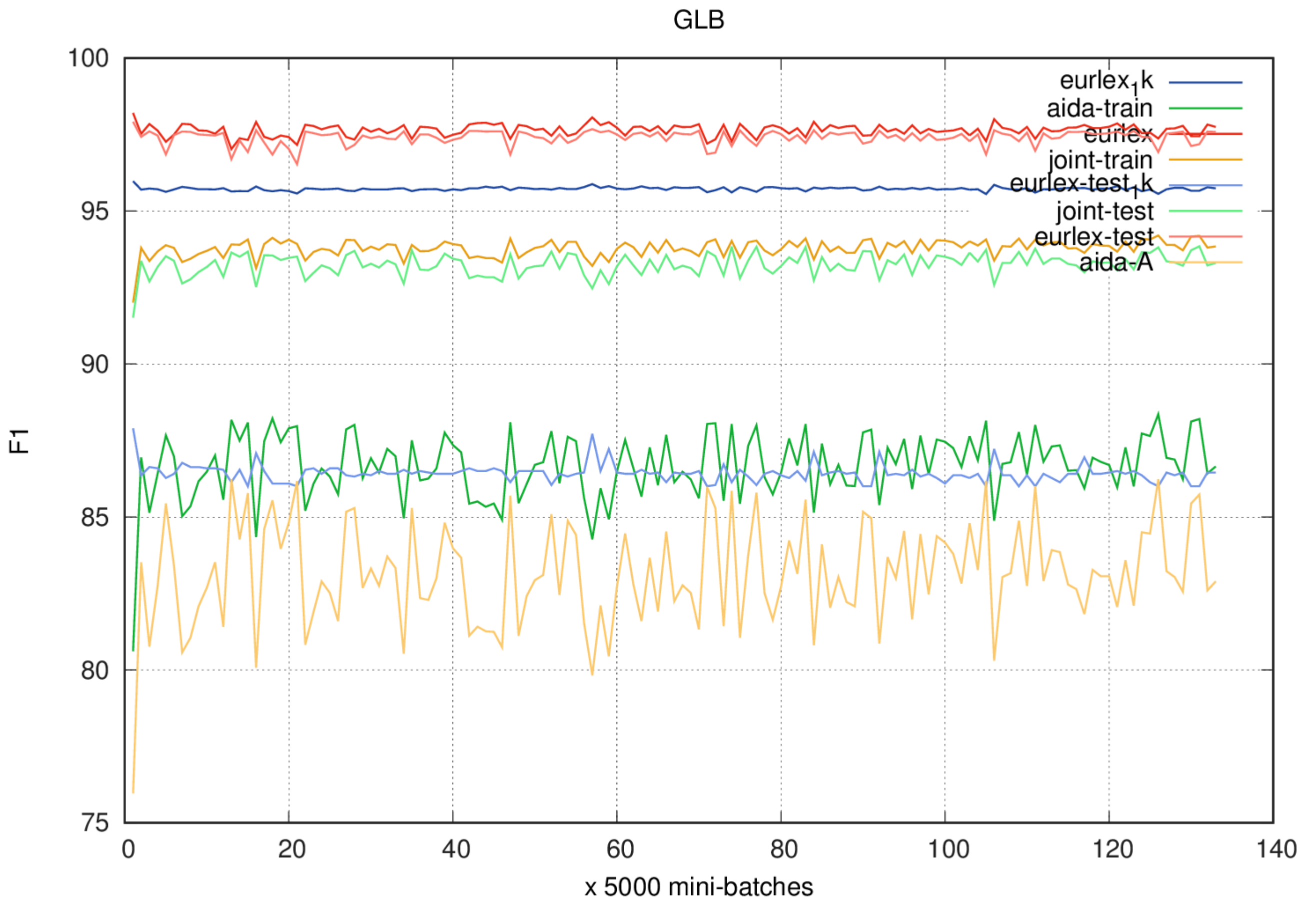
- AIDA-CoNLL
  - Manually annotated
  - F1 score: 90.1%

- EUR-Lex 20k
  - Automatically created
  - F1 score: 98.01%



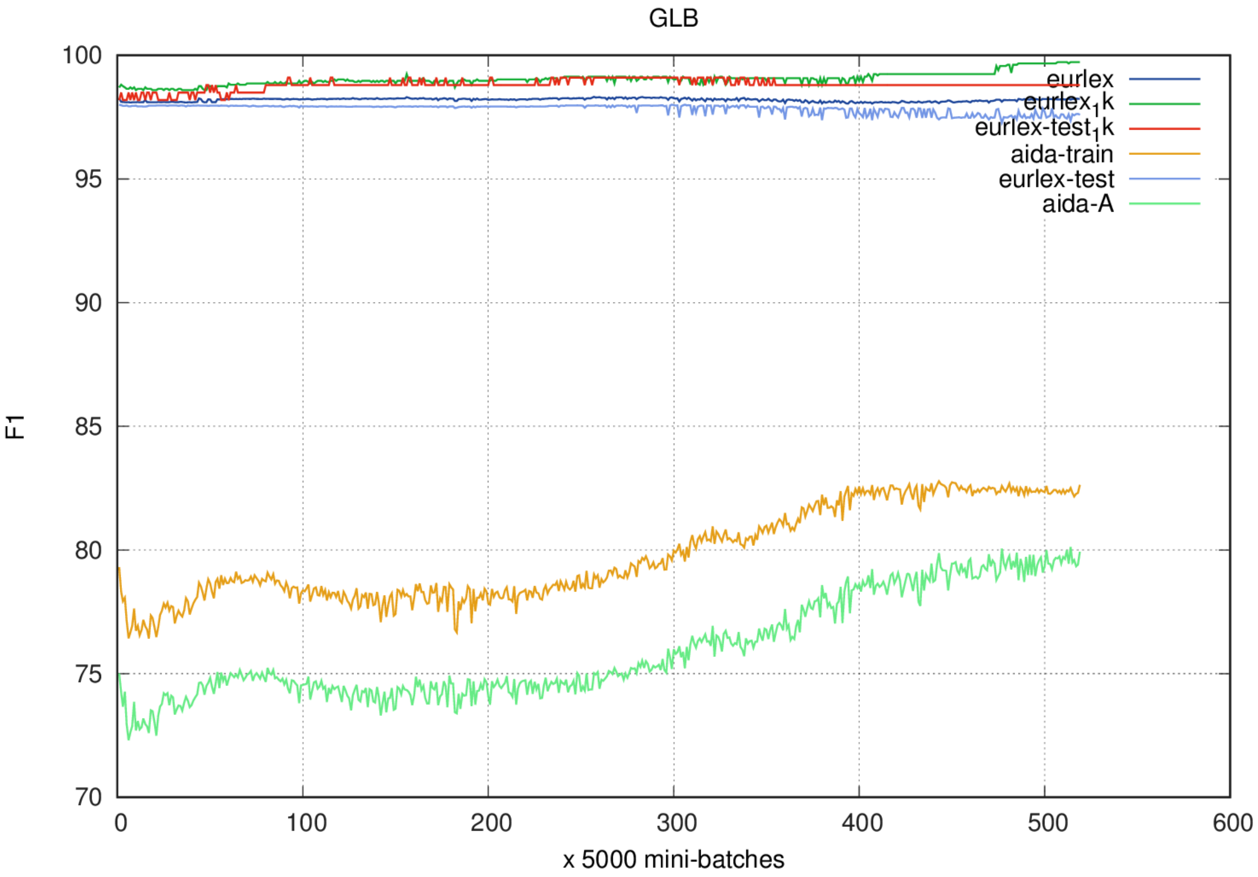
- Merged dataset of AIDA-CoNLL and EUR-Lex 20k
  - F1 scores
    - Joint test set: 93.73%
    - AIDA
      - Train: 87.50%
      - Test: 85.29%
    - EUR-Lex 20k
      - Train: 97.36%
      - Test: 97.19%

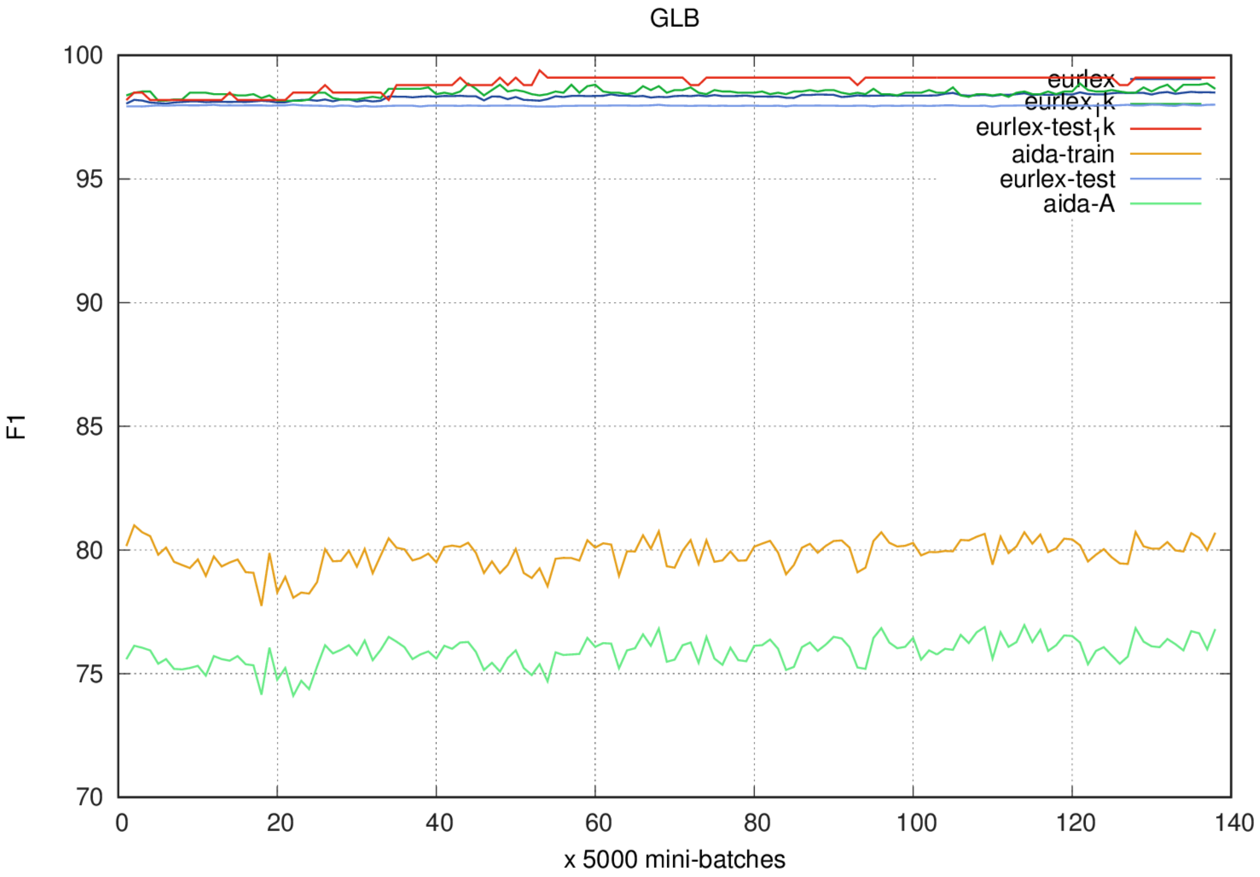


# Evaluation

## Transfer Learning AIDA-CoNLL → EUR-Lex

- EUR-Lex 1k
  - F1 score
    - Train: 99.73%
    - Test: 98.90%
- EUR-Lex 20k
  - F1 score
    - Train: 98.49%
    - Test: 98.01%

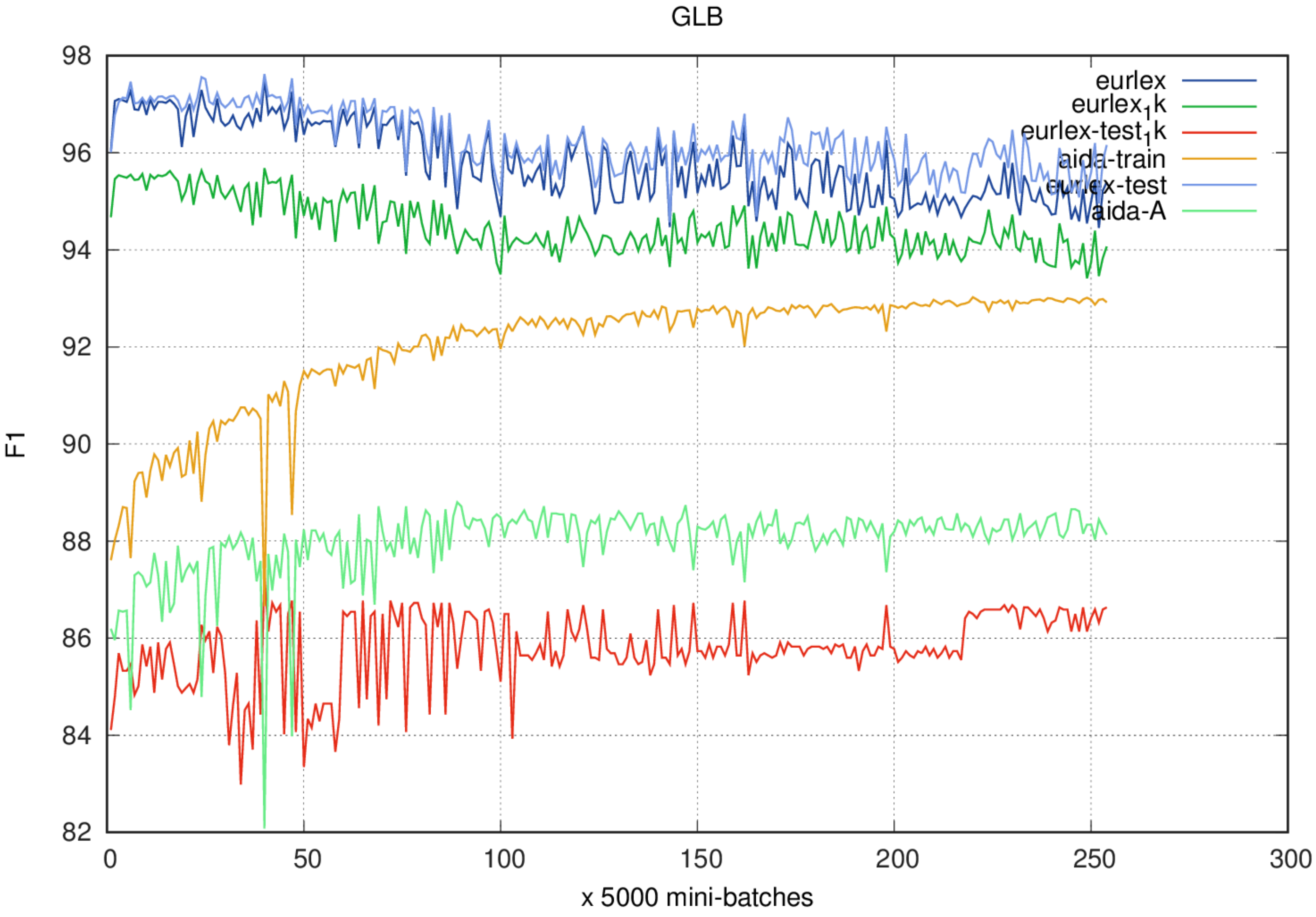


# Evaluation

## Transfer Learning EUR-Lex → AIDA-CoNLL

- F1 score
  - Train: 93.41%
  - Test: 88.80%



# Evaluation

## Discussion

| F1 Score Comparison | Single Training | Joint Training | Transfer Learning |
|---|---|---|---|
| AIDA-train | 92.36% | 87.50% | **93.41%** |
| AIDA-A | **90.1%** | 85.29% | 88.8% |
| EUR-Lex train 1k | 99.02% | 97.14% | **99.73%** |
| EUR-Lex train 20k | 98.34% | 97.36% | **98.49%** |
| EUR-Lex test 1k | 98.29% | 90.41% | **98.90%** |
| EUR-Lex test 20k | 98.01% | 97.19% | 98.01% |

# Outline

Motivation

Research Questions

Research Approach

Related Work

Implementation

Evaluation

**Conclusion**

# Conclusion

## Revisiting the RQ

sebis

### 1. What kind of existing approach should be used for transfer learning?

- Deep Joint NEL
  - Deep learning
  - High performance
  - State of the art
  - In contact with author

### 2. Which technique of transfer learning suits best?

- Employed state of the art
- Put datasets in relation: size & similarity
  → Pure fine tuning without layer adaption

### 3. Is the use of transfer learning with named-entity linking beneficial in the legal domain?

- Performance increase for AIDA-CoNLL → EUR-Lex 1k/20k
- Slight increase for EUR-Lex 20k → AIDA-CoNLL (only on training set)
- Legal domain benefits from transfer learning
- Implication: NEL systems can improve through transfer learning

# Thank you for your attention!

iteratec

KOMPETENZ, DIE ENTLASTET

# Literature

Sources:


- http://www.saupair.com/the-work-of-a-criminal-defense-lawyer/
- http://www.sokulskieng.com/
- https://sixfeetup.com/blog/an-introduction-to-beautifulsoup
- https://blog.theodo.fr/2018/02/scrape-websites-5-minutes-scrapy/
- https://www.pinterest.co.uk/pin/314477986464608806/
- https://www.digitaldoughnut.com/articles/2018/february/a-comparison-of-deep-learning-frameworks
- https://blog.dandelion.eu/
- https://sebastianraschka.com/faq/docs/multiclass-metric.html
- https://medium.com/greyatom/performance-metrics-for-classification-problems-in-machine-learning-part-i-b085d432082b
- https://arxiv.org/pdf/1711.04450.pdf
- https://indico.io/exploring-computer-vision-transfer-learning/
- https://dl.acm.org/citation.cfm?id=2390963
- https://arxiv.org/pdf/1604.00734.pdf
- https://www.aclweb.org/anthology/K/K17/K17-1008.pdf
- https://data-flair.training/blogs/transfer-learning/
- http://www.lua.org/about.html
- https://pytorch.org/
- https://www.iteratec.de/
- https://wwwmatthes.in.tum.de/pages/1a9cbqbo2o7p1/Master-s-Thesis-Ingo-Glaser
- https://wwwmatthes.in.tum.de/pages/1sy7ehcl1sz4z/Named-Entity-Recognition-Extraction-and-Linking-in-German-Legal-Contracts