

Using Natural Language Processing and Machine Learning to Assist First-Level Customer Support for Contract Management

Master thesis – Final presentation

Michael Legenc

Advisor: Daniel Braun

Munich, 08.01.2018

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

www.matthes.in.tum.de

Introduction

- Scalability issues of email customer supports.
- Acceleration by automation and assistance features.

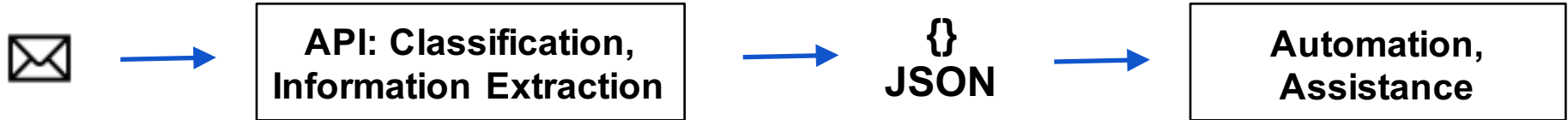


Hurdle: Free text.



Machine Learning and Natural Language Processing.





Cancellation of Contract

Contract team

High priority

Cancellation



My Customer <customer@mycompany.de>

Cancellations

14:00 PM (1 hour ago)



to ▼

To Whom It May Concern,

I would like to cancel my contract with customer number 381953212.

Reply:

Your request is being processed. We will get back to you shortly.

Sorry, your contract cannot be cancelled. Please visit this link for more information

Your contract has been successfully cancelled.

Contract Management

Customer: My Customer

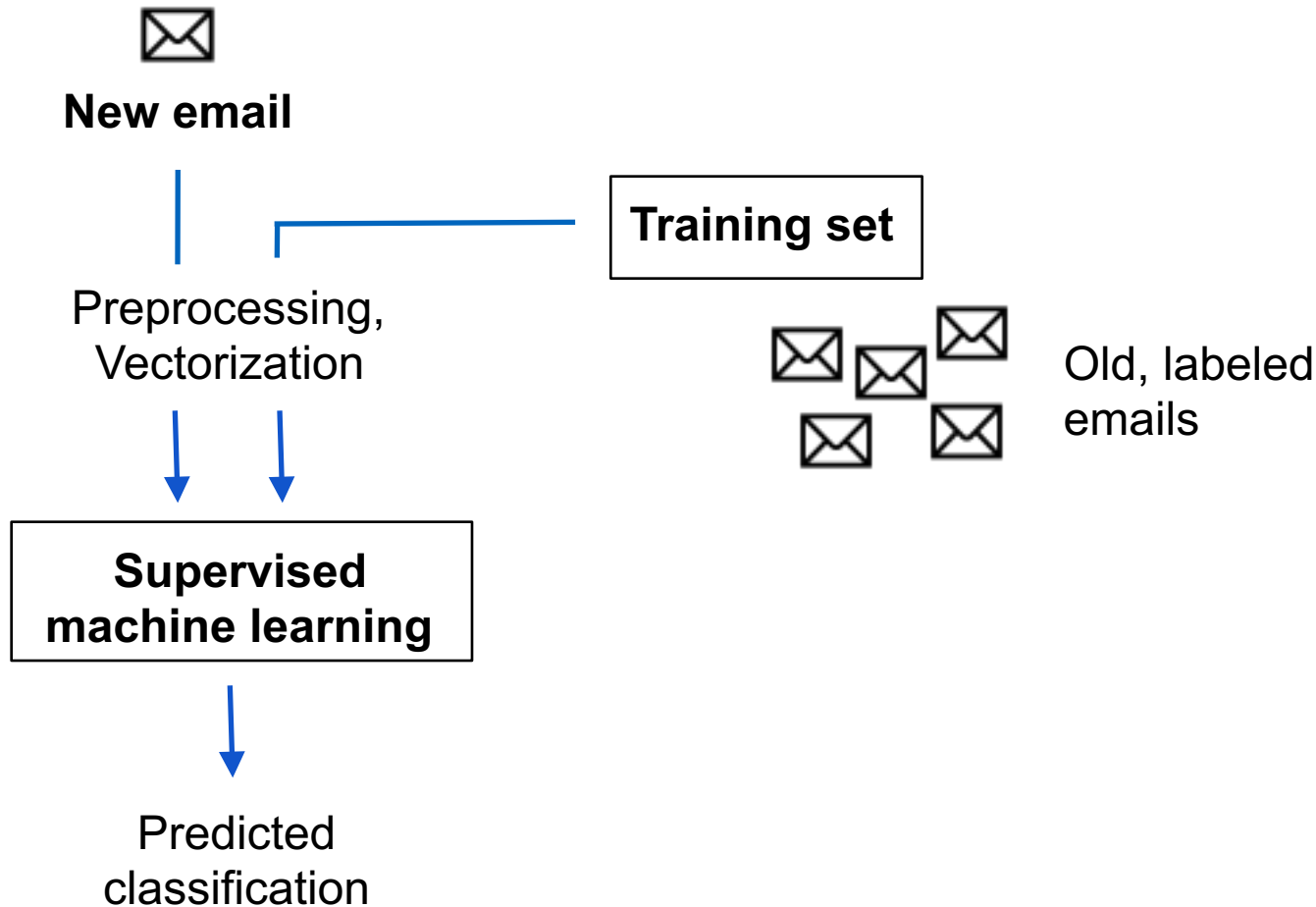
Email: customer@mycompany.de

Contracts: Red 3GB with iPhone 7

Cancel contract

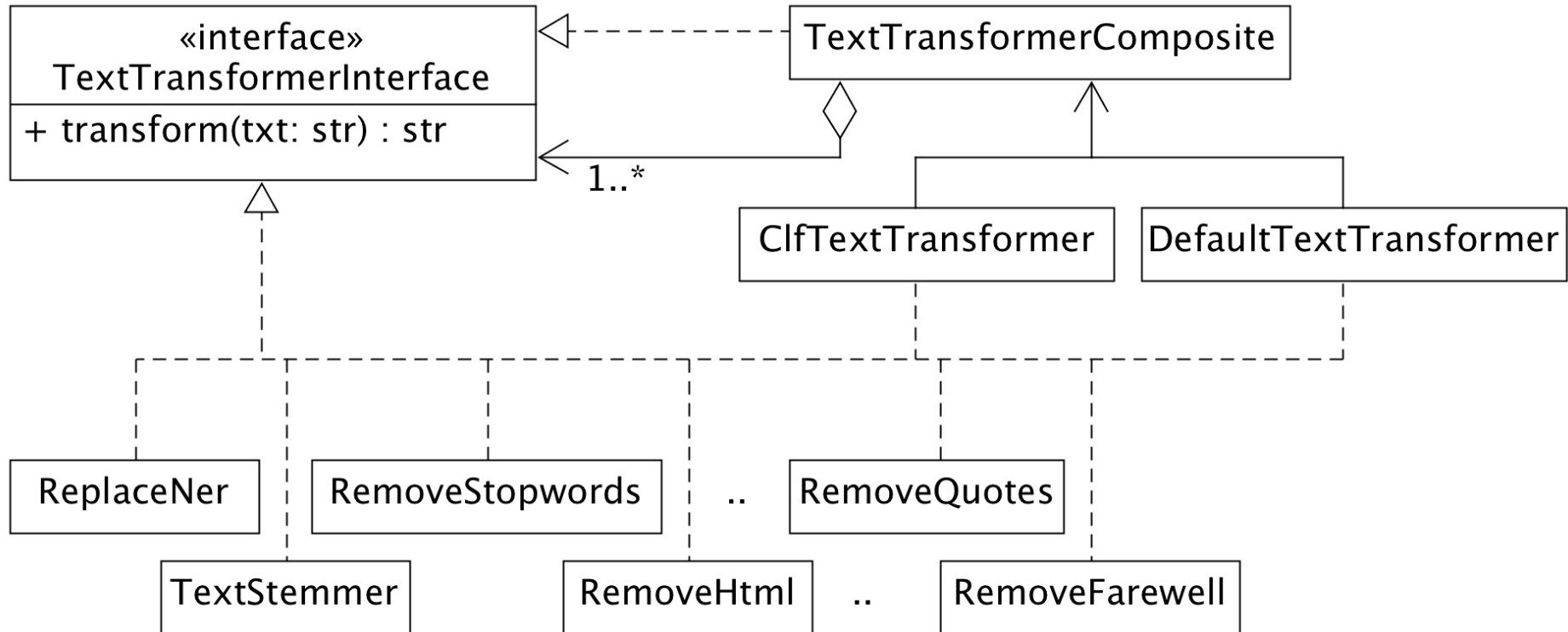
Demo

Classification: Machine Learning



Training set creation

- **Dataset:**
 - 18 Mio. unlabeled Emails.
 - Filtering
 - Retrievable by an API.
 - Caching
- **Command Line Tool: Demo.**



▪ Vectorization: Tf-idf

Classification: Evaluation

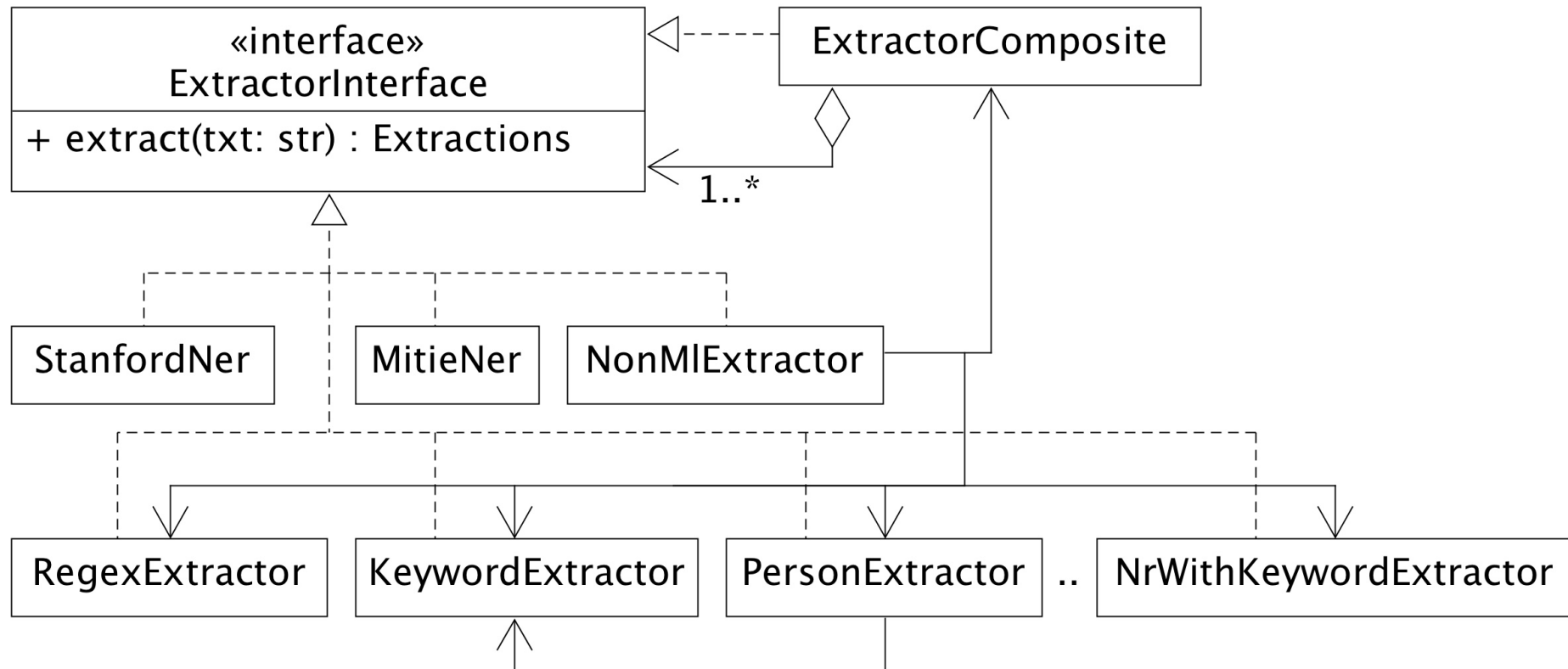
- **Final configuration:**

- Stochastic gradient descent
- Tf-idf thresholds: max. 0.4, min. 0.001

Classifier	Avg. Prec./Rec./F1-Score
k-nearest neighbors	0.68
Logistic Regression	0.65
Multinomial Naive Bayes	0.57
Passive Aggressive Classifier	0.83
Perceptron	0.73
Stochastic Gradient Descent	0.84

Information extraction

- **Implemented types:**
 - Person
 - Date
 - Time
 - Postcode, City
 - Money
 - Vendor
 - Order-, account- and contractnumber



- **Non-ML:** Regex, keyword lists and rule-based.

- Not much support.
- **Learns from the sentence context.**
 - Unknown or misspelled words are recognized by their context.
- **Training set:**
 - Needs massive input. Public data is not modifiable.
 - Creation supported by Command Line Tool and special file format.

- **Used training set:** Created by non-ml approach.
 - 1000 emails. 3361 entities.
- **Test set:** Manually created.

Stanford NER:

	precision	recall	f1-score	support
BILL_NR	N/A	0.00	N/A	1
CITY	1.00	0.76	0.86	21
CONTRACT_NR	0.00	0.00	N/A	2
DATE	0.57	0.10	0.17	39
METER_NR	1.00	0.56	0.71	9
MONEY	1.00	0.65	0.79	23
ORDER_NR	1.00	0.50	0.67	4
PERSON	0.94	0.82	0.88	102
POSTCODE	1.00	0.92	0.96	12
TIME	1.00	0.50	0.67	4
VENDOR	0.94	0.81	0.87	37
avg / total	0.93	0.67	0.78	254

Non-ML:

	precision	recall	f1-score	support
BILL_NR	1.00	1.00	1.00	1
CITY	1.00	0.90	0.95	21
CONTRACT_NR	0.33	0.50	0.40	2
DATE	0.91	0.77	0.83	39
METER_NR	1.00	0.56	0.71	9
MONEY	1.00	0.65	0.79	23
ORDER_NR	1.00	1.00	1.00	4
PERSON	0.94	0.87	0.90	102
POSTCODE	1.00	0.92	0.96	12
TIME	1.00	0.50	0.67	4
VENDOR	0.97	0.85	0.91	37
avg / total	0.95	0.82	0.88	254

Future work

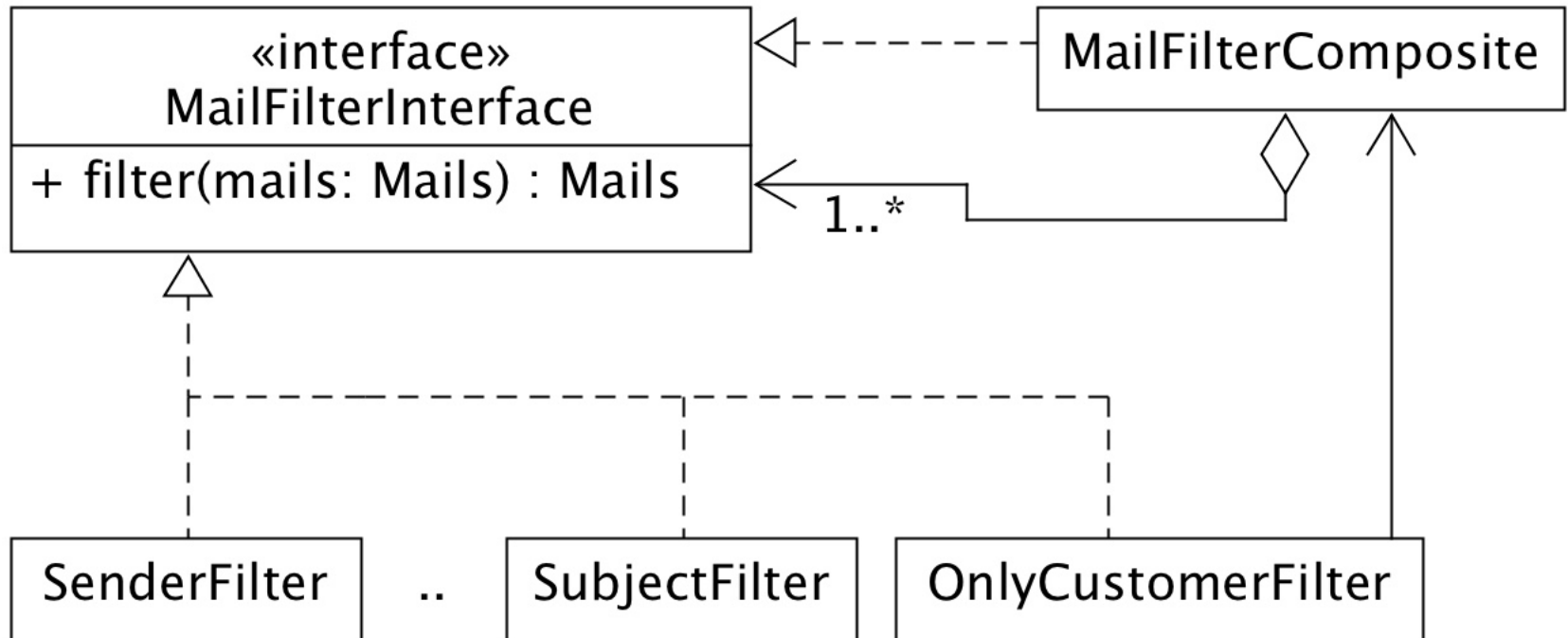
- Transparent training.
- Topic segmentation.
- Better understanding of coherences.
- Automation and assistance features.

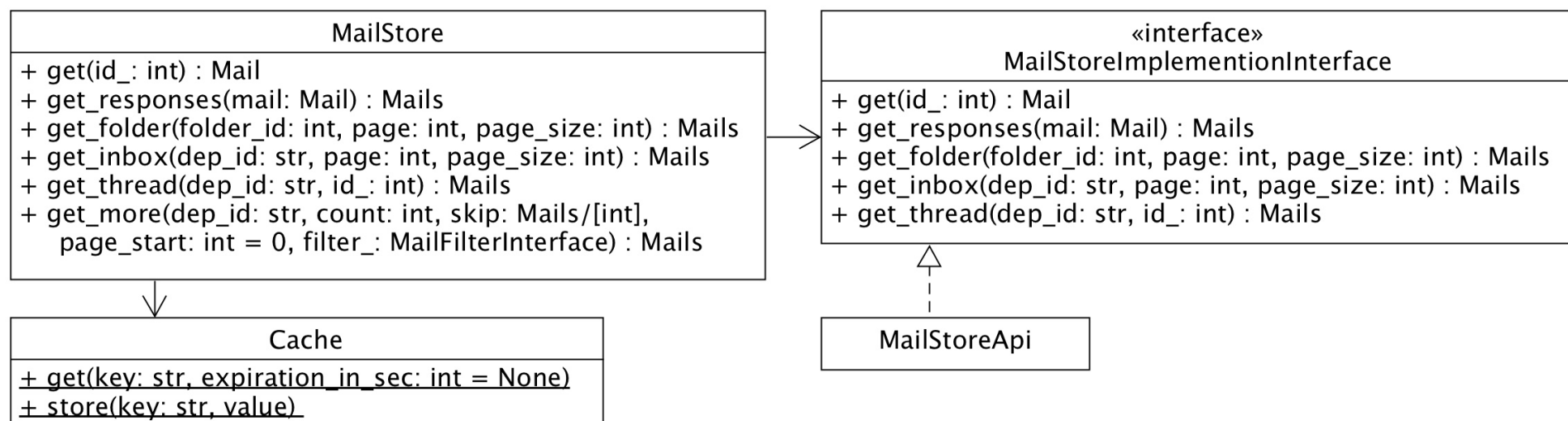
- **Email customer support benefits from automation and assistance:**
 - Time, thus cost saving.
 - Increased employee and customer satisfaction.

- **Free text accessibility by ML and NLP.**

- **Selection of algorithms, parameters and preprocessors depends on data set and concrete application:**
 - Interchangeable toolset approach.
 - Evaluation-based selection.

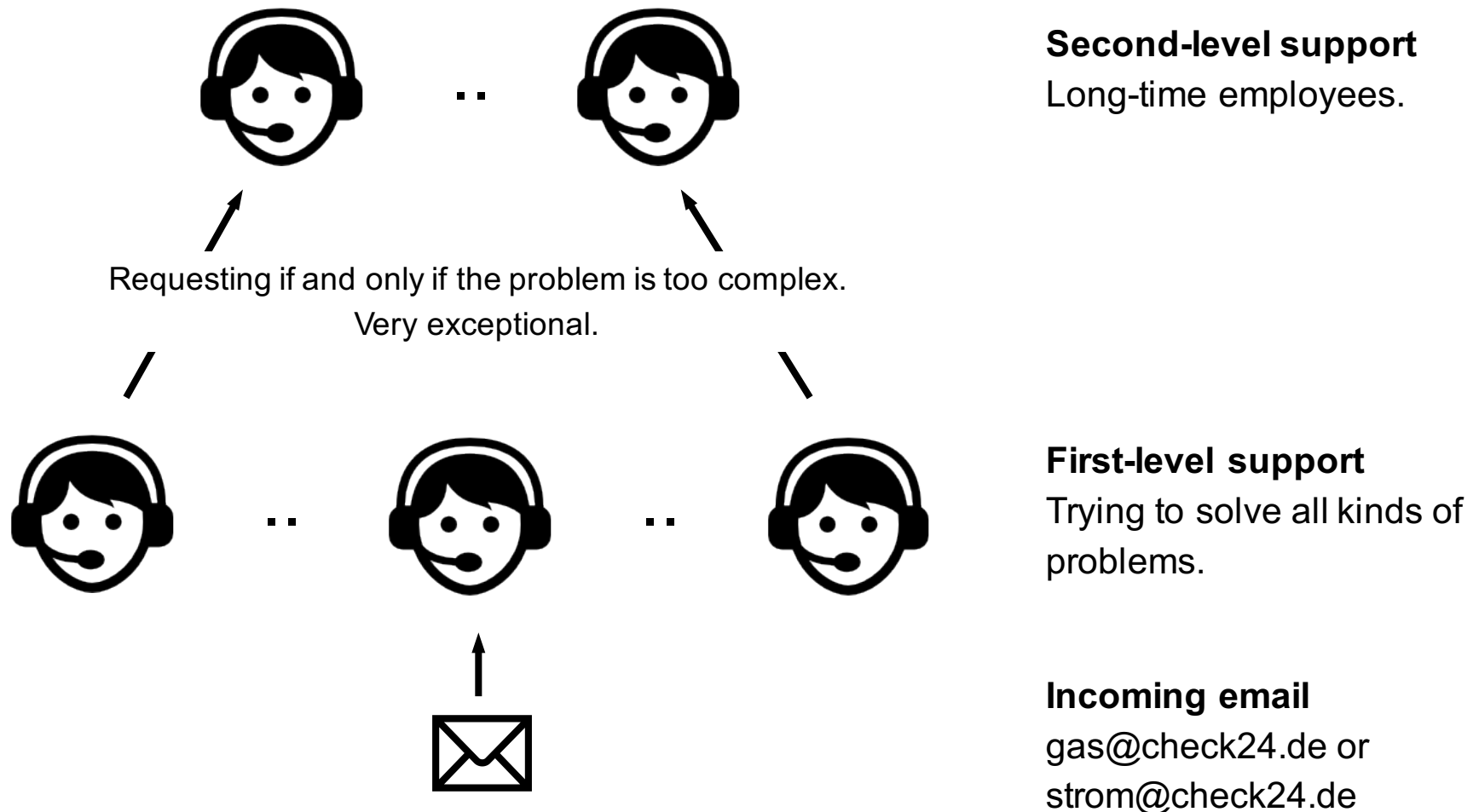
Thank you





Current customer support team organization

at check24's gas and electricity department



Current email preprocessing



Storage

MySQL Server

Long-time storing of all incoming and outgoing emails with the determined folder and customer mappings.



Customer detection

PHP-Server

Mapping mails to customers and simple folders like gas/electricity (unseen) inbox/spam.



Spam detection

Microsoft exchange server

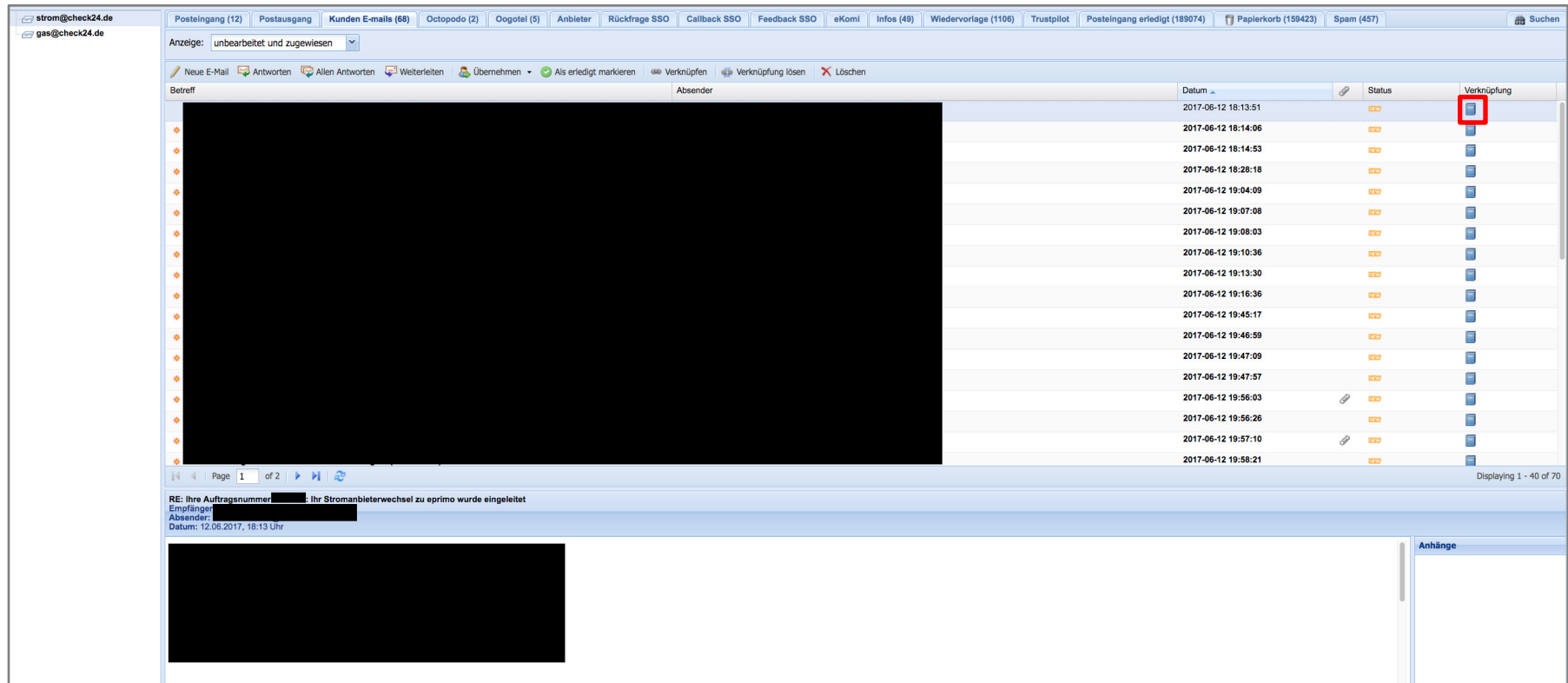


Incoming email

gas@check24.de or strom@check24.de

Software used by the customer support

Webmail: Superficial mail inspection.



Button (e.g. in the red box) leads to a more sophisticated solution.

Software used by the customer support

Proprietary software: View mail with further information and processing features.

Allgemein

Bestellbestätigung
E-Mail Standard
gewünschter Lieferbeginn

Rückfrage

Rückfrage SMS
Daten fehlen
Kundennummer fehlt
Zählernummer fehlt
Bankdaten fehlen
Vorversorger fehlt
Daten ungültig
Neueinzug ungültig
Kündigung liegt i. d. Vergangenheit
Unklare Daten
2 Zählernummern
Doppelantrag
Zählpunktnummer statt Zählernummer angegeben
Kundennr. und ZN identisch
Problemfälle
Storno intern
Vertragsangebot abgelaufen

nach Übermittlung

Widerruf
Statusabfrage Kunde
Sofortbonus Anbieter
Check24 Sofortbonus

Rückmeldung Versorger

ZN bereits in Belleferung
ZN bereits in Belleferung (Neueinzug)
1 PLZ, 2 Netzbetreiber
Wechseltermin noch nicht möglich
Anbieterwechsel fehlergeschlagen
Verzögerung bei Übermittlung
Auftrag erfolgreich bearbeitet

Bewertungs-E-Mail

Auftrags-Historie (1)

E-Mails (0)

Anhänge (0)

E-Mail / SMS-Autorisierungsverlauf, E-Mail / SMS für Daten unvollständig

Alte Anfrage stornieren und neue Anfrage
senden per: ☐ SMS ☐ E-Mail

■ Nachricht konnte nicht zugestellt werden.

Seite: 1

Einträge: 2

ID	Typ	Sender	Empfänger	Nachricht	Datum, Uhrzeit
1720326	out	energie-9343257@vergleich.check24.de		Sehr geehrter Herr Mustermann, vielen Dank für Ihr Vertrauen. Die Daten für Ihren Auftrag zum Stromtarifwechsel wurden telefonisch erfasst, Ihren Auftrag finden Sie als PDF anbei. Um den Auftrag zu bestätigen, antworten Sie bitte auf diese E-Mail mit JA oder klicken Sie hier. Sollten Ihnen bei der Überprüfung des Auftrages fehlerhafte Angaben aufgefallen sein oder haben Sie noch Fragen zu Ihrem Auftrag, so wenden Sie sich hierfür bitte an energie@check24.de.	12.06.2017 13:15:27
				Vielen Dank für Ihren Auftrag (Vattenfall Sales, Easy12 Strom, Auftragsnr.: 9343257). Um	

Contribution

- **Toolset:** Classification, Information Extraction
- **Command Line Tool:**
 - Training set creation
 - Inspection
 - Optimization
 - Evaluation
- **API:** Enables integration