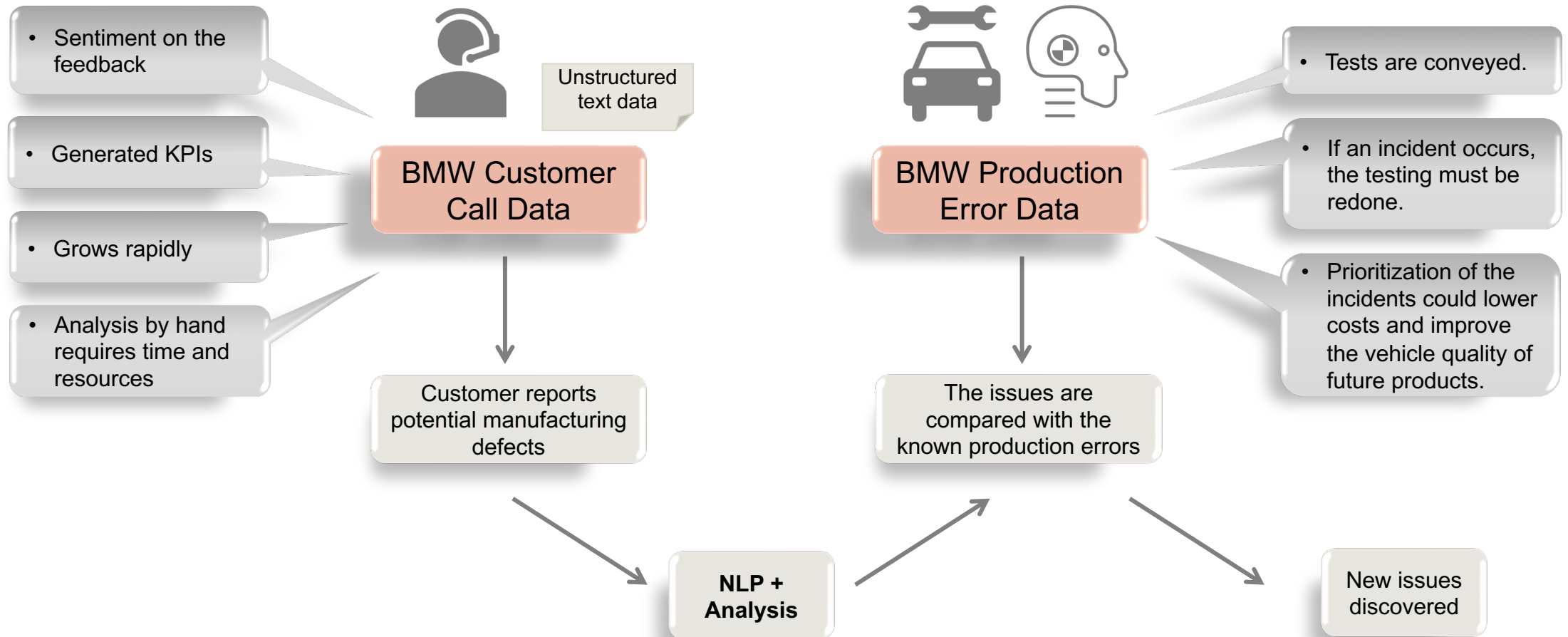


Early Detection of Issues in New Car Models using Natural Language Processing on Customer Call Data

Dilara Ademoglu, 07.11.2022, Final Presentation of Master's Thesis

Chair of Software Engineering for Business Information Systems (sebis)
Faculty of Informatics
Technische Universität München
www.matthes.in.tum.de







Motivation



Research Questions



Background

Topic Modeling



Datasets



Methodology



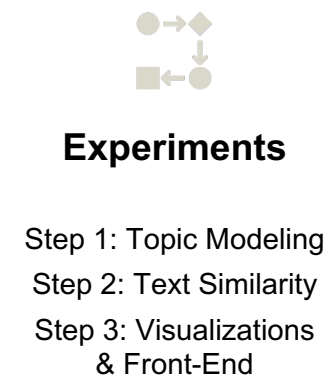
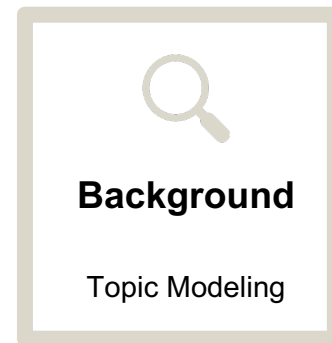
Experiments

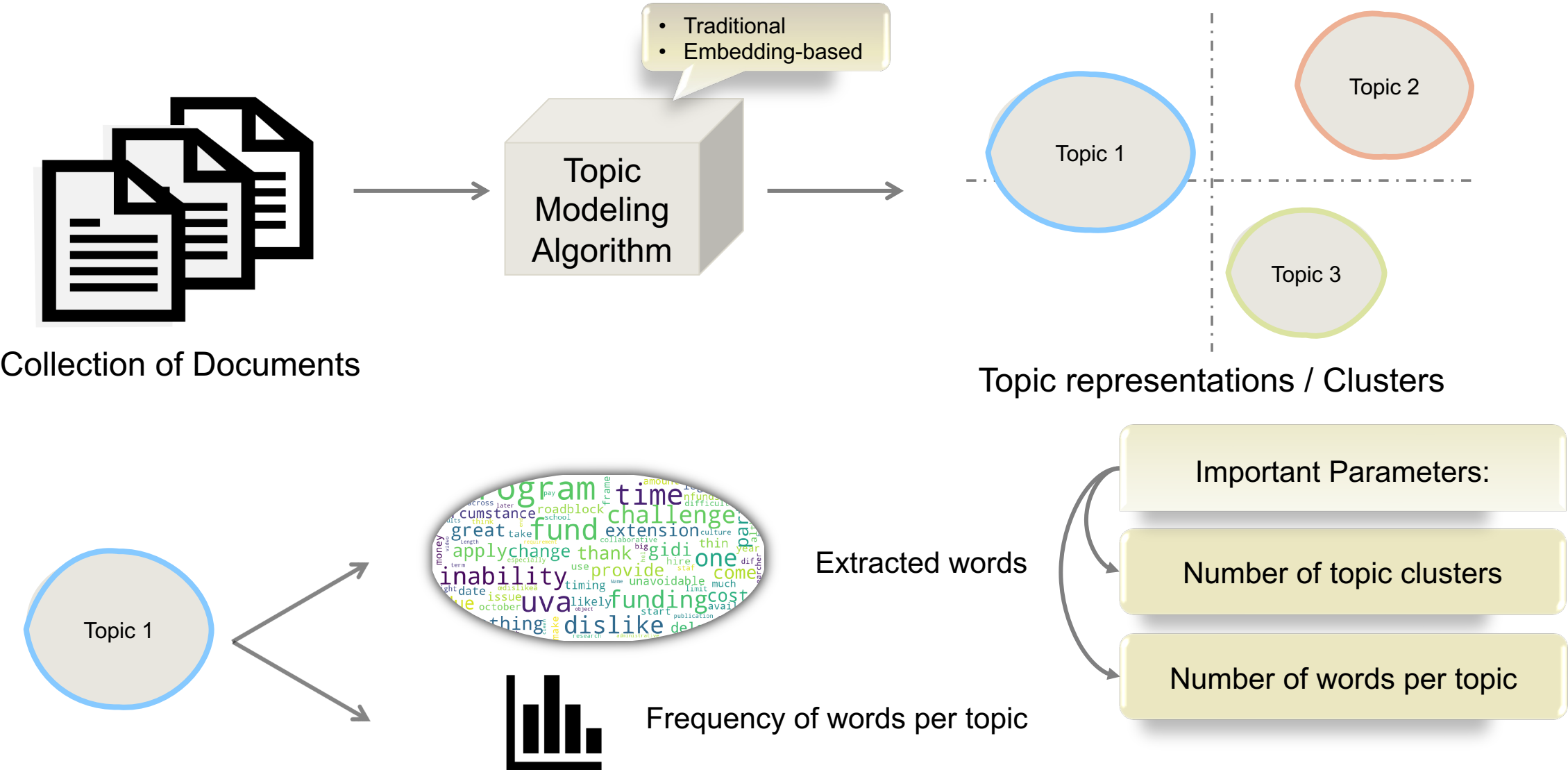
Step 1: Topic Modeling
Step 2: Text Similarity
Step 3: Visualizations
& Front-End



Evaluation Results

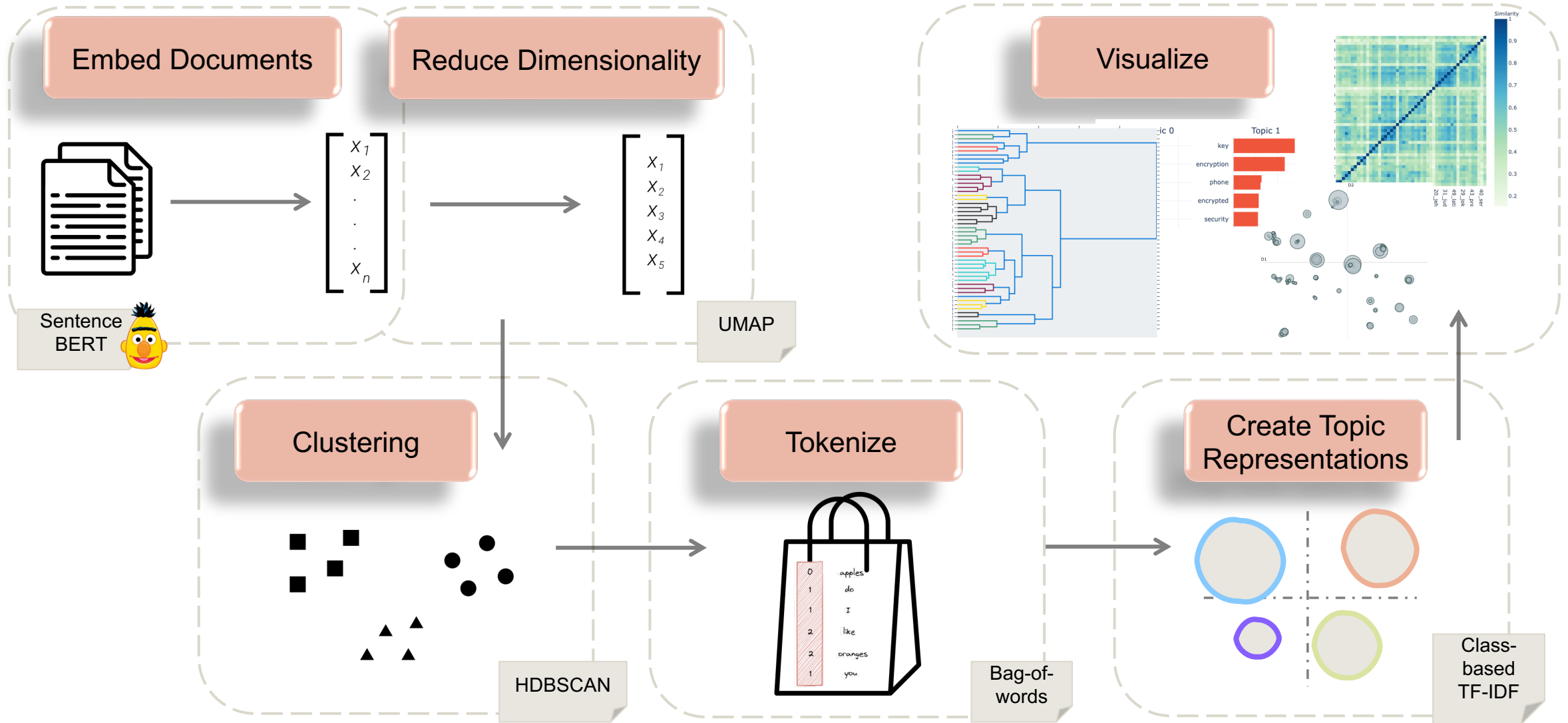
- 1 Which state-of-the-art **topic modeling** approaches would provide better insight into BMW customer feedback datasets?
- 2 Which **text similarity** techniques give better matches between BMW customer feedback datasets?
- 3 How to support quality control departments with interactive topic **visualizations** of BMW customer feedback datasets?





Embedding-based Topic Model

BERTopic





Motivation



Research Questions



Background

Topic Modeling



Datasets



Methodology



Experiments

Step 1: Topic Modeling

Step 2: Text Similarity

Step 3: Visualizations
& Front-End



Evaluation Results

Column Name	Description
van_17	The unique anonymized vehicle identifier. (text)
call_date	The date of the call. (timestamp)
category	Hand-labeled category of the customer feedback. (text)
customer_feedback	The transcribed customer feedback text. (text)
feedback_type	Labeled type of feedback, four categories available: Likes, Difficult to Use, Wants, Defects. (categorical)
production_series	The model name during the production. (text)

Customer call data.

Number of samples: 236.878

- 
- Customer stated he wants a lighter trunk because it's too heavy.
 - Customer stated that he feels that the Navigation is hard to read at times.
 - Customer stated he wanted a UV Sunshade to protect his interior. Writer advised they can be found on the BMW shop site.

Column Name	Description
pqm_incident_number	The unique incident identifier. (numeric)
pqm_problem_number	The unique problem identifier. (numeric)
pqm_problem_assignment_date	The date of incident assignment. (timestamp)
pqm_incident_title	The title of the PQM incident. (text)
pqm_incident_description	The description of the PQM incident. (text)
pqm_model_range	The model name during the production. (text)
pqm_problem_title	The title of the PQM problem. (text)
pqm_problem_description	The description of the PQM problem. (text)

Production error data.

Number of samples: 247.608

Incident:

1. Customer complains about steering wheel being off center respectively car pulling to one side.
2. Steering wheel not straight.
3. Front Axle, Steering Wheel misalignment

Problem:

- Front Axle, Steering Wheel misalignment and vehicle pulling.



Motivation



Research Questions



Background

Topic Modeling



Datasets



Methodology

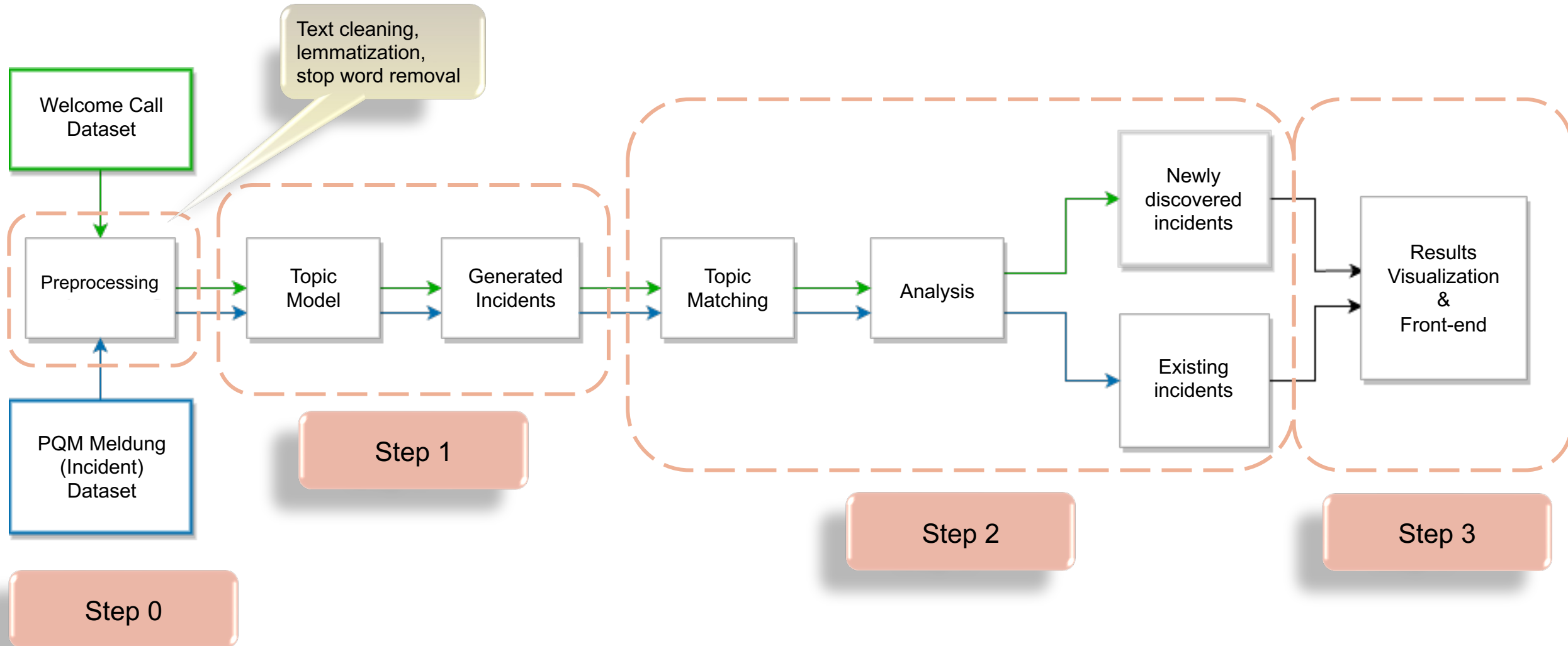


Experiments

Step 1: Topic Modeling
Step 2: Text Similarity
Step 3: Visualizations
& Front-End



Evaluation Results





Motivation



Research Questions



Background

Topic Modeling



Datasets



Methodology



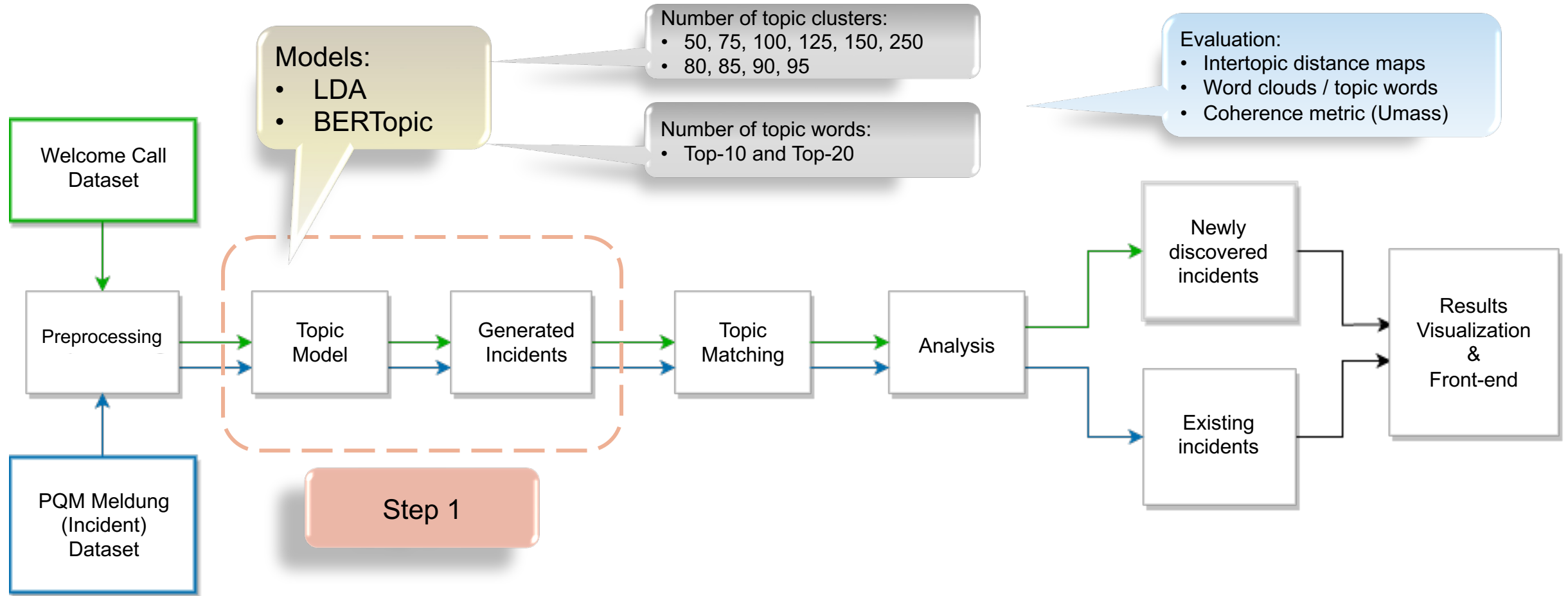
Experiments

Step 1: Topic Modeling
Step 2: Text Similarity
Step 3: Visualizations
& Front-End



Evaluation Results

Step 1 – Topic Modeling



Experimental Results – Topic Modeling

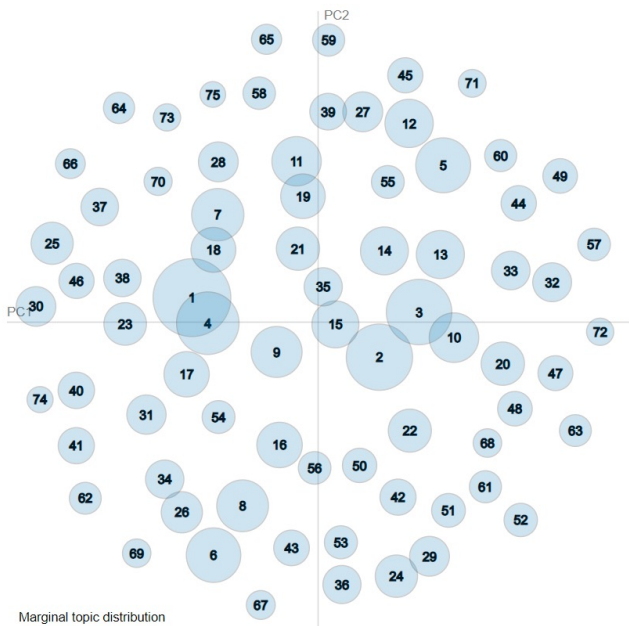
Welcome Call – 75 topic clusters (top-20-words)

Intertopic distance map

Topic words: **seat**

Umass coherence score

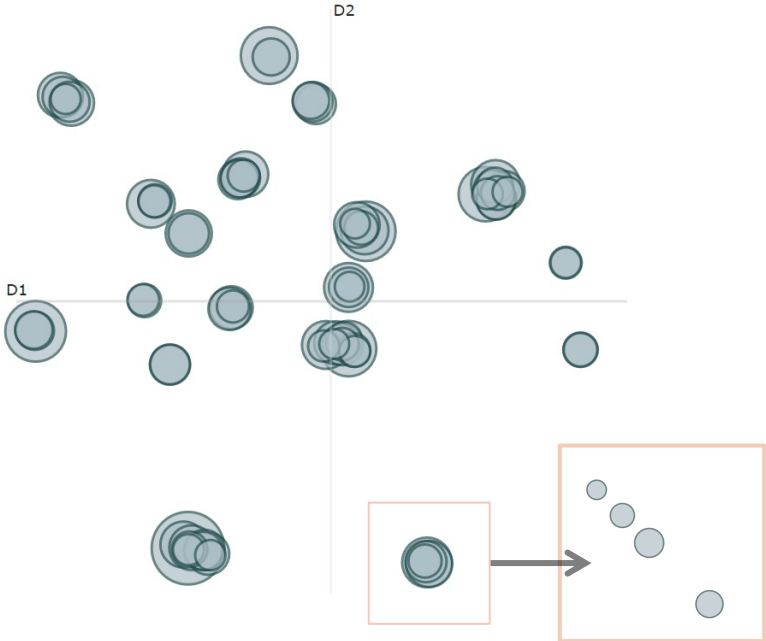
LDA



seat, *like*, passenger, belt, want, heat, sun, area, higher, leather, sit, easily, material, *able*, inch, visor, adjustment, easy, acceleration, smoother

-4.49

BERTopic

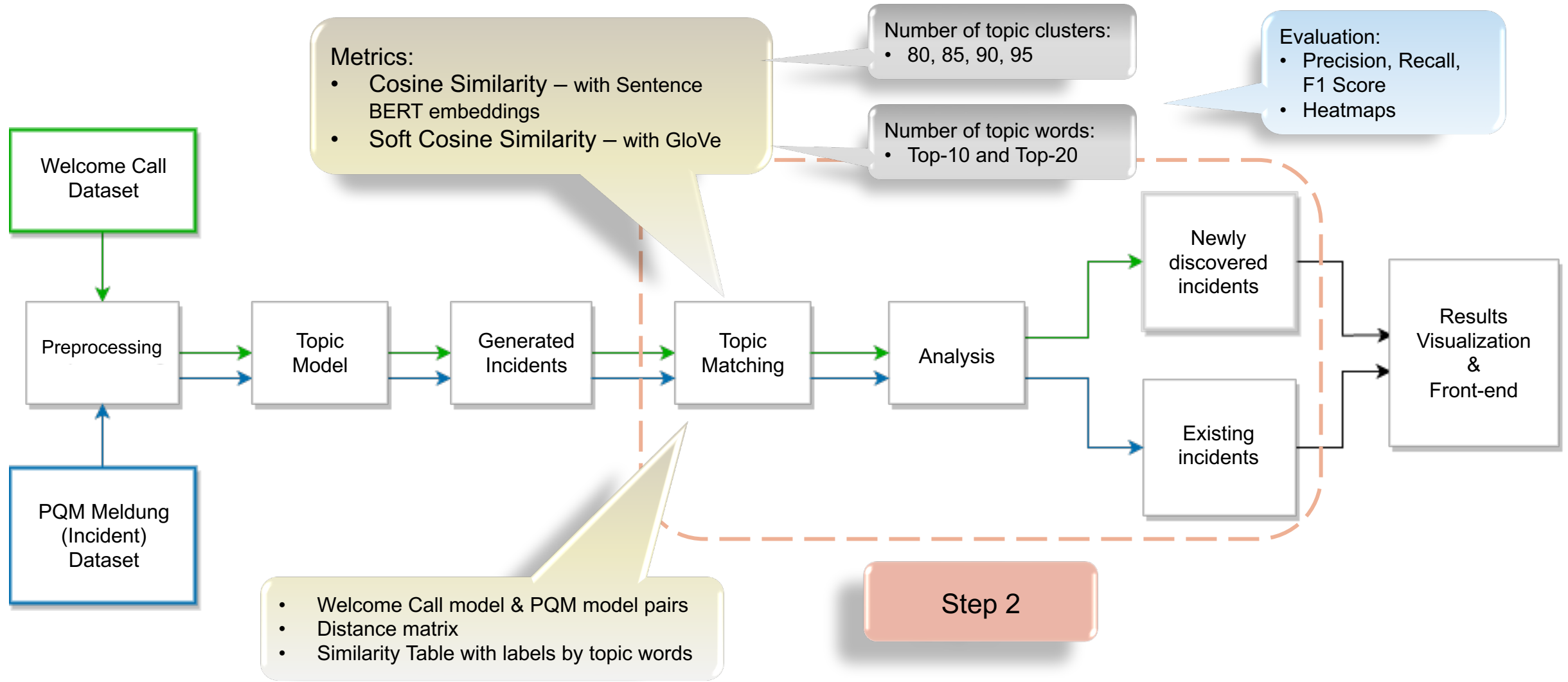


seat, row, massage, room, rear, adjustment, leg, passenger, adjust, space, recline, message, second, lock, interior, position, forward, chair, like, child

leather, interior, **seat**, color, upholstery, material, look, stitch, beige, black, jean, like, picture, stain, plastic, white, passenger, cognac, red, clean

-3.12

Step 2 – Topic Matching / Text Similarity



Experimental Results – Text Similarity

Analytical Measures – F1 Score

Number of Welcome Call clusters	Number of PQM clusters	F1 Score: Cosine with Sentence-BERT	F1 Score: Soft-Cosine with GloVe
80	80	0.873	0.665
	85	0.865	0.656
	90	0.921	0.671
	95	0.886	0.727
85	80	0.763	0.723
	85	0.868	0.748
	90	0.854	0.769
	95	0.899	0.778
90	80	0.784	0.684
	85	0.881	0.743
	90	0.886	0.693
	95	0.921	0.727
95	80	0.810	0.704
	85	0.827	0.766
	90	0.867	0.758
	95	0.857	0.760

F1 Scores of Cosine vs Soft-Cosine models with top-20 topic words, for 80, 85, 90, 95 topic clusters. 85-95 pair is selected for next step.

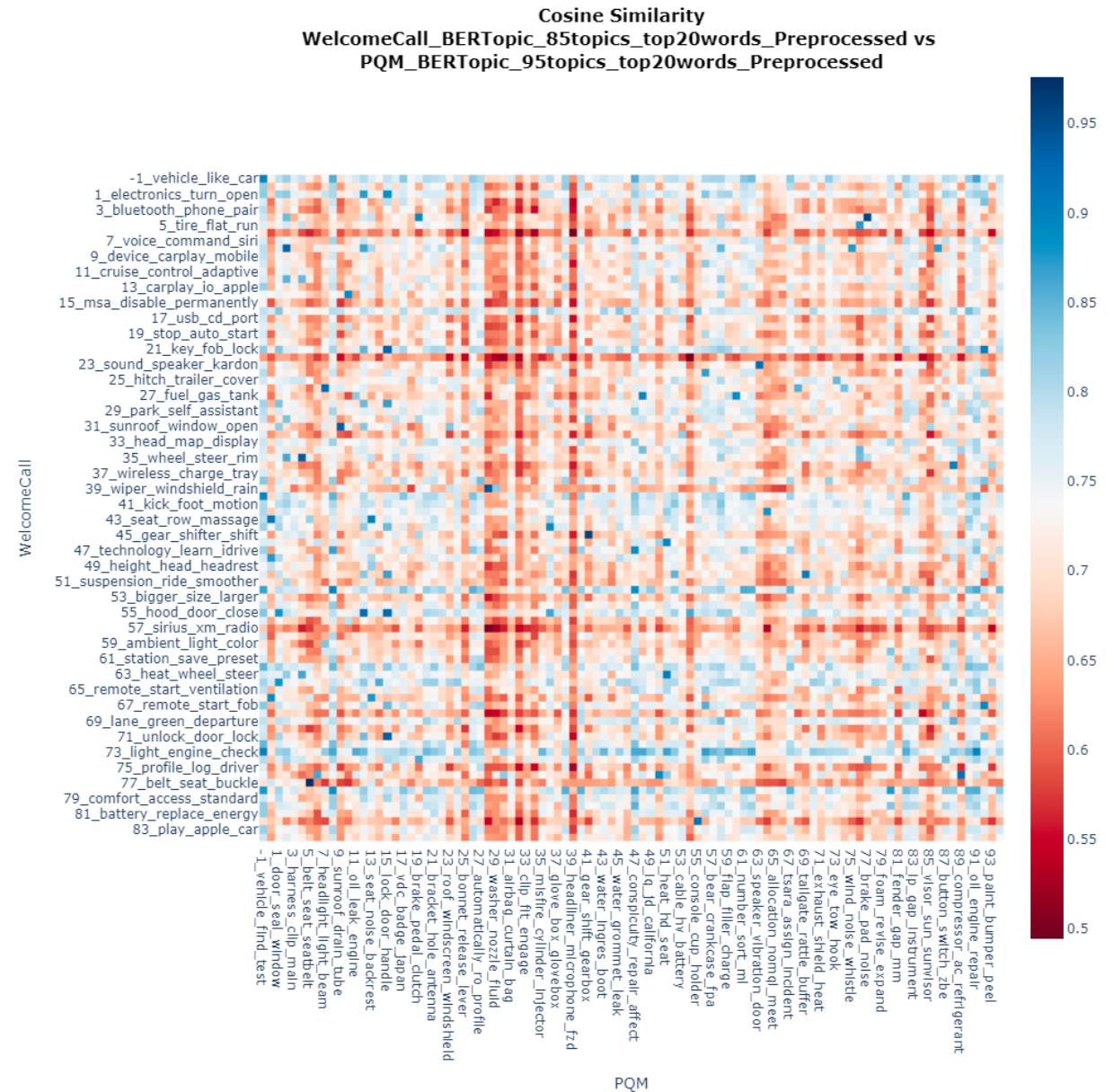
Experimental Results – Text Similarity

Heatmap

Welcome Call 85 topic clusters & PQM 95 topic clusters

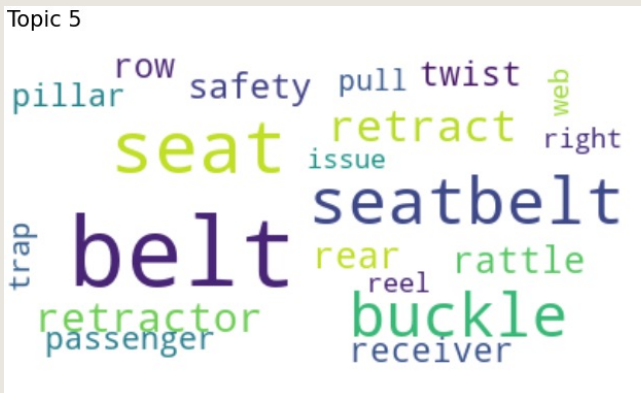
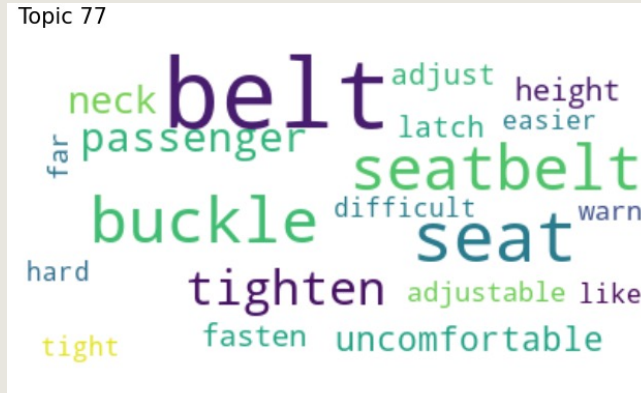
Blue = Similar topics

Red = Dissimilar topics



Welcome
Call

Match 1



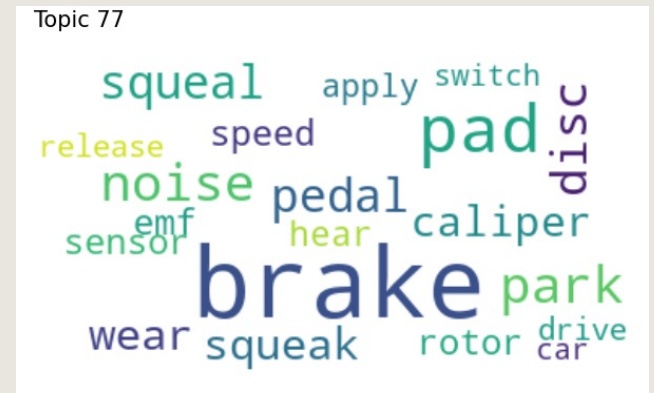
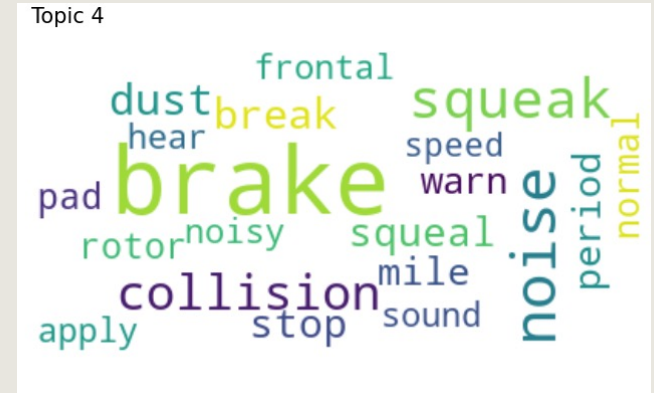
Seatbelt buckle rattling noise,
height adjustment

Match 2



Reverse gear shift in
manual/auto transmission

Match 3

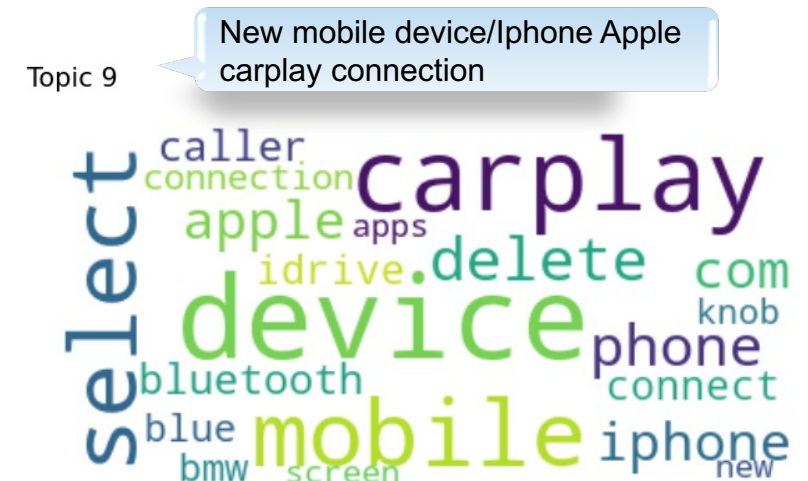
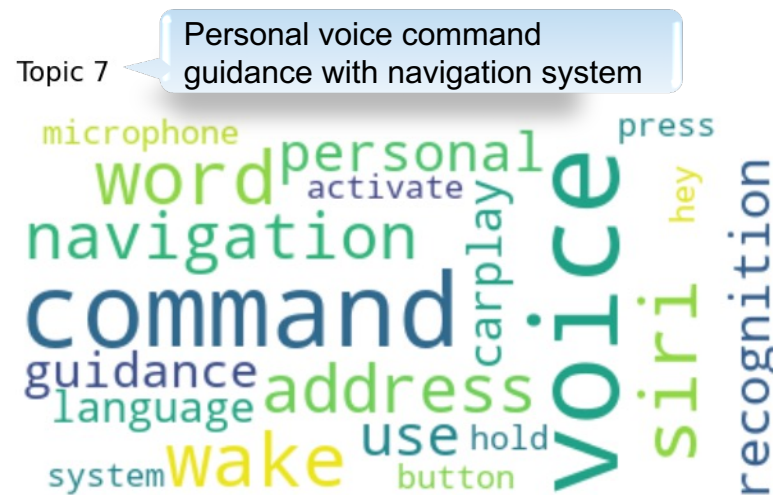
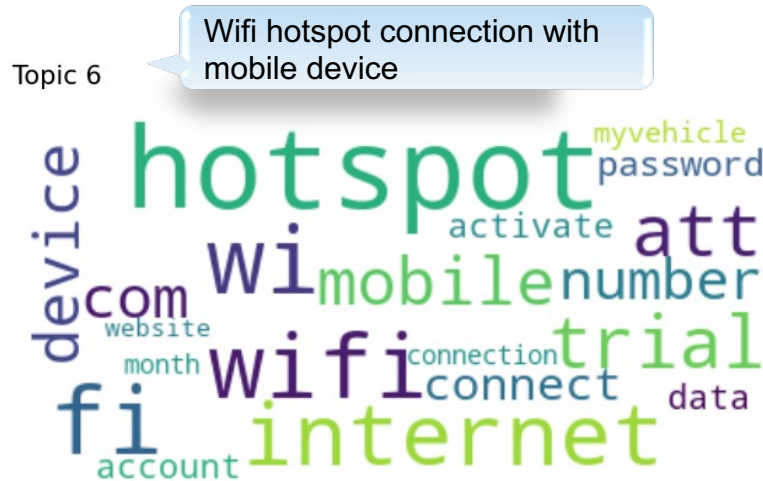
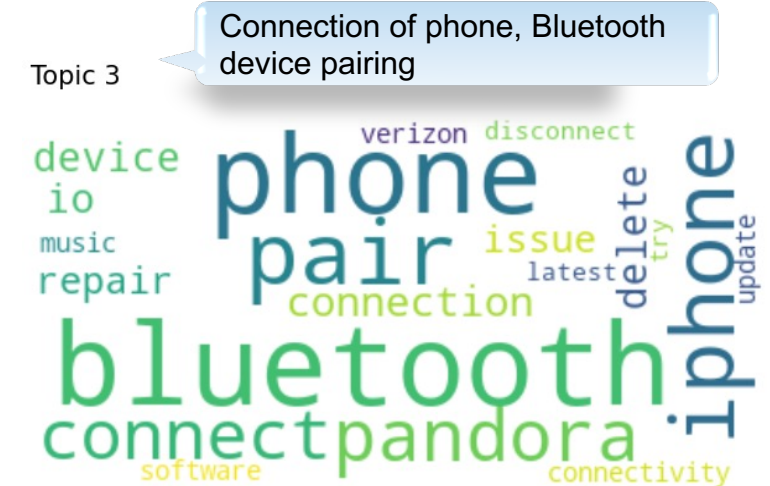
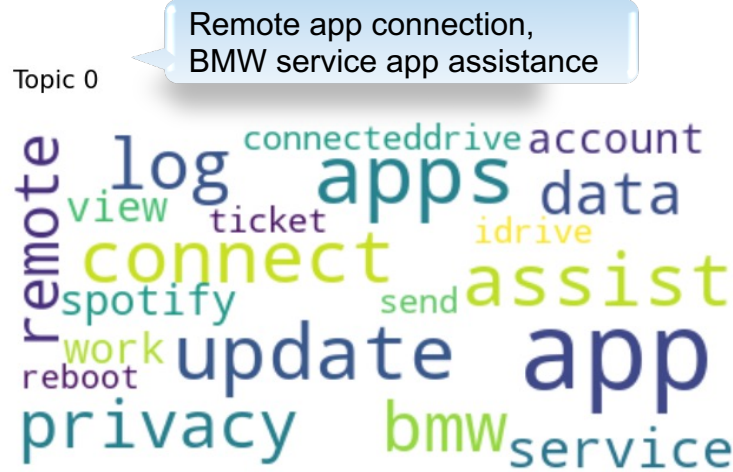


Squeaking noise coming from
brakes

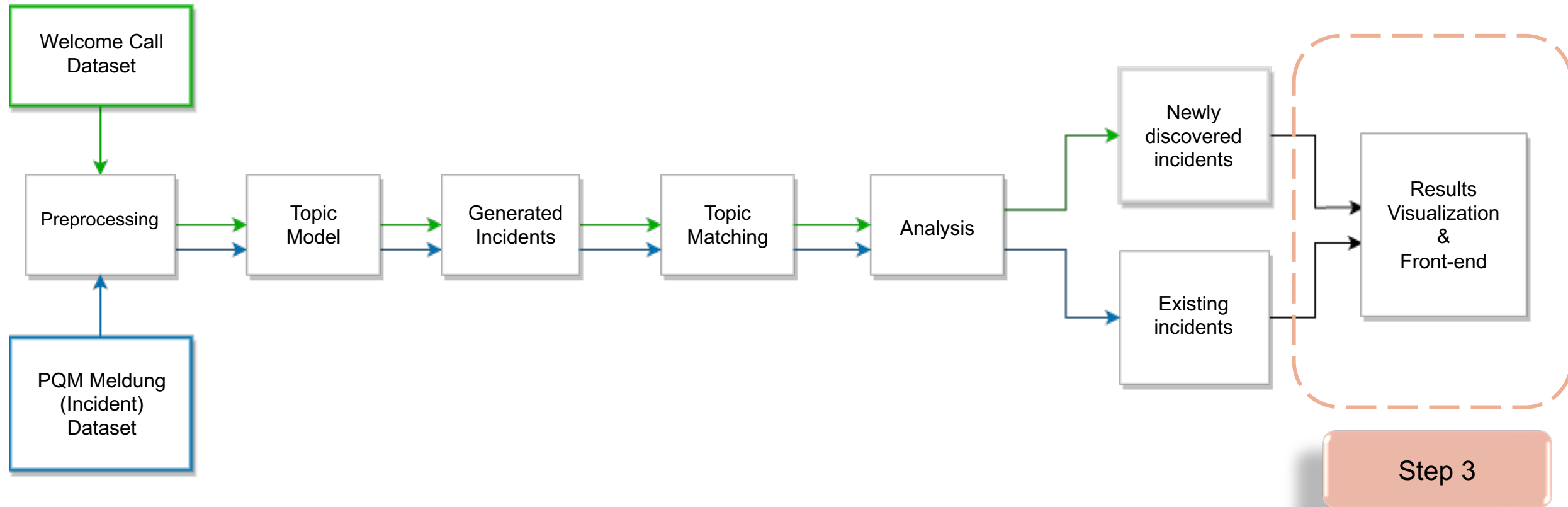
PQM

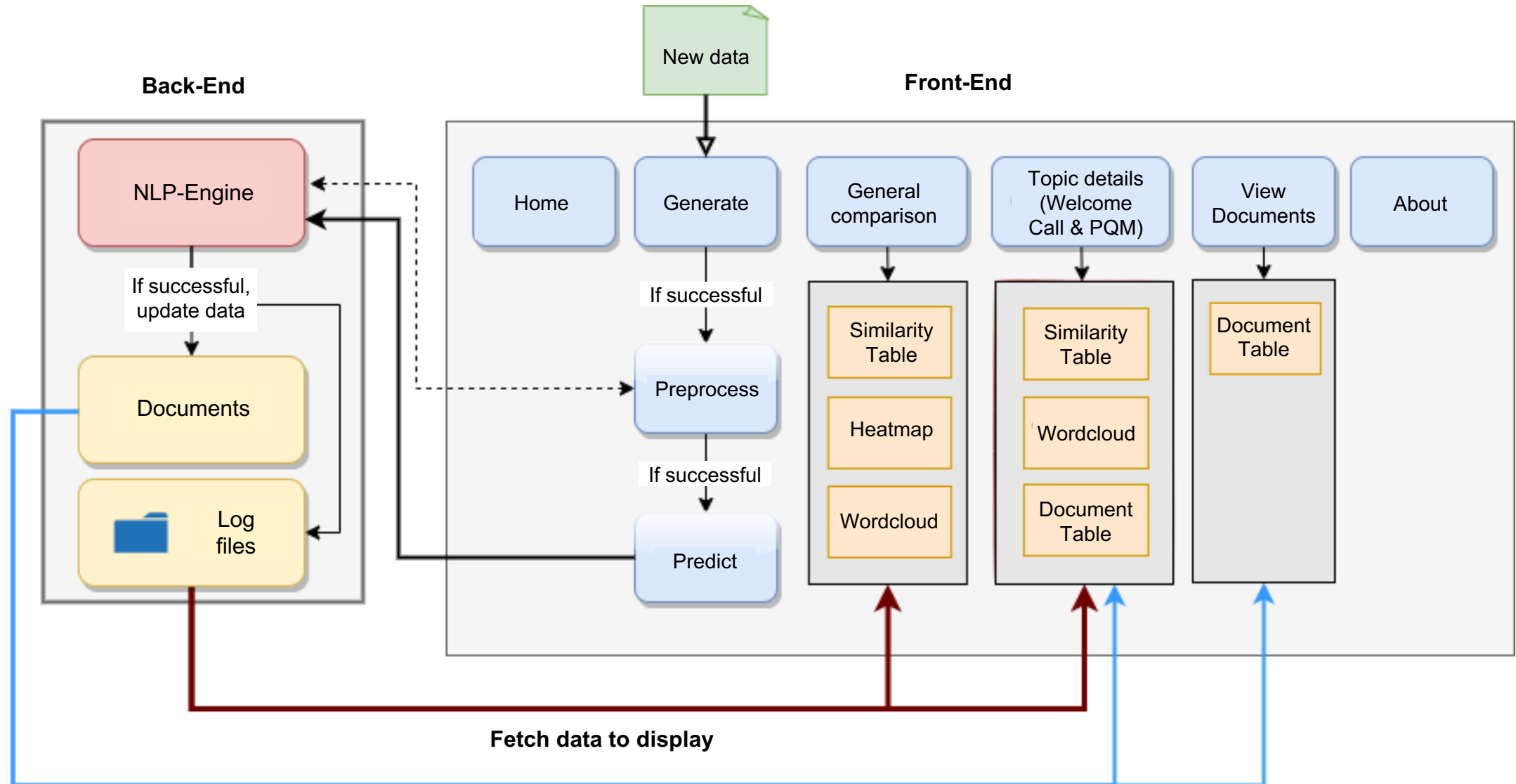
Results

Newly Discovered Issues



47 new issues were found in Customer Call dataset, that does not occur in the production data. (out of 85 topics)







Motivation



Research Questions



Background

Topic Modeling



Datasets



Methodology



Experiments

Step 1: Topic Modeling

Step 2: Text Similarity

Step 3: Visualizations & Front-End



System Architecture



Evaluation Results

User study with 14 statements, 4 aspects.

1-5 scale (Strongly Disagree – Strongly Agree)

Model Comparison
Evaluation



Views in General Comparison page are useful:

- Table View
- Heatmap View
- Word Cloud View

User Interface Quality



- Website is easy to use.
- Appearance of website
- Speed of loading results

Result Quality



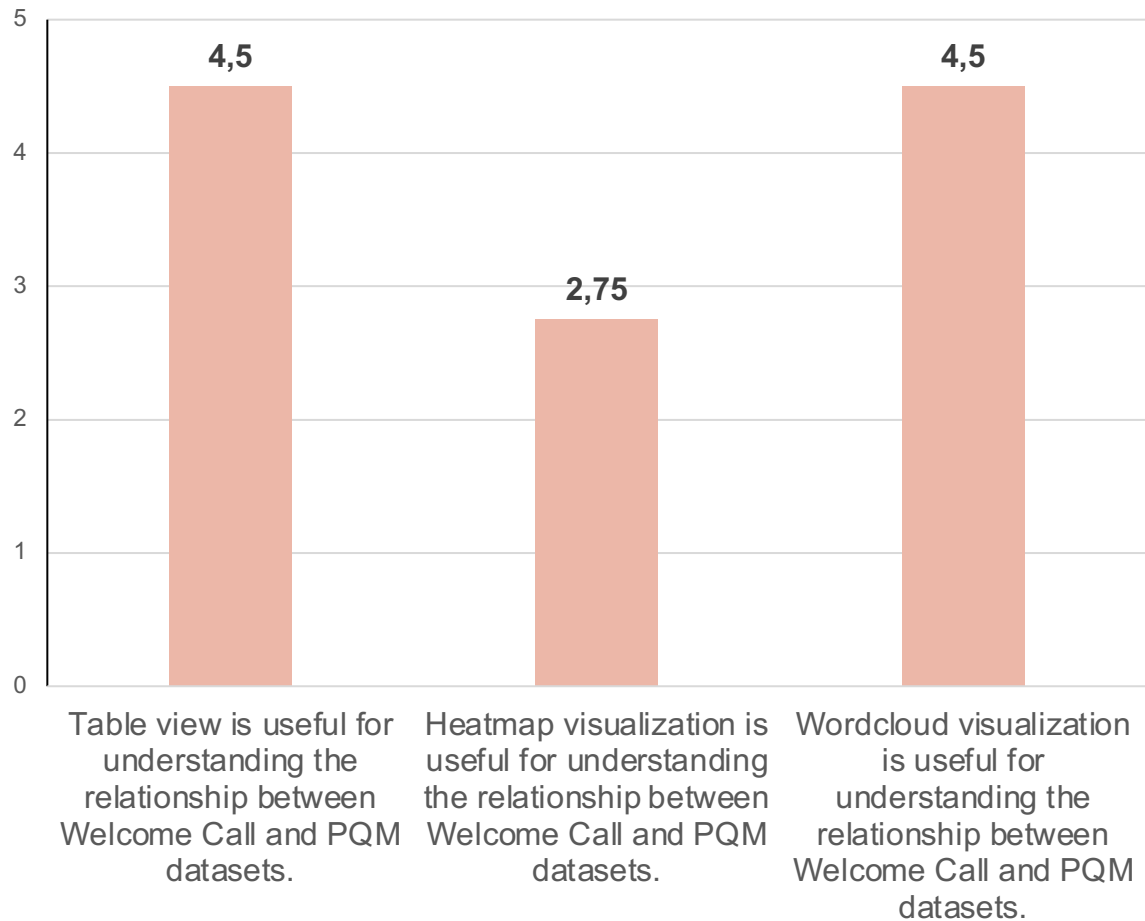
- Topic words of models made sense. – Welcome Call & PQM
- Document-wise topic predictions – Welcome Call & PQM
- Accuracy of similarity matching
- System is helpful to understand relationship between processes

Recommendations

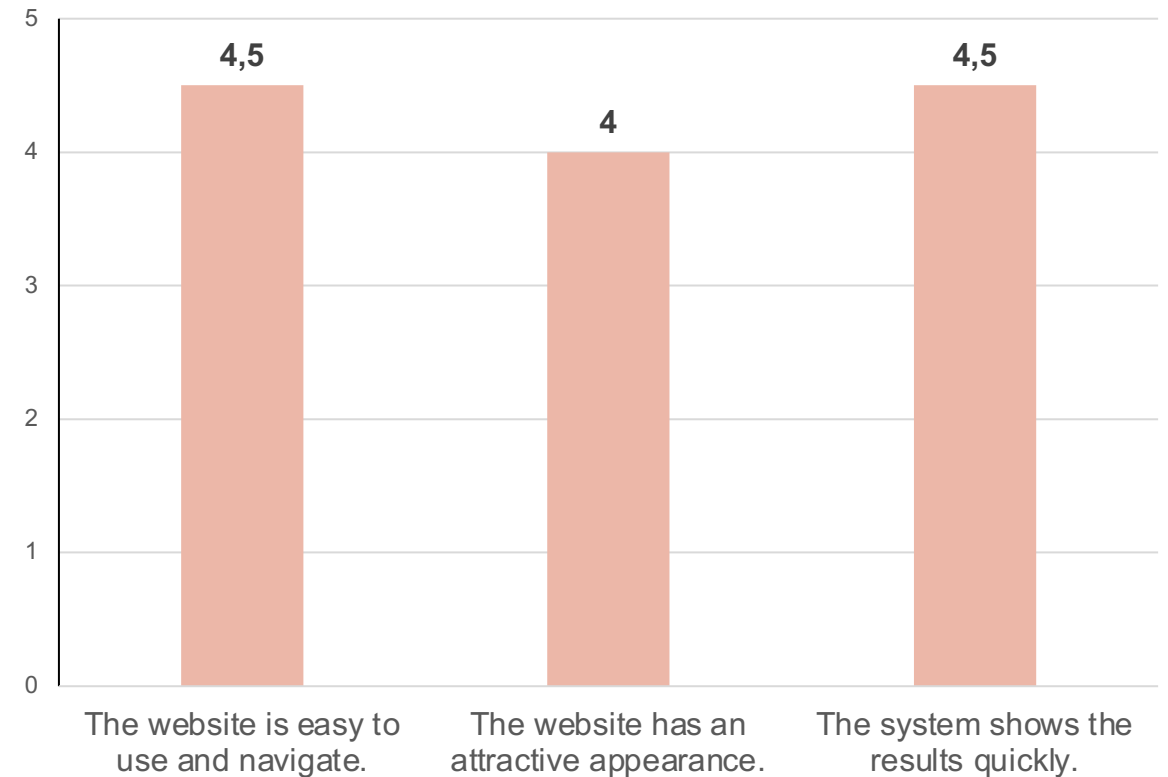


- Would they use it?
- Would they recommend it to colleagues?

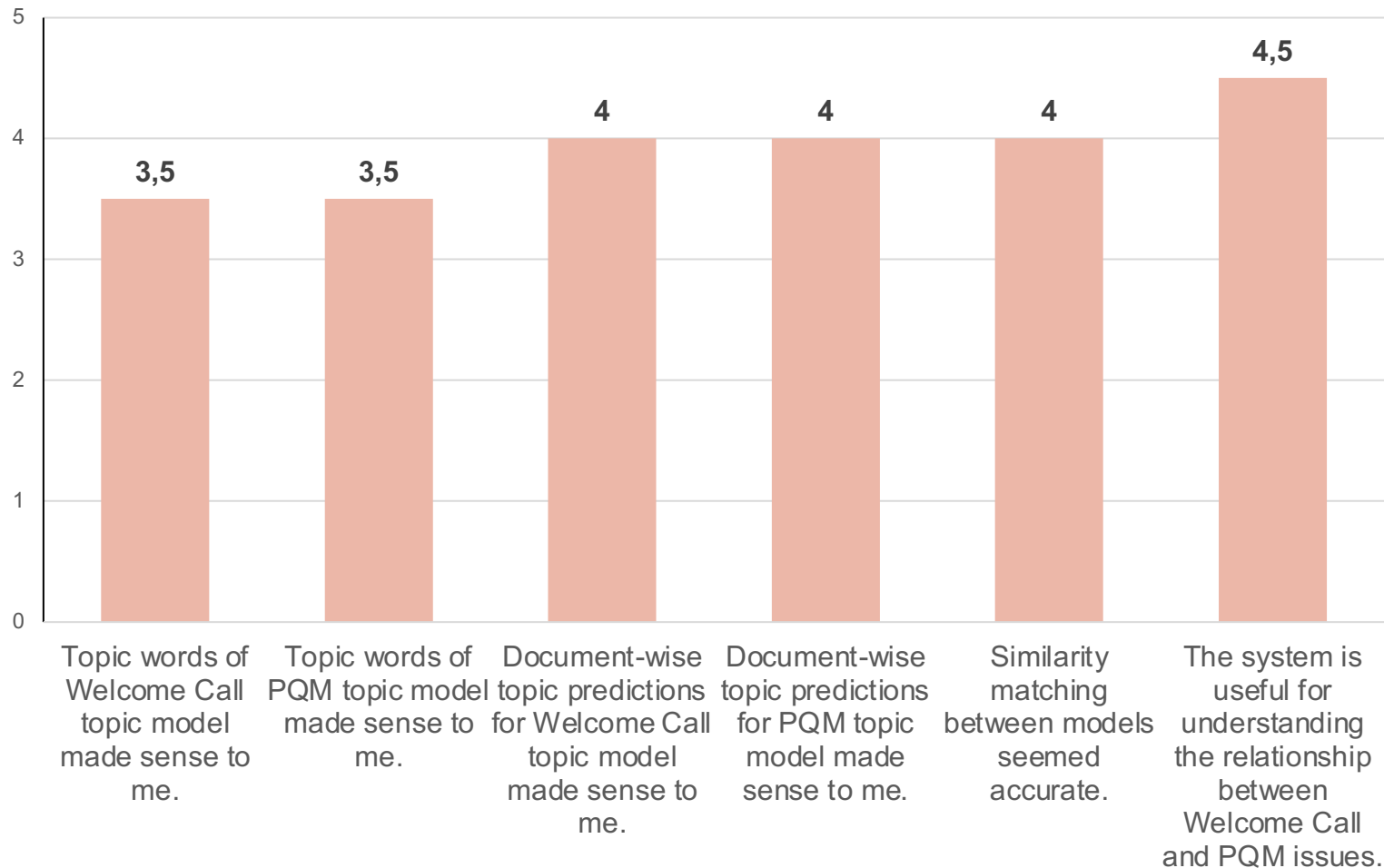
Model Comparison Evaluation



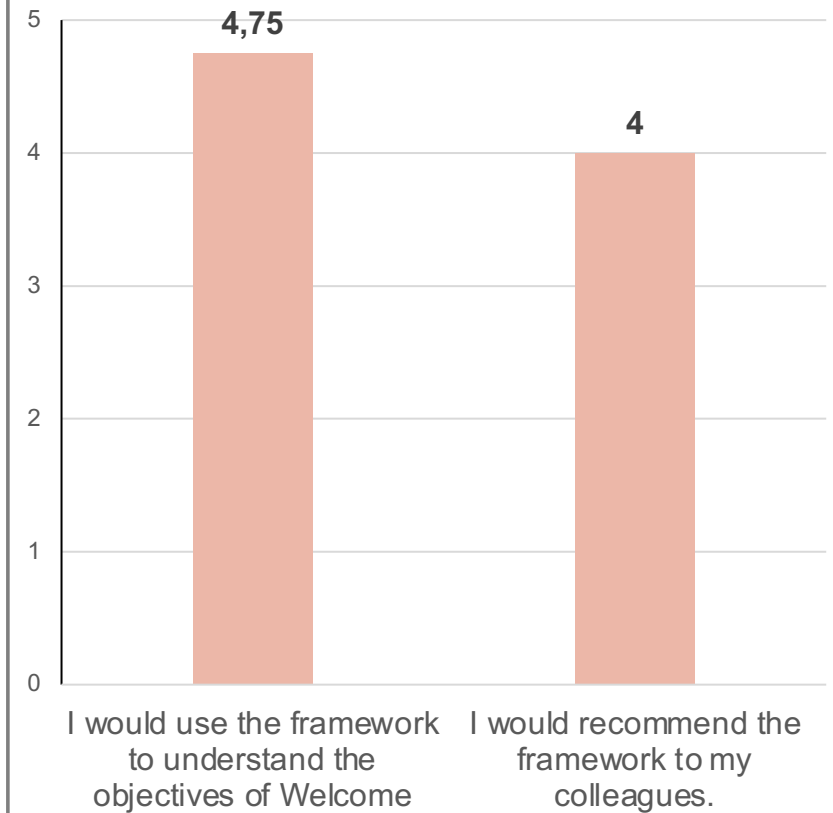
User Interface Quality Evaluation



Result Quality Evaluation



Recommendation Evaluation



1

Which state-of-the-art **topic modeling** approaches would provide better insight into BMW customer feedback datasets?

Embedding-based model BERTopic provides better insights into BMW customer feedback datasets by generating more domain-specific topics than **traditional model LDA** which generates more general topic representations.

2

Which **text similarity** techniques give better matches between BMW customer feedback datasets?

Cosine Similarity with contextual embeddings provided better topic matches over **Soft Cosine Similarity** with pretrained word embeddings.

3

How to support quality control departments with interactive topic **visualizations** of BMW customer feedback datasets?

Word clouds were useful for interpreting topic quality and representations.
Heatmaps were helpful for inspecting topic similarity.

Thank you for listening!



Prof. Dr.

Florian Matthes

Technische Universität München
Faculty of Informatics
Chair of Software Engineering for Business
Information Systems

Boltzmannstraße 3
85748 Garching bei München

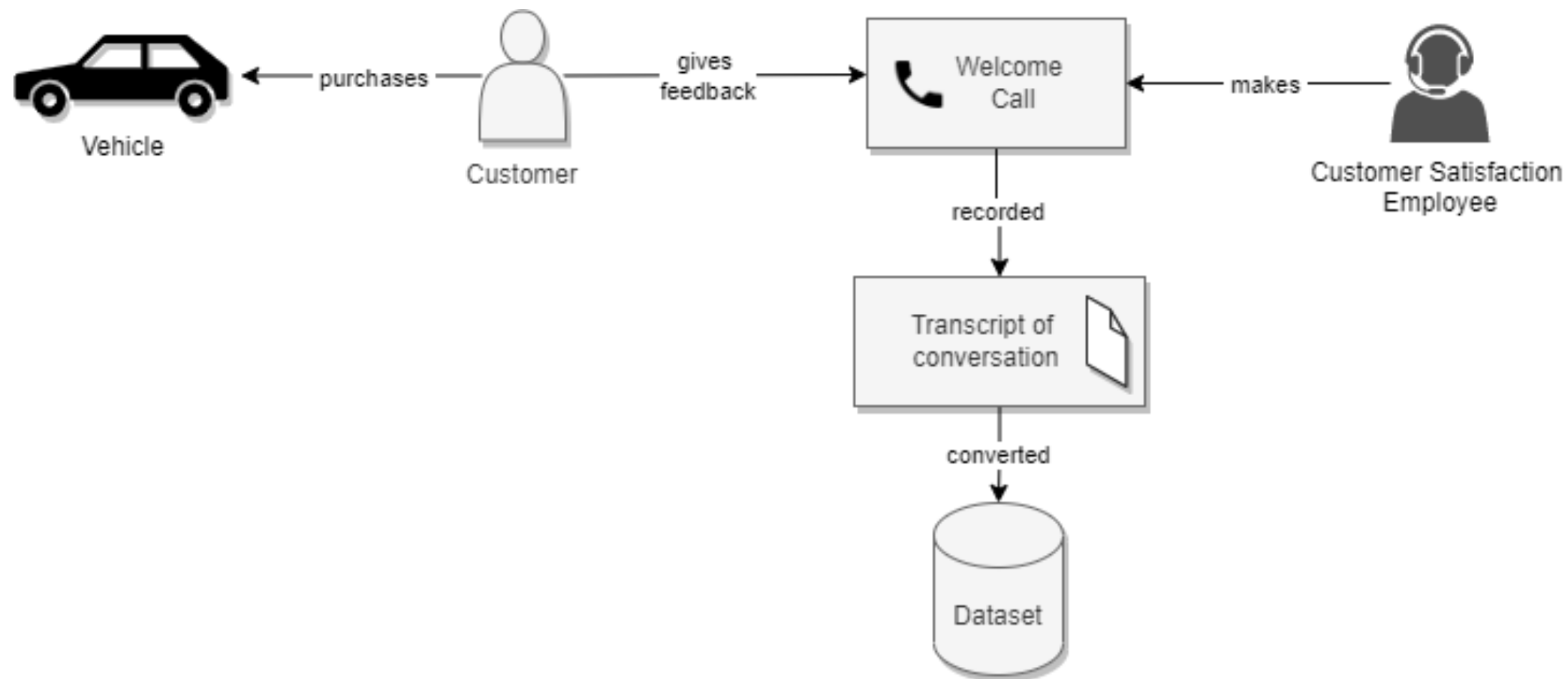
Tel +49.89.289. 17132

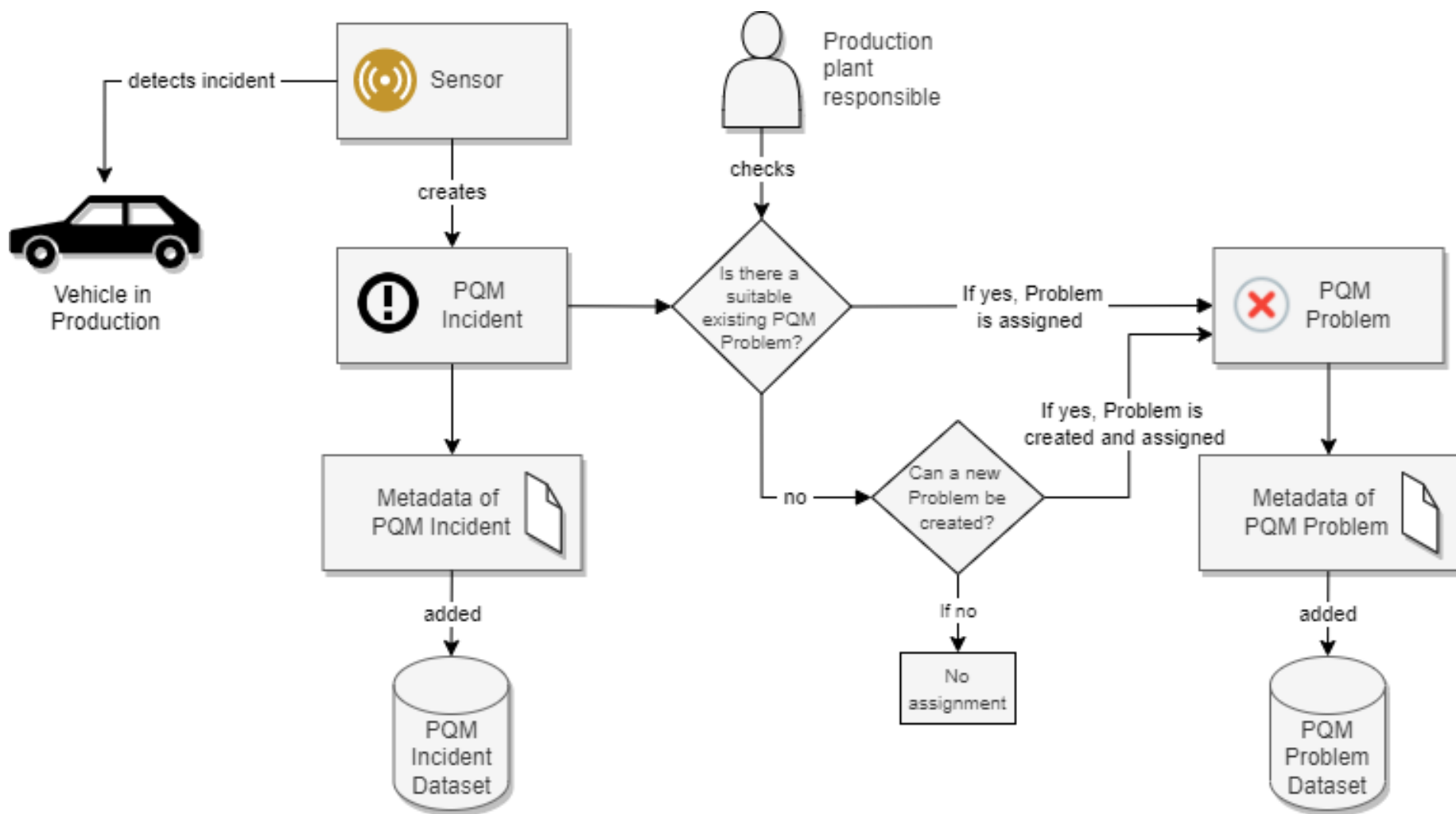
Fax +49.89.289.17136

matthes@in.tum.de
www.matthes.in.tum.de

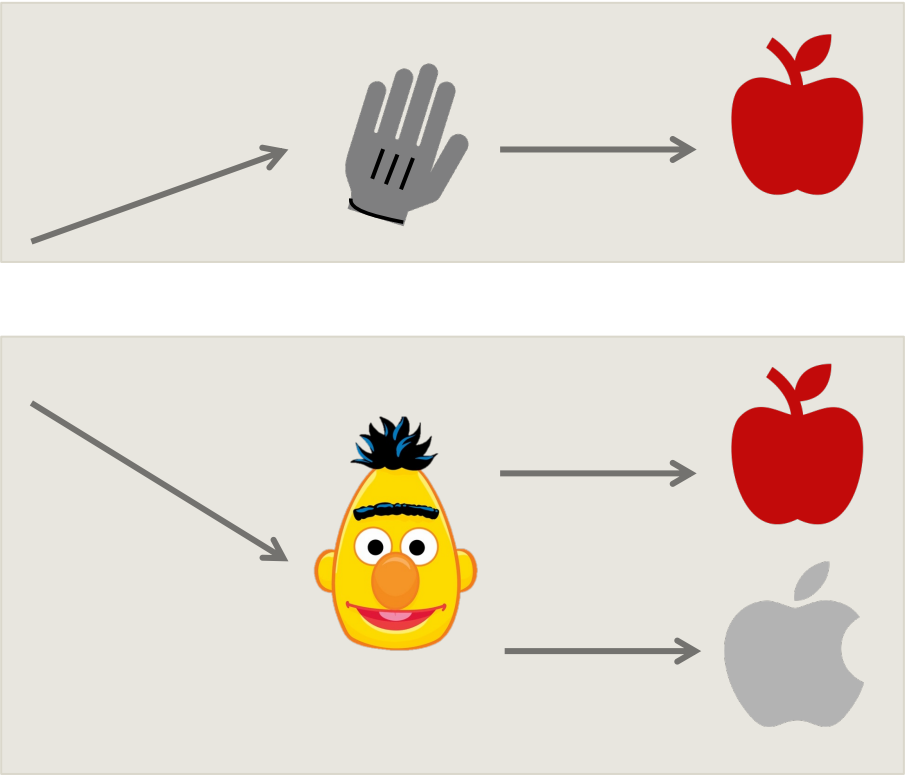
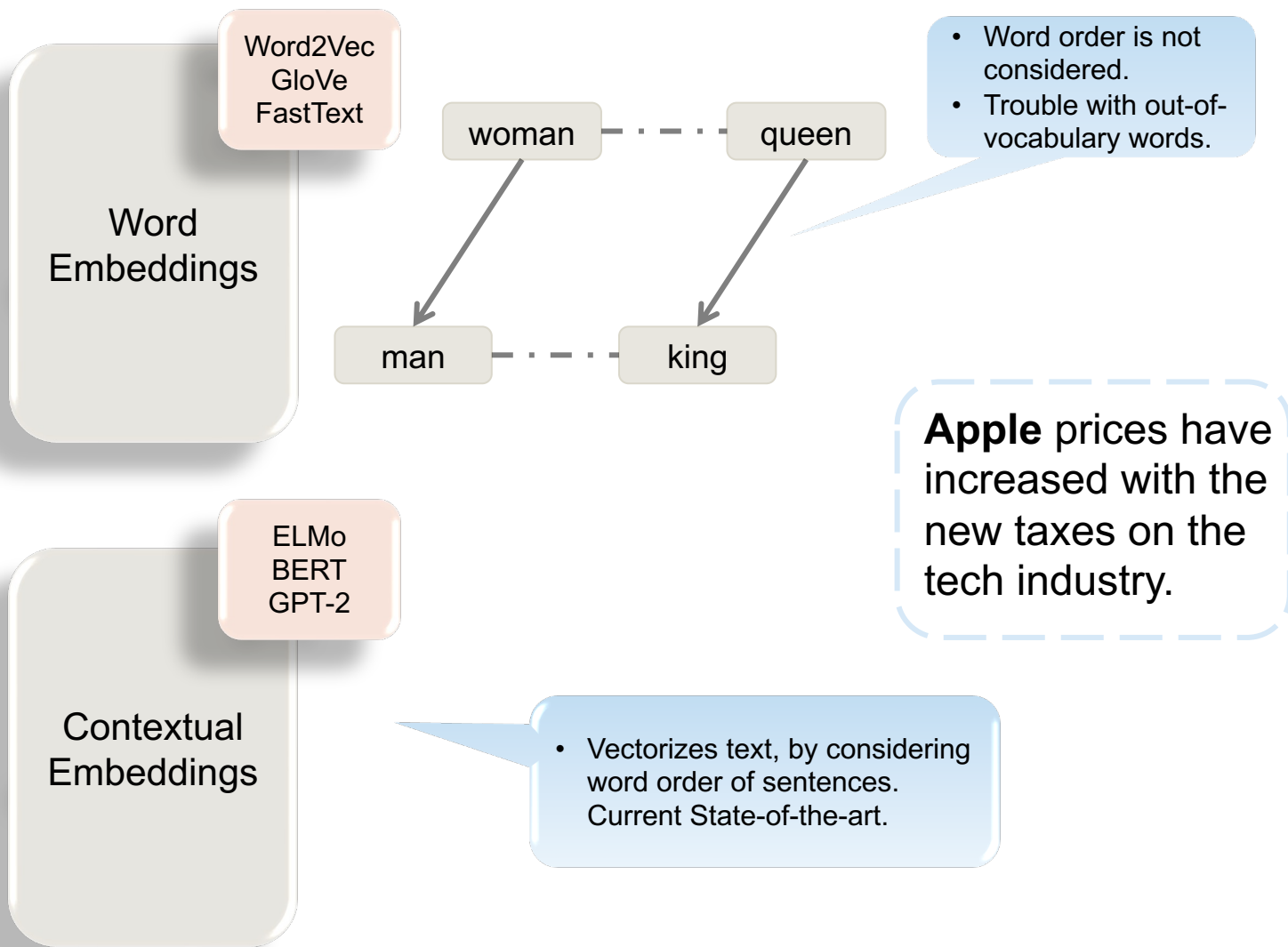


Bonus - Welcome Call Process





Word Embeddings vs. Contextual Embeddings



[GloVe] Pennington J., Socher R., and Manning C. 2014. Global Vectors for Word Representation

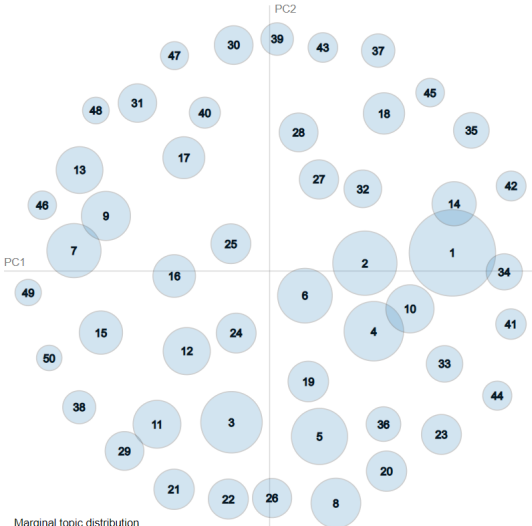
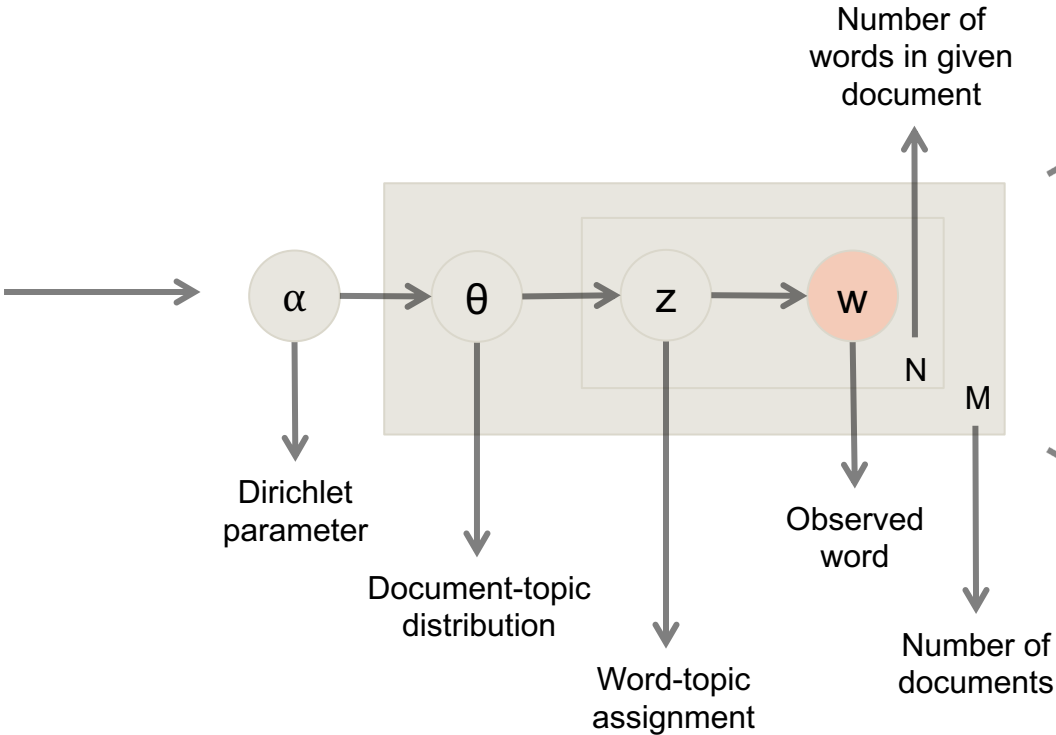
[BERT] Devlin, Jacob and Chang, Ming-Wei and Lee, Kenton and Toutanova, Kristina. 2018. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

Traditional Topic Model

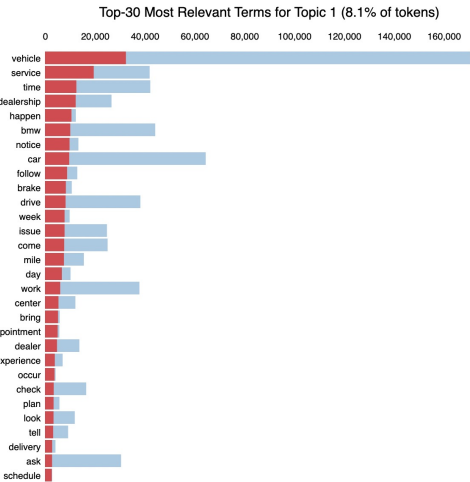
Bonus – Latent Dirichlet Allocation (LDA)



Collection of Documents



Topic representations / Clusters



Frequency of topics per documents

[LDA] Blei, David M. and Ng, Andrew Y. and Jordan Michale I. 2003. Latent Dirichlet Allocation.

Experimental Results – Topic Modeling

Welcome Call – 75 topic clusters (top-20-words)

	LDA	BERTopic
Intertopic distance map		
Topic words: <i>mirror</i>	mirror , automatically, inside, fold, view, stick, switch, <i>like</i> , vehicle, passenger, want, curb, soft, exterior, position, manually, bar, lift, <i>able</i> , <i>countryman</i>	fold, mirror , curb, tilt, passenger, monitor, reverse, automatically, exterior, lock, automatic, switch, position, seat, driver, view, feature, door, deactivate, button
Umass coherence score	-4.49	-3.12

Experimental Results – Topic Modeling

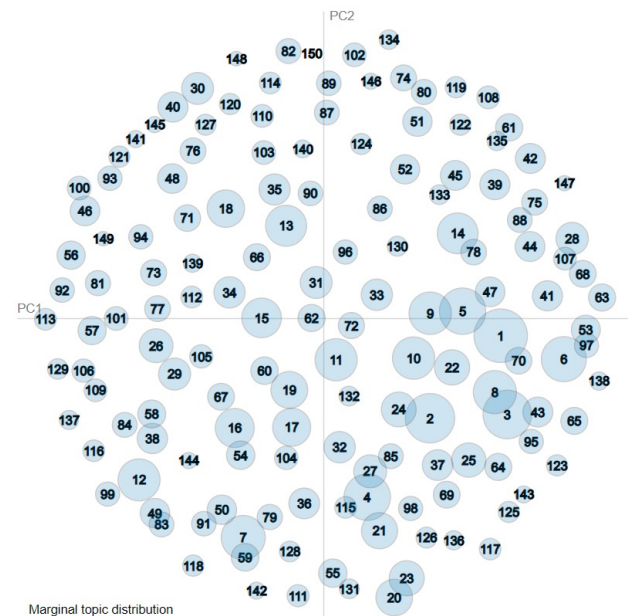
Welcome Call – 150 topic clusters (top-20-words)

Intertopic distance map

Topic words: *mirror*

Umass coherence score

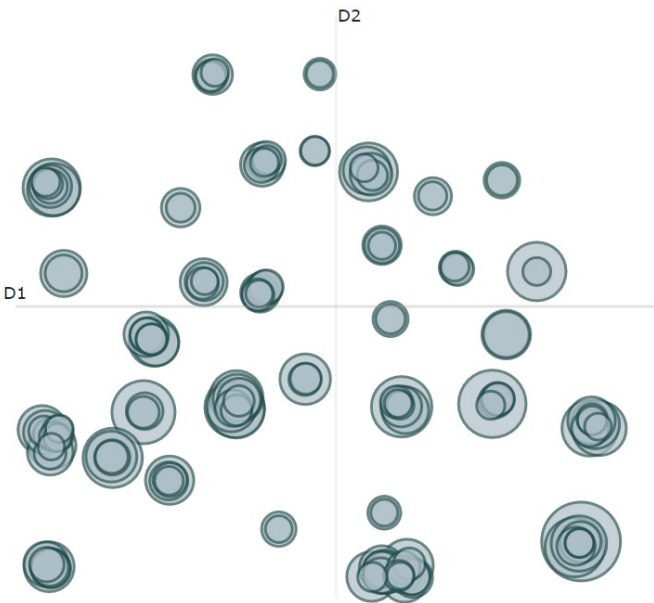
LDA



mirror, heat, position, utilize, tilt, curb, passenger, resource, slide, like, automatic, *briefly*, *ask*, backrest, distribution, vehicle, able, *september*, *use*, look

-6.76

BERTopic



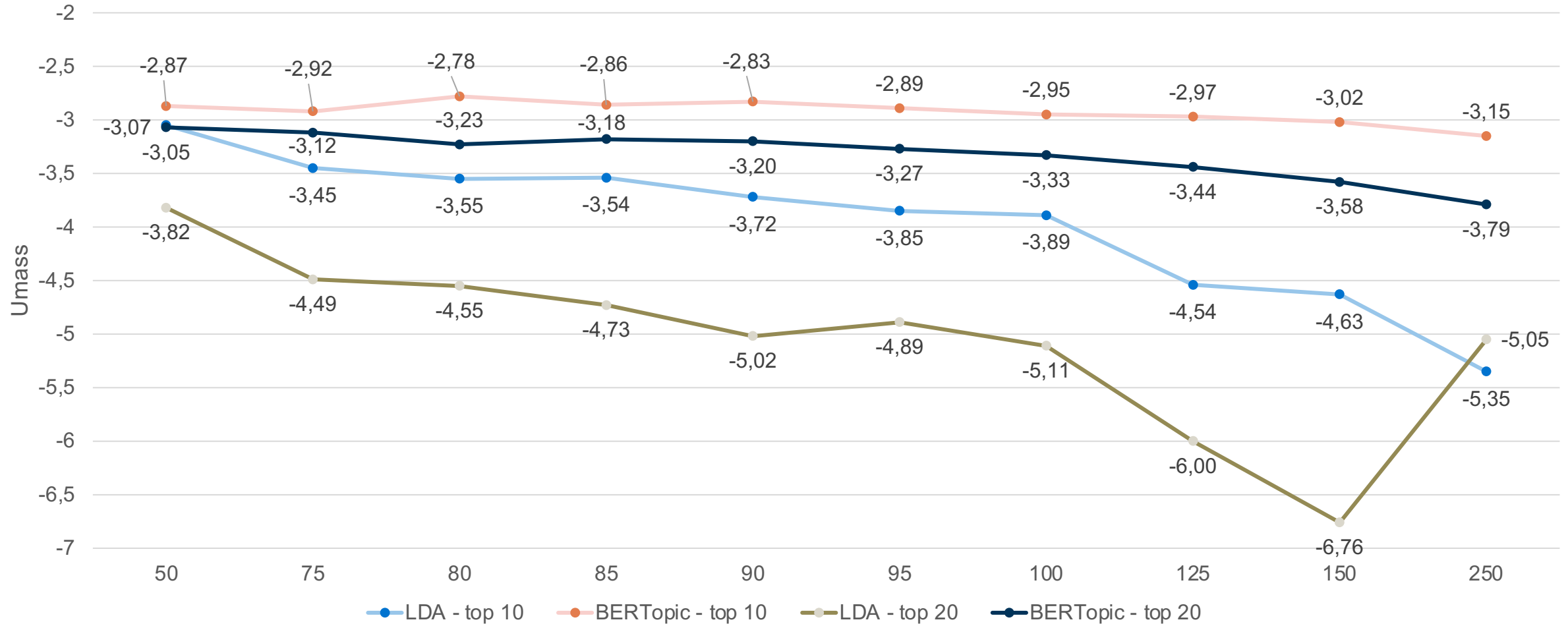
- fold, **mirror**, automatically, lock, exterior, seat, unfold, row, rear, captain, cargo, chair, door, access, button, automatic, manually, want, flat, power
- **mirror**, screen, entertainment, android, video, iphone, watch, apple, rear, iphones, compatible, stream, hdmi, miracast, device, use, youtube, carplay, phone, able

-3.58

Experimental Results – Topic Modeling

Bonus – Analytical Measures – Umass Coherence

Topic Model Comparison for Welcome Call dataset

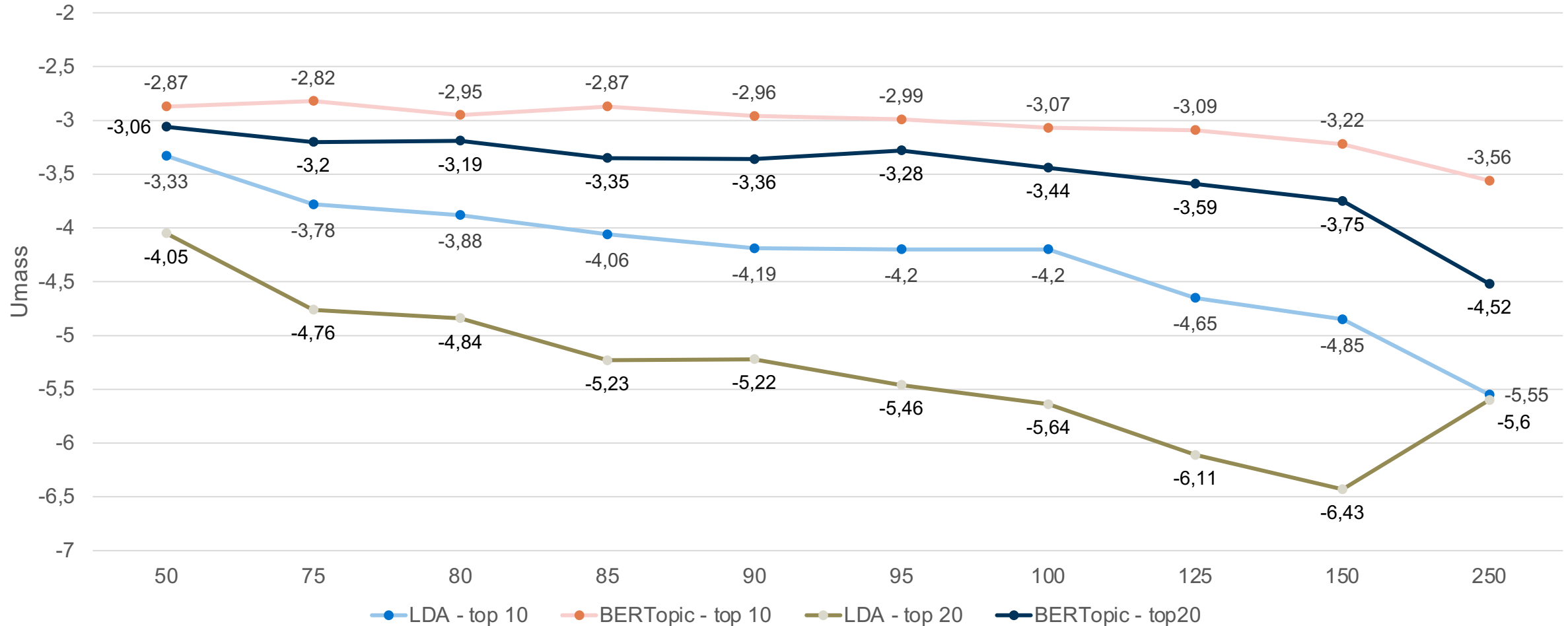


The closer Umass score is to zero the better the topic coherence of the model.

Experimental Results – Topic Modeling

Bonus – Analytical Measures – Umass Coherence

Topic Model Comparison for PQM dataset



The closer Umass score is to zero the better the topic coherence of the model.

Experimental Results – Topic Modeling

Bonus – Analytical Measures – Umass Coherence



Number of topic clusters	LDA – Welcome Call	BERTopic – Welcome Call	LDA – PQM	BERTopic – PQM
50	-3.82	-3.07	-4.05	-3.06
75	-4.49	-3.12	-4.76	-3.20
80	-4.55	-3.23	-4.84	-3.19
85	-4.73	-3.18	-5.23	-3.35
90	-5.02	-3.20	-5.22	-3.36
95	-4.89	-3.27	-5.46	-3.28
100	-5.11	-3.33	-5.64	-3.44
125	-6.00	-3.44	-6.11	-3.59
150	-6.76	-3.58	-6.43	-3.75
250	-5.05	-3.79	-5.60	-4.52

Umass scores for LDA and BERTopic models with top-20 topic words, for each topic cluster number.
The closer Umass score is to zero the better the topic coherence of the model.

Experimental Results – Text Similarity

Bonus – Similarity Table



Similarity Score	Welcome Call topics	PQM topics	Label	Prediction
0.976	77 belt seat buckle seatbelt tighten passenger neck uncomfortable height fasten latch adjust adjustable warn far difficult easier hard tight like	5 belt seat seatbelt buckle retract retractor rear rattle safety row receiver twist pillar passenger trap pull right web reel issue	Match	Match
0.962	45 gear shifter shift paddle selector reverse park transmission forward gearshift mode button manual use lever push backwards drive neutral drivelogic	41 gear shift gearbox shifter transmission selector lever knob select gws reverse clutch gearshift auto egs engage drive manual park test	Match	Match
0.895	52 offer learn assistance question technology decline primary carplay apple genius reach dealership figure moment wife appointment need encore feature assist	8 repair history conspicuity affect case perception content request sa component german pre entry title comment single perform additional code available	Non-Match	Match
0.892	14 fold mirror curb tilt passenger monitor reverse automatically exterior lock automatic switch position driver seat view feature door deactivate unfold	10 mirror exterior view fold door cap switch glass outside rearview leave triangle base right function driver noise interior adjust rear	Match	Match
0.888	66 seat support lumbar thigh comfortable uncomfortable extender leg bolster cushion like wider adjustable sport lever little adjust feel softer extend	0 seat backrest leather cushion headrest rear armrest crease bolster wavy cover rest audit wrinkle row leave stitch visible gap issue	Match	Match
0.886	47 technology learn idrive intuitive offer assistance complicate menu user question friendly decline interface system specific controller wish navigate overwhelm difficult	47 conspicuity repair affect history case bluetooth period code selective sa cost component country far perception content perform entry single step	Non-Match	Match
0.880	64 trim grill kidney piece black grille door fiber carbon loose strip plastic chrome rubber mold seal glare wood passenger rear	1 door seal window glass frame gap finisher rear panel pillar trim waist guide rubber quarter lh fit chrome rh inner	Non-Match	Match
0.875	48 leather seat interior upholstery jean material stitch picture stain look passenger wrinkle cognac clean plastic quality color notice vernasca dealership	0 seat backrest leather cushion headrest rear armrest crease bolster wavy cover rest audit wrinkle row leave stitch visible gap issue	Match	Match
0.776	15 msa disable permanently default feature stop sport turn mode start like function shut auto engine button basis emission trip able	47 conspicuity repair affect history case bluetooth period code selective sa cost component country far perception content perform entry single step	Non-Match	Non-Match
0.752	53 bigger size larger smaller visibility mirror small like little wider rear body view wish style big look screen bite design	1 door seal window glass frame gap finisher rear panel pillar trim waist guide rubber quarter lh fit chrome rh inner	Non-Match	Non-Match

Detail topic view for Welcome Call model

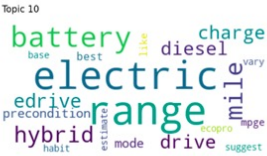
Select Topic no to display for Welcome Call model:
10_range_electric_battery_mile_hybrid_drive_drive_charge_diesel

Apply

Selected topic id is 10.

Scroll for more details.

Topic 10



4.0 number of documents are classified as Topic 10.

Similarity scores of PQM model topics with the selected topic:

This table lists each topic from PQM model and shows the similarity with the selected topic. (based on Cosine Similarity metric).

Show 5 entries

Search:

Similarity Score	Topic id	Topic words	Document frequency
0.822827	56	starter, start, engine, eml, light, power, motor, illuminate, stall, check, warn, crank, fault, breakdown, vehicle, run, customer, lamp, msa, loss	6.0
0.7806	47	conspicuity, repair, affect, history, case, bluetooth, period, code, selective, sa, cost, component, country, far, perception, content, perform, entry, single, step	1.0
0.778626	62	production, month, kmgerman, vin, mileage, title, state, km, date, request, case, perform, replace, component, work, single, detail, performance, ri, create	1.0
0.777766	51	heat, hd, seat, test, cycle, temperature, disassembly, disturbance, ptce, run, damage, rotor, htoc, circuit, detect, lin, heater, analysis, machine, mech	1.0
0.770197	35	misfire, cylinder, injector, plug, spark, injection, coil, ignition, gas, pan, perform, exhaust, fault, engine, breakdown, memory, german, test, entry, title	26.0

Similar topics from PQM model to Welcome Call model topic 10

Showing 1 to 5 of 95 entries

Previous 1 2 3 4 5 ... 19 Next

Topic Documents

Left table shows documents belonging to selected topic 10 from Welcome Call documents.

Right table shows documents from the closest matched topic from PQM model. Both tables are sortable and searchable.

Show 5 entries

Search:

Welcome Call Documents for Topic 10

van17

Welcome Call Document

Customer stated when driving mode should he be in to only use the . Writer advised to put the vehicle in . which allows the vehicle to stop using . Writer advised putting the vehicle in sport mode will use a lot of the combustion engine. Customer stated what driving mode should he normally be in? Writer advised to maximize fuel efficiency, use the mode or adaptive driving mode. Writer advised if the customer is worried about accelerating in mode, advised to push all the way down on the accelerator pedal to activate .

Customer stated that he had read on some Forums that people would get more than previous days and inquired on what that meant. Writer advised that that is presented is an estimated guess after being fully charged. Writer advised that the estimation is based off the .

Show 5 entries

Search:

PQM Documents for Topic 56

PQM Problem Number

PQM Document

The engine warning lamp lit up.

Engine and engine light on.

Customer states he went to start the bike up and gas started pouring out.

Check engine not starting.

Showing 1 to 5 of 6 entries

Previous 1 2 Next

Uploaded data is used for PQM model. Uploaded data is used for Welcome Call model.

Topic pairs have been classified into three categories for better understanding: **Definite matches**, **Possible matches** and **Non matches**.

Table

Heatmap

Wordcloud

Wordcloud view of matched topics

You can navigate through three tabs and observe wordclouds for the topic match pairs.

Definite Matches

Possible Matches

Non Matches

Welcome Call

PQM

Topic 77

Topic 5



Welcome Call

PQM

Topic 45

Topic 41



Uploaded data is used for PQM model. Uploaded data is used for Welcome Call model.

221107 Dilara Ademoglu Final Presentation Master's Thesis

© sebis

40