Topics in Opinion Mining

Dr. Paul Buitelaar
Data Science Institute, NUI Galway

Opinion: Sentiment, Emotion, Subjectivity

After a slide by Carlos Iglesias, UPM
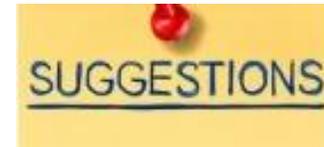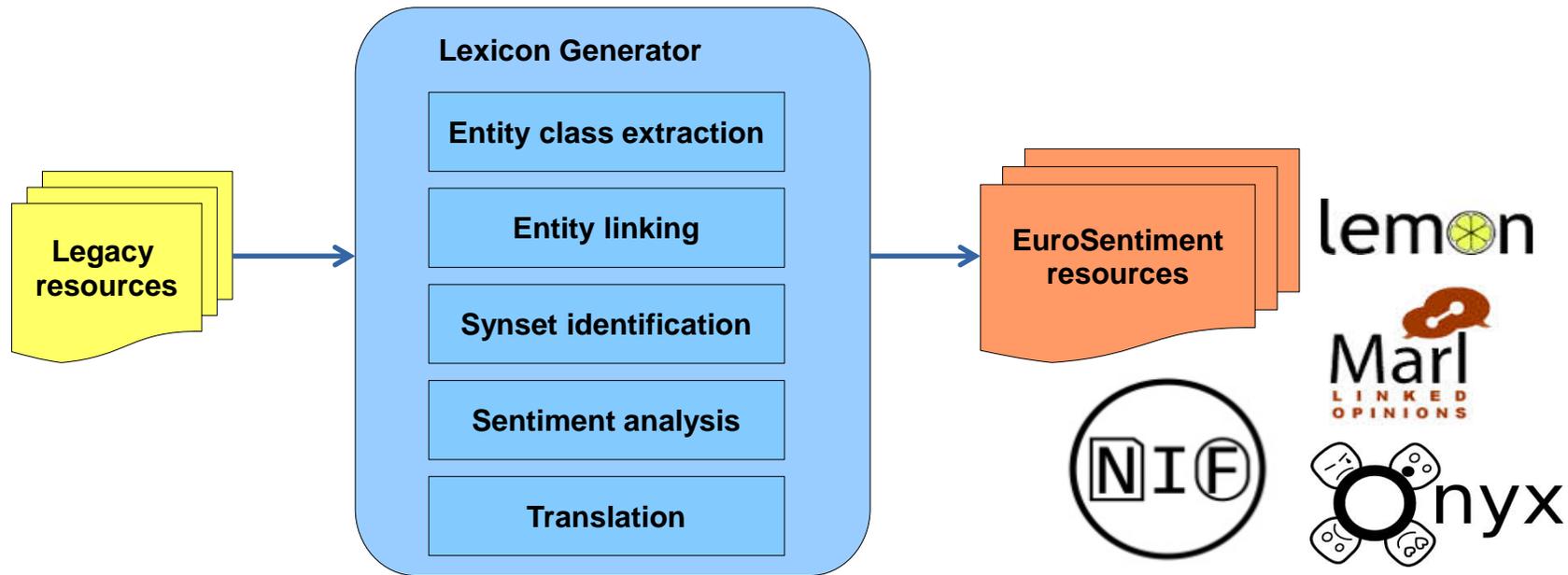
Language Resource Pool for Sentiment Analysis in European Languages

Services	Resources

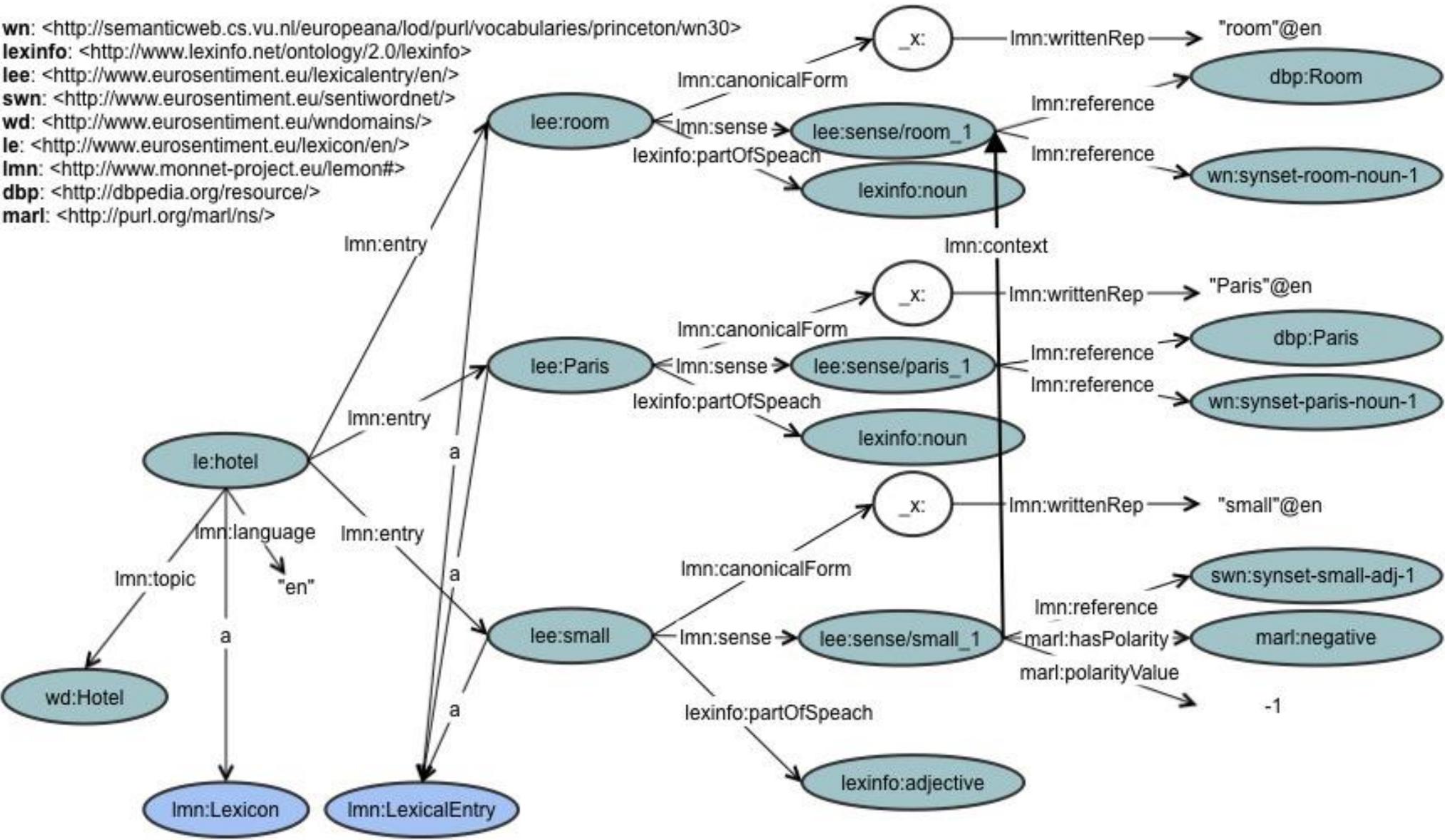| Name	Description	Language	Domain	Free
13919 Catalan human-annotated triplets - Hotel	13919 Catalan triplets human-annotated triplets for domain of hotel with sentiment score	Catalan	Hotels	Yes
22599 English human-annotated triplets - Hotel	22599 English triplets human-annotated triplets for domain of hotel with sentiment score	English	Hotels	Yes
8595 Spanish human-annotated triplets - Hotel	8595 Spanish triplets human-annotated triplets for domain of hotel with sentiment score	Spanish	Hotels	Yes
2648 French human-annotated triplets - Hotel	2648 French triplets human-annotated triplets for domain of hotel with sentiment score	French	Hotels	Yes
15369 Italian human-annotated triplets - Hotel	15369 Italian triplets human-annotated triplets for domain of hotel with sentiment score	Italian	Hotels	Yes

Generate Sentiment Lexicons

Domain-specific, multilingual, Linked Data based

Gabriela Vulcu, Paul Buitelaar, et al. (2014) **Generating Linked-Data based Domain-Specific Sentiment Lexicons from Legacy Language and Semantic Resources** 5th International Workshop on EMOTION, SOCIAL SIGNALS, SENTIMENT & LINKED OPEN DATA, co-located with LREC 2014.

wn: <http://semanticweb.cs.vu.nl/europeana/lod/purl/vocabularies/princeton/wn30>
lexinfo: <http://www.lexinfo.net/ontology/2.0/lexinfo>
lee: <http://www.eurosentiment.eu/lexicalentry/en/>
swn: <http://www.eurosentiment.eu/sentiwordnet/>
wd: <http://www.eurosentiment.eu/wndomains/>
le: <http://www.eurosentiment.eu/lexicon/en/>
lmn: <http://www.monnet-project.eu/lemon#>
dbp: <http://dbpedia.org/resource/>
marl: <http://purl.org/marl/ns/>

wn: <http://semanticweb.cs.vu.nl/europeana/lod/purl/vocabularies/princeton/wn30>
lee: <http://www.eurosentiment.eu/lexicalentry/en/>
led: <http://www.eurosentiment.eu/lexicalentry/de/>
wd: <http://www.eurosentiment.eu/wndomains/>
le: <http://www.eurosentiment.eu/lexicon/en/>
ld: <http://www.eurosentiment.eu/lexicon/de/>
lmn: <http://www.monnet-project.eu/lemon#>
dbp: <http://dbpedia.org/resource/>
marl: <http://purl.org/marl/ns/>
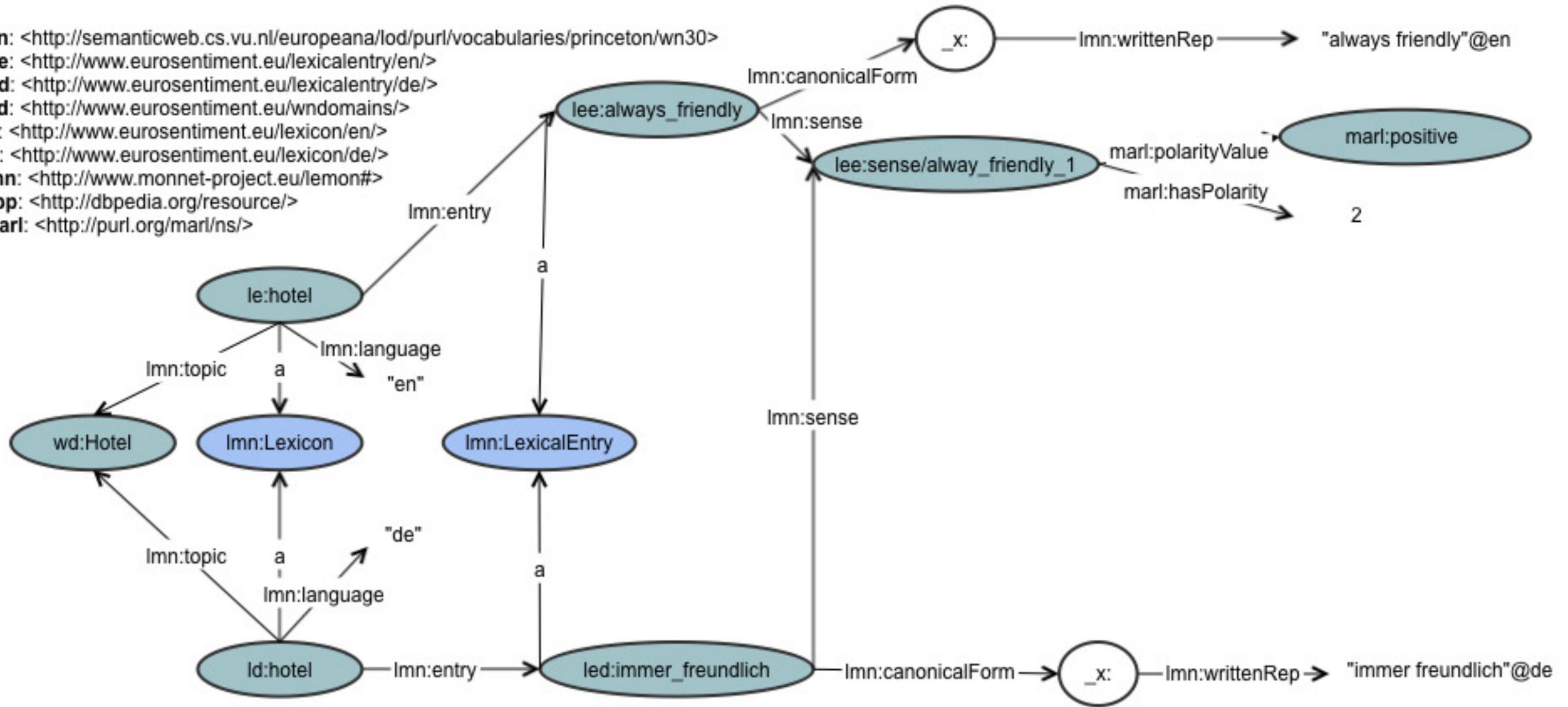
Social Semantic Emotion Analysis for Innovative Multilingual Big Data Analytics Markets

Beta

MixedEmotions

⚙ Modules   ⚗ Example Use Cases   📖 Documentation   ☰ Run a Task
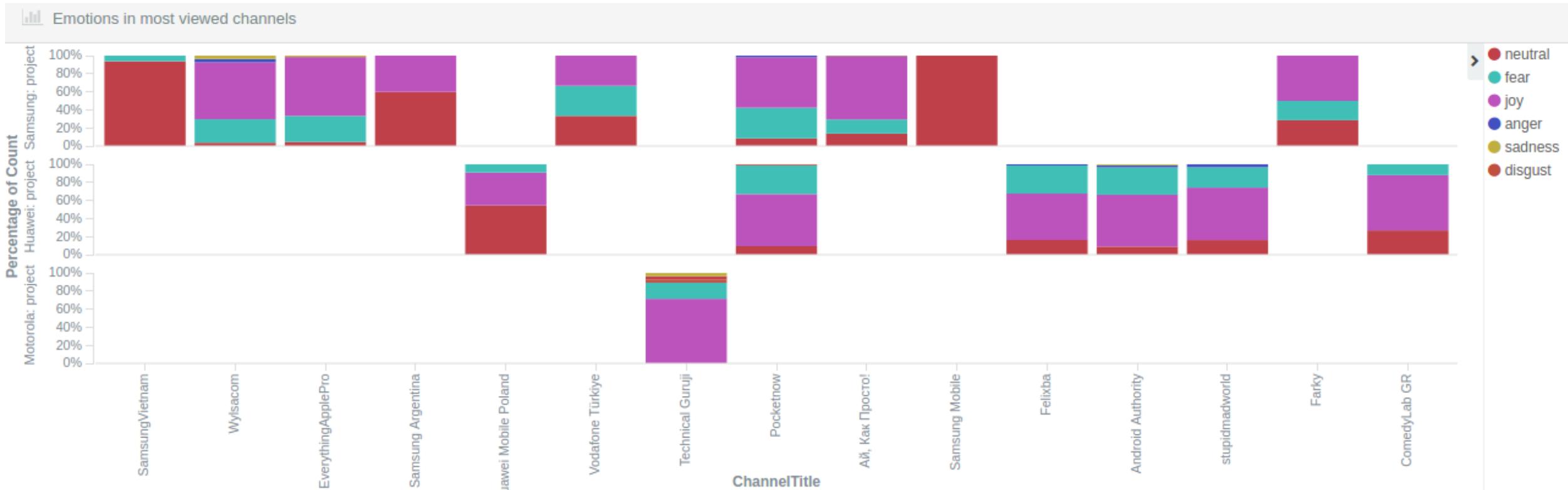
Welcome to **MixedEmotions** Platform

NAVIGATE THE WORLD OF EMOTIONS

The **MixedEmotions** platform is a Big Data Toolbox for multilingual and multimodal emotion extraction and analysis. It can extract emotions from text, audio and video. However, it also has many other capabilities, such as sentiment analysis, social network analysis and knowledge graphs visualization among others.

MixedEmotions develops innovative multilingual multi-modal Big Data analytics applications. Our tools analyse a more complete emotional profile of user behavior using data from mixed input channels:

- ▶ multilingual text data sources
- ▶ A/V signal input (multilingual speech, audio, video)
- ▶ social media (social network, comments)
- ▶ and structured data.

NUI Galway
OÉ Gaillimh

http://mixedemotions.insight-centre.org/

Use Case: Brand Reputation Management

Emotions in most viewed channels

Text Emotion Analysis (Distant Learning)	EN				Run a Task

Distant learning models for emotion detection based on Twitter hash tags for Ekmans six emotions.

Text Emotion Analysis (Lexical)	Multi-lingual				Run a Task

Uses English text emotion lexical resources to produce values for basic emotion categories.

Text Sentiment Analysis (Deep Learning)	EN				Run a Task
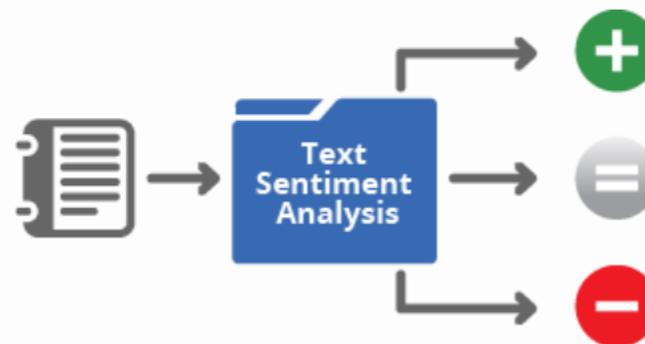
Sentiment analysis models using LSTM neural networks trained on Stanford treebank movie review data.

Paul Buitelaar, Ian D. Wood, et al. (2018). **MixedEmotions: An Open-Source Toolbox for Multi-Modal Emotion Analysis**. IEEE Transactions on Multimedia, pages 1–12.

NUI Galway
OÉ Gaillimh

Suggestion Mining

Sapna Negi (now at Genesys)

Data set development

	Source	Class ratio	Phase 2 agreement (kappa)
Re-tagged Datasets			
Microsoft tweets (original)	Twitter	238 / 2762	1.0
Travel forum (original)	Fodors, Insight Vacations	2192 / 3007	0.76
Travel forum (re-tagged)	Fodors, Insight Vacations	1314 / 3869	0.72
New Datasets			
Hotel train	Tripadvisor	448 / 7086	0.86
Hotel test	Tripadvisor	404 / 3000	0.86
Electronics train	Amazon	324 / 3458	0.83
Electronics test	Amazon	101 / 1070	0.83
Travel test	Fodors	229 / 871	0.72
Software Train	Uservoice	1428 / 4296	0.81
Software Test	Uservoice	296 / 742	0.81

NUI Galway
OÉ Gaillimh

Features

P1: Lexical

Keywords (expanded seed words using WordNet)

P2 – P5: Syntactic

PoS n-grams

PoS patterns in imperative and subjunctive sentences

Subject

P6 – P7: Sentiment

Sentiment score (SentiWordnet)

Normalised score

P8: Semantic

Presence and type of named entity

NUI Galway
OÉ Gaillimh

Negi, S. and Buitelaar, P., 2015. **Towards the extraction of customer-to-customer suggestions from reviews**. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (EMNLP), pp. 2159-2167.

Experiment setup

Domain-specific

Train	Test
Hotel	Hotel
Electronics	Electronics
Travel	Travel
Software	Software

Open domain

Train	Test
Hotel Electronics Travel Software	Hotel
	Electronics
	Travel
	Software

Cross-domain

Train	Test
Hotel	Hotel
Electronics Travel Software	Electronics
	Travel
	Software

Results

Domain Specific

Open Domain

Cross Domain
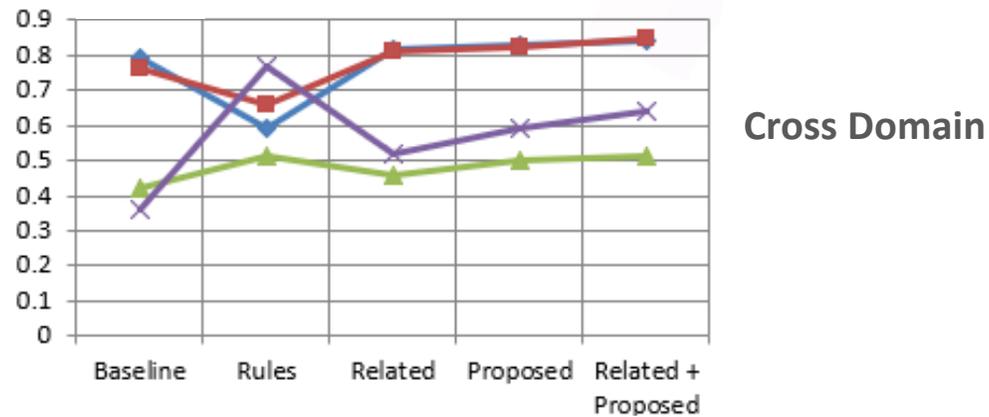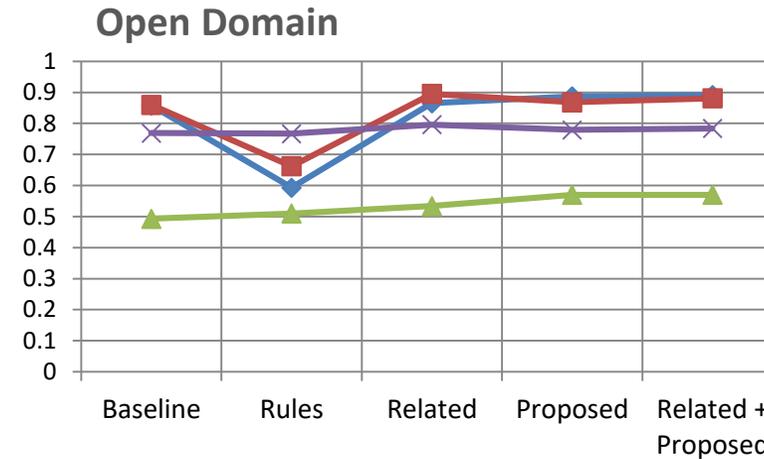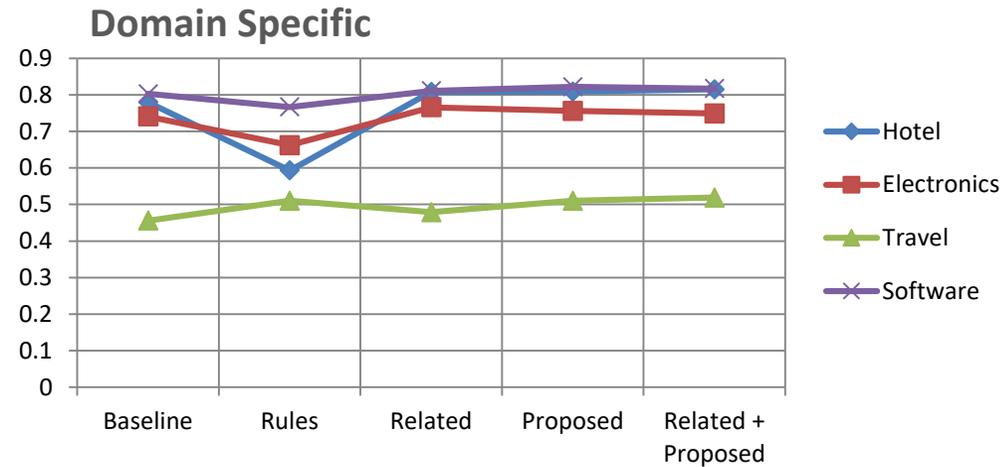
Metaphor Identification in Twitter

Omnia Zayed

Metaphors in Tweets

Peace&Respect @sathishk_raj · 3h

Replying to @bbctamil

@bbctamil , finally u found the trick to **grab attention** for ur posts 😂👍😉

Kindle ✔ @AmazonKindle · Apr 2

Well, that's one way of looking at it! Do you give up on books that don't **capture your interest** or do you power through? amzn.to/2GMKNQu via @ElectricLit

Paul @Kolology · 1 Feb 2015

I'm starting to **catch the idea** on how this game goes. Fair enough.

1   2   2

NUI Galway
OÉ Gaillimh

Metaphor Annotation approach

tweet dataset → **Metaphor Identification Baseline** → weakly annotated dataset → **Crowdsourcing** → human annotated dataset
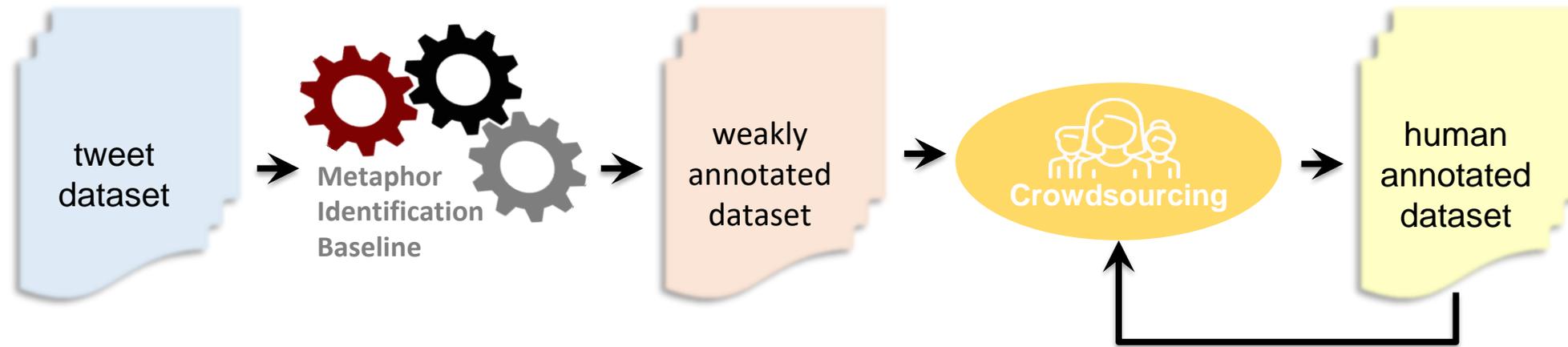
Distributional Metaphor Identification

Candidate	Metaphoric Seed	Cosine Similarity	Candidate	Metaphoric Seed	Cosine Similarity
break promise	break agreement	0.6376	break glass	break agreement	0.5304
	hold back truth	0.4560		hold back truth	0.3435
	fix term	0.3653		frame question	0.3109
	spell out reason	0.3385		face hour	0.2949
	seize moment	0.3384		block out thought	0.2701
	glimpse duty	0.3224		seize moment	0.2677
	grasp term	0.3019		throw remark	0.2583
	frame question	0.2959		skim over question	0.2509
	accelerate change	0.2927		mend marriage	0.2375
	throw remark	0.2776		spell out reason	0.2354

Table 1: The cosine similarity between the candidates "break promise" and "break glass" and the top 10 metaphoric seeds in the seed set using a pre-trained Word2Vec word embedding model on Google News dataset.

Data sources for tweet data set

Semeval-2018 Task 1: Affect in Tweets.

  tweets collected using emotion-related words: *angry, happy, surprised, …*

  100M tweets, retrieved 10K

Irish Elections Twitter Dataset (collected by UCD)

  tweets collected around the recent Irish elections.

  40K tweets (17K unique)

NUI Galway
OÉ Gaillimh

Metaphor Annotation – Some Results

Top-10 Highest Agreement	Top-10 Lowest Agreement
control terrorism	have democracy
take away sadness	leave call
clean up nation	match outfit
drown faith	invade space
put smile	hold prisoner
educate people	take revenge
wear smile	talk sense
twist facts	get chance
bred dragon	change fact
push poland	give anxiety

NUI Galway
OÉ Gaillimh

Market Sentiment Analysis

Tobias Daudert

More holistic view on Market Sentiment

NUI Galway
OÉ Gaillimh
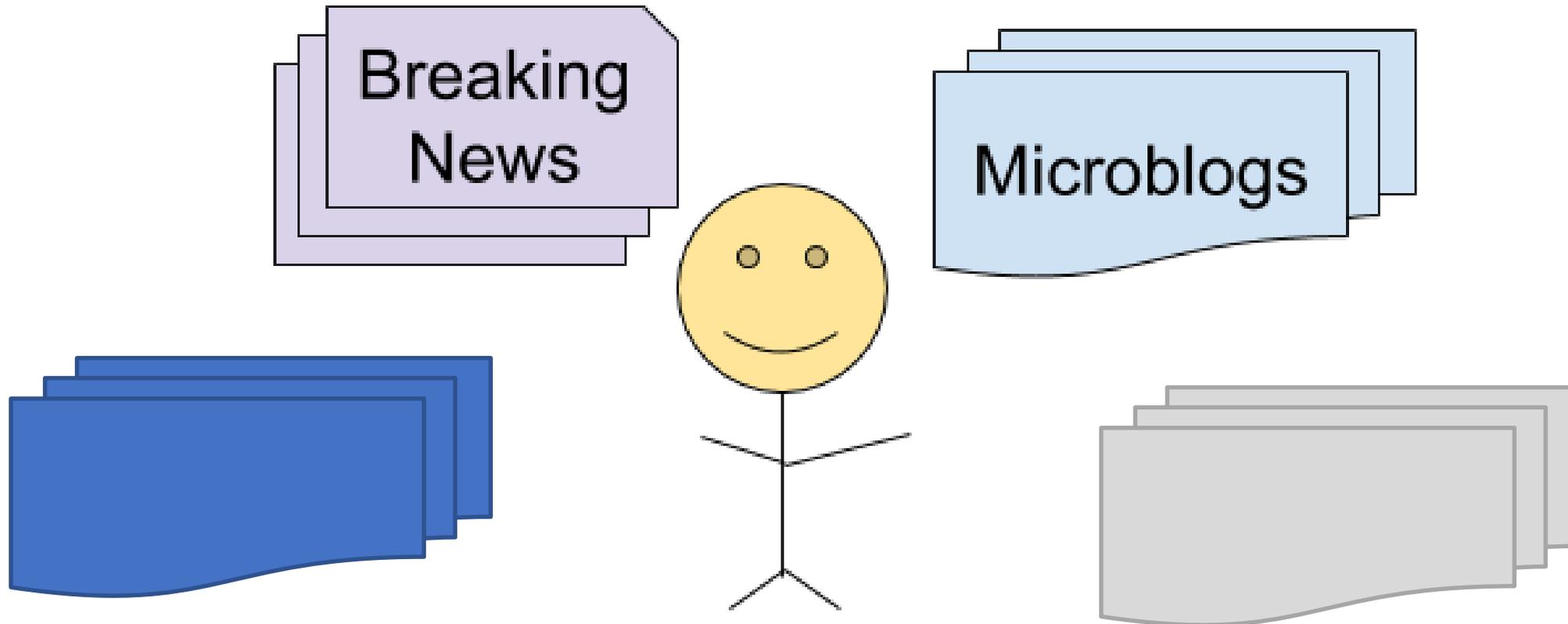
Sentiment sources and influences

News

Microblogs

$AAPL signs up with Google cloud, to spend less with Amazon Web Services. $AMZN drops late in the day from resistance, Slims target 540

$AMZN dropping ~$7B of market cap because they're losing $400M of $AAPL's business may be just a tad bit of an overreaction

Apple Using Google's Cloud, Moving Some Data From Amazon
Major victory for the Alphabet Inc. unit in its battle against Amazon and Microsoft

who would of guessed $AMZN would kill it and $GOOG $GOOGL would suck...not me lol

Google Software Notches Third Straight Win Over Human Go Champion
Artificial intelligence project AlphaGo clinched the best-of-five series against South Korea's Lee Se-dol

Report: Apple signs up for Google's cloud, uses much less of Amazon's $AAPL $GOOG $GOOGL $AMZN $DROPB https://t.co/zN3KDGYvGT

Go Champion Beats AlphaGo Software on Fourth Try
South Korean Go champion Lee Se-dol beat the software after three straight losses last week

After Go victories, Alphabet looks to give DeepMind more complex challenges $GOOG $GOOGL https://t.co/p1xPgypRzT

Sentiment 'conveyance'

News

Microblogs

Apple Using Google's Cloud, Moving Some Data From Amazon
Major victory for the Alphabet Inc. unit in its battle against Amazon and Microsoft

$AAPL signs up with Google cloud, to spend less with Amazon Web Services. $AMZN drops late in the day from resistance, Slims target 540

$AMZN dropping ~$7B of market cap because they're losing $400M of $AAPL's business may be just a tad bit of an overreaction

who would of guessed $AMZN would kill it and $GOOG $GOOGL would suck...not me lol

Report: Apple signs up for Google's cloud, uses much less of Amazon's $AAPL $GOOG $GOOGL $AMZN $DROPB https://t.co/zN3KDGYvGT

NUI Galway
OÉ Gaillimh

Linking

News

Microblogs

Apple Using Google's Cloud, Moving Some Data From Amazon
Major victory for the Alphabet Inc. unit in its battle against Amazon and Microsoft

$AAPL signs up with Google cloud, to spend less with Amazon Web Services. $AMZN drops late in the day from resistance, Slims target 540

$AMZN dropping ~$7B of market cap because they're losing $400M of $AAPL's business may be just a tad bit of an overreaction

who would of guessed $AMZN would kill it and $GOOG $GOOGL would suck...not me lol

Report: Apple signs up for Google's cloud, uses much less of Amazon's $AAPL $GOOG $GOOGL $AMZN $DROPB https://t.co/zN3KDGYvGT

NUI Galway
OÉ Gaillimh

Assigning

QA