

A decorative graphic on the left side of the slide consisting of a network of black and white lines and small circles, resembling a circuit board or a neural network diagram.

DOES AI PERFORM BETTER?

METADATA ENHANCEMENT THROUGH DEEP LEARNING

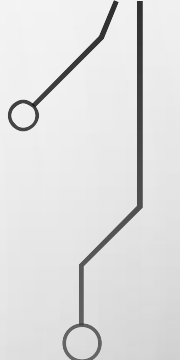
BY CHRISTIAN STUBER AND MANUEL KOCHER, STUDENTS AT BERN UNIVERSITY OF APPLIED SCIENCES



GOAL

- Which techniques and methods of deep learning can be practically applied to support heritage institutions in recognizing objects in digital images and publishing them on the open semantic web?

BY CHRISTIAN STUBER AND MANUEL KOCHER, STUDENTS AT BERN UNIVERSITY OF APPLIED SCIENCES



PROCESS



BY CHRISTIAN STUBER AND MANUEL KOCHER, STUDENTS AT BERN UNIVERSITY OF APPLIED SCIENCES

Acquire
images



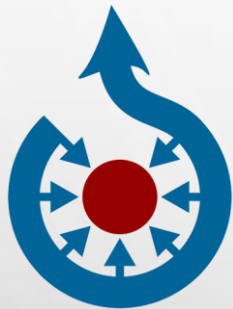
Object
detection

Filter tags

Matching
Wikidata
items

Select
unclear
matches

Upload
depicts

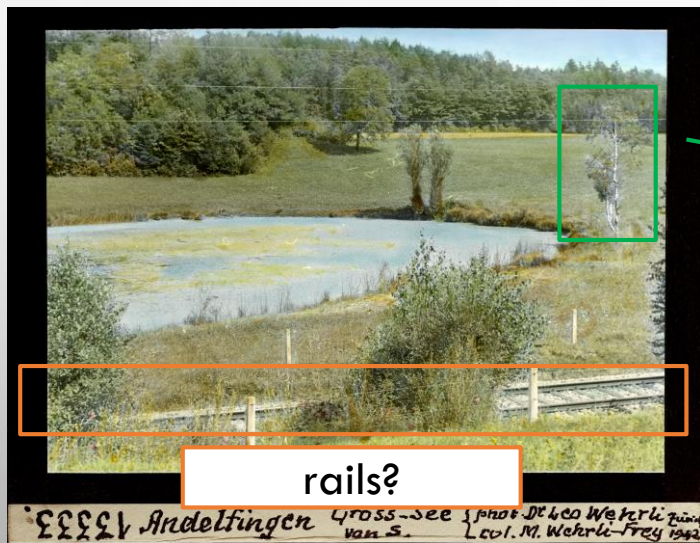


Category:
ETH-BIB Leo
Wehrli

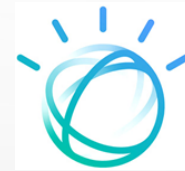


Service conditions:

- Format (.jpg, .jpeg)
- Size < 4 MB
- no west-european accents
(ä, ö, ü, é, è, ê ...)



rails?



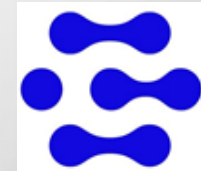
IBM Watson

Tag	Confidence
riverbed	0.54
nature	0.71
oxbow (river)	0.54
detention basin	0.52
passe-partout (frame)	0.51
framework	0.51
supporting structure	0.51
riverbank	0.51
slope	0.51
meander	0.50
pale yellow color	0.92



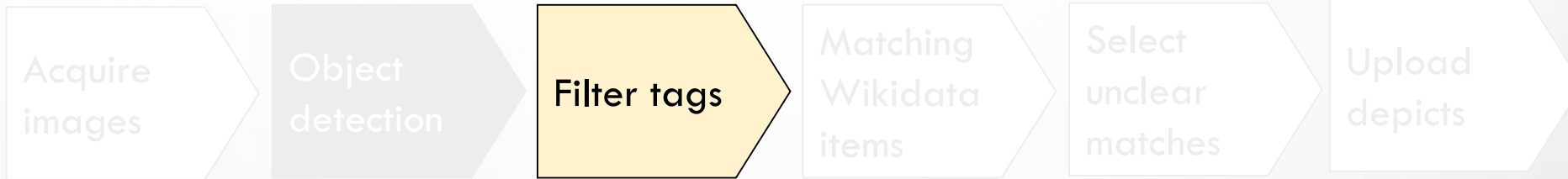
MS Azure

Tag	Confidence
text	1.00
grass	0.98
tree	0.91
plant	0.91
water	0.71
lake	0.59



Clarifai

Tag	Confidence
nature	0.99
landscape	0.99
water	0.99
no person	0.98
grass	0.98
river	0.95
outdoors	0.95
lake	0.95
tree	0.95
pool	0.94
travel	0.91
sky	0.90
...	



Tag	Confidence
grass	1.96
tree	1.86
nature	1.70
water	1.70
lake	1.54
text	1.00
landscape	0.99
no person	0.98
river	0.95
outdoors	0.95
pool	0.94
pale yellow color	0.92
travel	0.91
...	

≥ 1.66



- Query is performed using a python script on the SPARQL endpoint of Wikidata
 1. Returns item with wikipedia article and the most sitelinks
 2. If the input tag matches the item label or its alias, the tag will be used automatically
 3. Otherwise a new query is performed: Returns all matched results (For manual selection)

	clear match	no clear match
Item semantically correct	Can be used automatically	Manual selection needed: Data can be used
Item semantically wrong	Wrong data will be used	Manual selection needed: Data is lost

Acquire
images

Object
detection

Filter tags

Matching
Wikidata
items

Select
unclear
matches

Upload
depicts



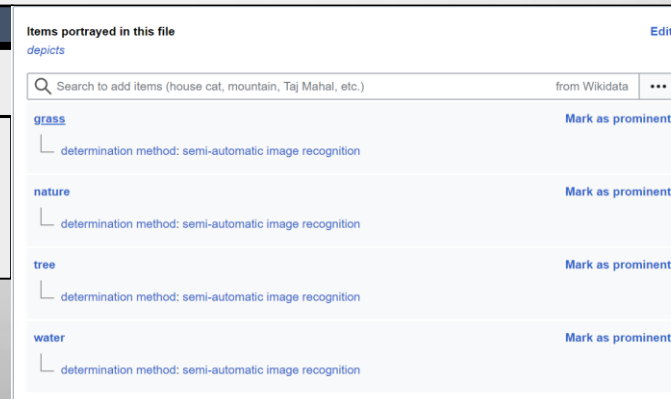
counter	tag input	korrekt	manualwork	item	item label	item description	item known as	item sitelinks
374	grass			1 http://www.wikidata.org/entity/Q43238	Poaceae	family of plants	Gramineae	99
374	grass			1 http://www.wikidata.org/entity/Q43238	Poaceae	family of plants	the grasses	99
374	grass			1 http://www.wikidata.org/entity/Q43238	Poaceae	family of plants	true grass	99
374	grass			1 http://www.wikidata.org/entity/Q2845	cannabis	a psychoactive drug from the Cannabis plant used for medical or recreational purposes	flower	99
374	grass			1 http://www.wikidata.org/entity/Q2845	cannabis	a psychoactive drug from the Cannabis plant used for medical or recreational purposes	Maria	99
374	grass			1 http://www.wikidata.org/entity/Q2845	cannabis	a psychoactive drug from the Cannabis plant used for medical or recreational purposes	Cannabis	99
374	grass			1 http://www.wikidata.org/entity/Q2845	cannabis	a psychoactive drug from the Cannabis plant used for medical or recreational purposes	pot	99
374	grass			1 http://www.wikidata.org/entity/Q2845	cannabis	a psychoactive drug from the Cannabis plant used for medical or recreational purposes	Mota	99
374	grass			1 http://www.wikidata.org/entity/Q2845	cannabis	a psychoactive drug from the Cannabis plant used for medical or recreational purposes	grass	99
374	grass			1 http://www.wikidata.org/entity/Q2845	cannabis	a psychoactive drug from the Cannabis plant used for medical or recreational purposes	herb	99
374	grass			1 http://www.wikidata.org/entity/Q2845	cannabis	a psychoactive drug from the Cannabis plant used for medical or recreational purposes	marihuana	99
374	grass			1 http://www.wikidata.org/entity/Q2845	cannabis	a psychoactive drug from the Cannabis plant used for medical or recreational purposes	marijuana	99
374	grass			1 http://www.wikidata.org/entity/Q2845	cannabis	a psychoactive drug from the Cannabis plant used for medical or recreational purposes	weed	99
374	grass			1 http://www.wikidata.org/entity/Q2845	cannabis	a psychoactive drug from the Cannabis plant used for medical or recreational purposes	Mostaza	99
374	grass			1 http://www.wikidata.org/entity/Q2845	cannabis	a psychoactive drug from the Cannabis plant used for medical or recreational purposes	ganja	99
374	grass			1 http://www.wikidata.org/entity/Q2845	cannabis	a psychoactive drug from the Cannabis plant used for medical or recreational purposes	buds	99
374	grass			1 http://www.wikidata.org/entity/Q2845	cannabis	a psychoactive drug from the Cannabis plant used for medical or recreational purposes	Hierba	99
374	grass			1 http://www.wikidata.org/entity/Q2845	cannabis	a psychoactive drug from the Cannabis plant used for medical or recreational purposes	La Kimberly	99
374	grass			1 http://www.wikidata.org/entity/Q2845	cannabis	a psychoactive drug from the Cannabis plant used for medical or recreational purposes	mostaza	99
374	grass			1 http://www.wikidata.org/entity/Q2845	cannabis	a psychoactive drug from the Cannabis plant used for medical or recreational purposes	reefer	99
374	grass			1 http://www.wikidata.org/entity/Q2845	cannabis	a psychoactive drug from the Cannabis plant used for medical or recreational purposes	tokes	99
374	grass			1 http://www.wikidata.org/entity/Q43238	Poaceae	family of plants	grass	99
374	grass		1	1 http://www.wikidata.org/entity/Q643352	grass	type of herbaceous plant		45
374	grass			1 http://www.wikidata.org/entity/Q2289565	marijuana	name for the cannabis plant	pot	5
374	grass			1 http://www.wikidata.org/entity/Q2289565	marijuana	name for the cannabis plant	grass	5
374	grass			1 http://www.wikidata.org/entity/Q2289565	marijuana	name for the cannabis plant	BOOF	5
374	grass			1 http://www.wikidata.org/entity/Q2289565	marijuana	name for the cannabis plant	cannabis	5
374	grass			1 http://www.wikidata.org/entity/Q2289565	marijuana	name for the cannabis plant	ganga	5
374	grass			1 http://www.wikidata.org/entity/Q2289565	marijuana	name for the cannabis plant	marihuana	5
374	grass			1 http://www.wikidata.org/entity/Q2289565	marijuana	name for the cannabis plant	mary jane	5
374	grass			1 http://www.wikidata.org/entity/Q2289565	marijuana	name for the cannabis plant	weed	5

BY CHRISTIAN STUBER AND MANUEL KOCHER, STUDENTS AT BERN UNIVERSITY OF APPLIED SCIENCES



- Data enrichment on wikimedia commons
- The data enrichment was made using «Quickstatements V2»
 - # Tags: 835 / # Images: 295

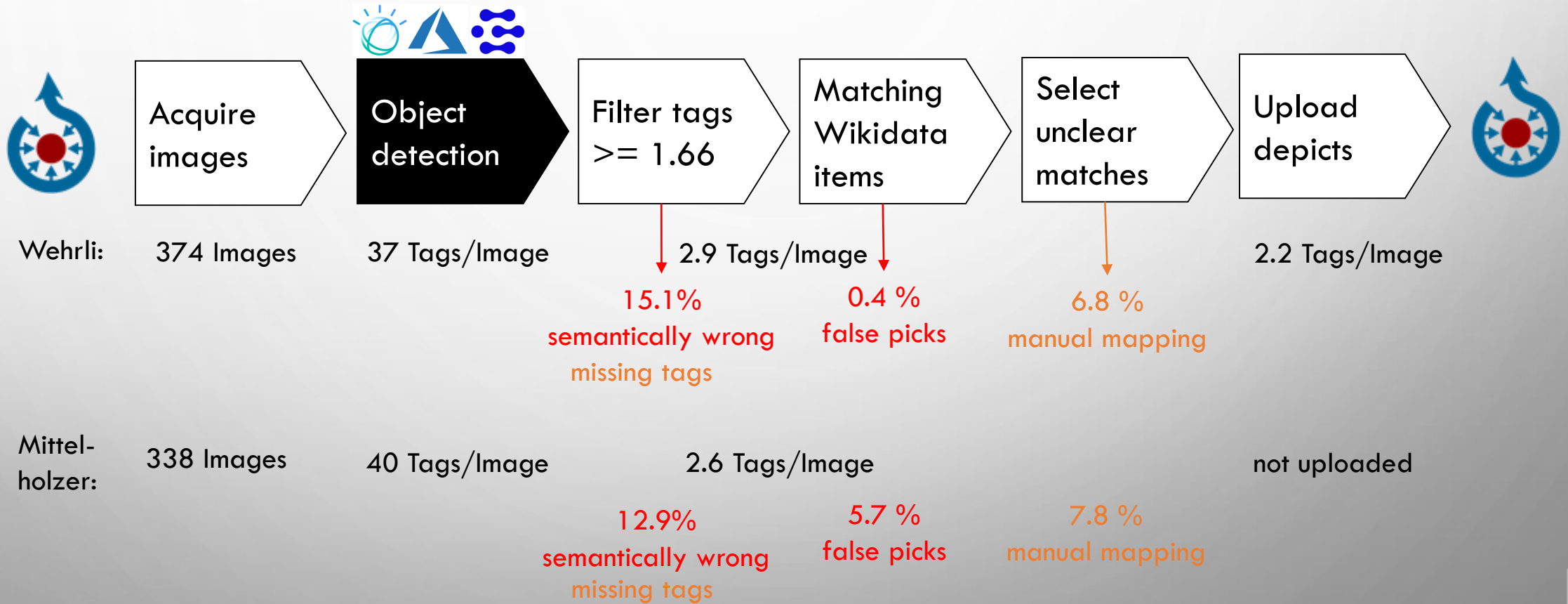
Media Item	Statement	Item	Qualifier	Item
M68548819	P180	Q1229765	p459	Q95076531
Picture	«depicts»	Tag	«determ. method»	«semi-automatic image recognition»



Source: https://commons.wikimedia.org/wiki/File:ETH-BIB-Andelfingen,_Gross-See_von_S%C3%BCden-Dia_247-13333.tif

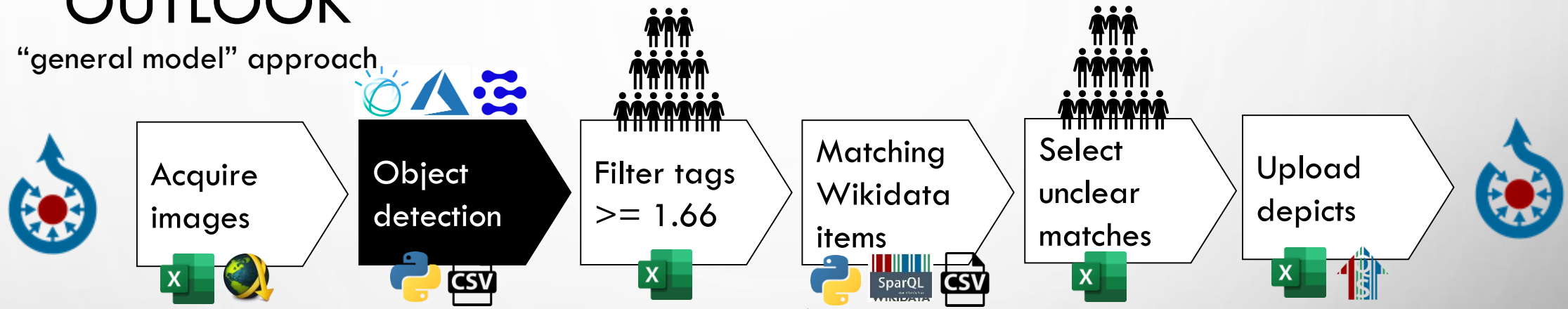
BY CHRISTIAN STUBER AND MANUEL KOCHER, STUDENTS AT BERN UNIVERSITY OF APPLIED SCIENCES

RESULT

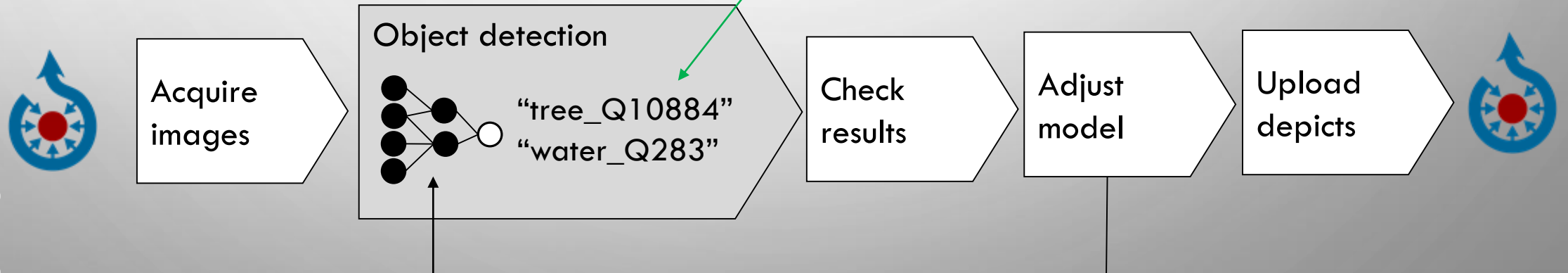


OUTLOOK

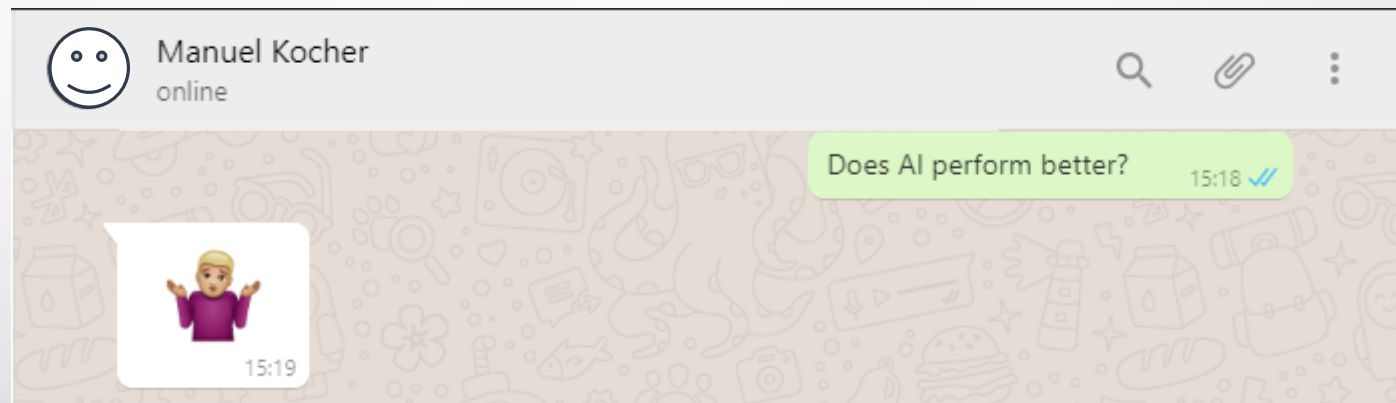
“general model” approach



“custom model” approach



QUESTIONS



BY CHRISTIAN STUBER AND MANUEL KOCHER, STUDENTS AT BERN UNIVERSITY OF APPLIED SCIENCES

IMAGE SOURCES

- <https://upload.wikimedia.org/wikipedia/commons/thumb/4/4a/Commons-logo.svg/2000px-Commons-logo.svg.png> (Logo Wikimedia Commons)
- https://www.underconsideration.com/brandnew/archives/ibm_watson_logo_before_after_a.jpg (Logo IBM Watson)
- <https://windowsarea.de/2017/09/microsoft-azure-bekommt-ein-neues-logo-und-manifest/> (Logo Microsoft Azure)
- <https://gal.imgix.net/logo/o/130043-1580197345-7764952?auto=format&q=50&fit=fill> (Logo Clarifai)
- https://jdownloader.org/_media/knowledge/wiki/jdownloader.png?cache= (Logo JDownloader)
- <https://upload.wikimedia.org/wikipedia/commons/c/c3/Python-logo-notext.svg> (Logo Python)
- https://www.flaticon.com/de/kostenloses-icon/csv-dateierweiterung_28842 (Logo csv)
- https://upload.wikimedia.org/wikipedia/commons/7/7f/Microsoft_Office_Excel_%282018%E2%80%93present%29.svg (Logo Excel)
- https://commons.wikimedia.org/wiki/File:SparQL_-_Wikidata_-_%C3%89cole_de_Biblioth%C3%A9caires_Documentalistes_-_2016-04-04.pdf (Logo SparQL)
- <https://www.wikidata.org/wiki/Help:QuickStatements/de> (Logo Quickstatements)

BY CHRISTIAN STUBER AND MANUEL KOCHER, STUDENTS AT BERN UNIVERSITY OF APPLIED SCIENCES