



NLP@Amazon in Berlin

Making Alexa smarter and more international

*Hagen Fürstenau
Sr. Manager Applied Science*

ML and NLP at Amazon in Berlin

- Development Center established in Berlin in 2013
- Scientists and developers work on a broad range of ML&NLP topics
 - Alexa, Amazon Music, Amazon Robotics, AWS, Community Shopping, Fire TV, Search, ...
- Applied Science
 - Close interaction with software development teams
 - See impact of research on systems & products used around the world

Our Berlin Alexa group

- Enable customers to **access information** through Alexa
- **Question Answering**
 - Answer general or shopping-related questions
 - Leverage internal and external (web) data
 - Build ML models that scale to web-size data and answer in milliseconds
 - Make Alexa more conversational and natural
- **Machine Translation**
 - Allow customers to access any information in their own language
 - State-of-the-art translation quality for questions, answers, and offline data

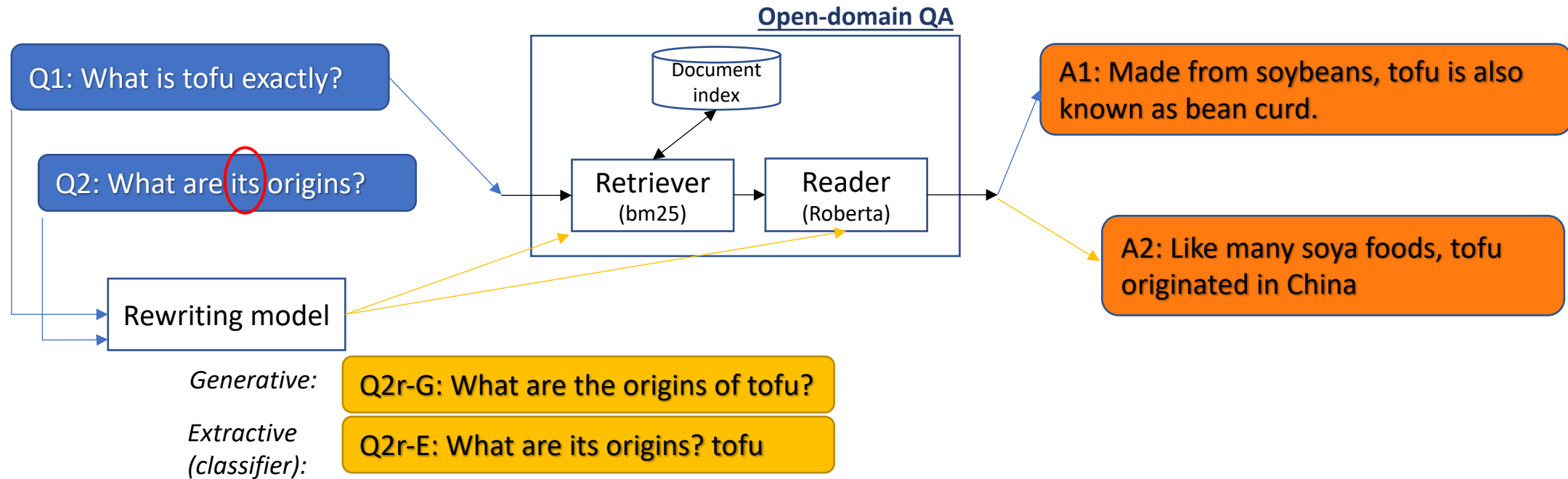
Applied Science at Amazon

- Our projects always start with the **customer**
- Develop a **vision** for a product/feature “working backwards”
- Identify **scientific challenges** that need to be solved
- Keep close contact with the scientific community
 - **Publish** results in conference papers
 - **Internships** for PhD students to work with us
 - Support conferences and events
- Next: Some examples of recently published research...

Open-Domain Conversational QA (1)



- Question Rewriting (QR): Reformulate questions that rely on conversational context
- Reduce risk of providing an answer out of context & reduce required context size as conversation length grows
- **But how does it impact Open-Domain QA system?**



- QR is much more important for retriever than for reader!
- Extractive QR: better for retrievers based on lexical similarity & more efficient, but Generative QR achieves better reader results
- QR solves easy anaphoric cases, but not practical for very long rewrites:

Was anyone opposed to him in this?

Was anyone opposed to Ira Hayes in revealing the truth that Harlon Block was still being misrepresented publicly as Hank Hansen?

Open-Domain Conversational QA (2)

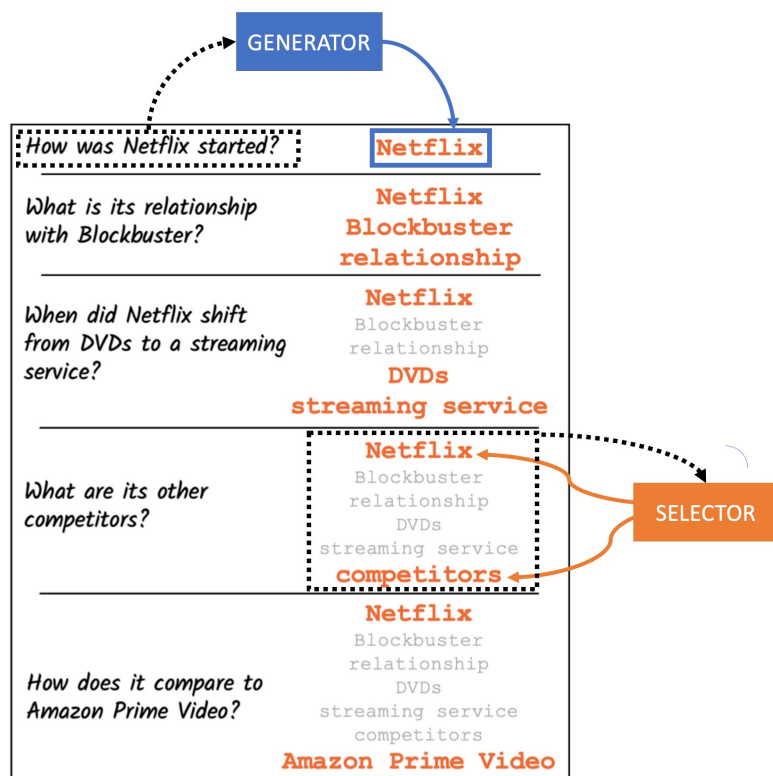


Alternative approach: Embed conversational history along with input question

Q1: What is tofu exactly? A1: Made from soybeans, tofu is also known as bean curd. Q2: What are its origins?

Open-Domain
QA

Keeping whole conversation does not scale. Can we remember only key concepts and attend to them?




- At each turn, **generate** the key phrases to remember in the future, and add them to a set (the Common Ground)
- Given a new question, our model **selects** only elements from the Common Ground
- Human rewrites are used to automatically label the Common Ground for training the generator.
- This history embedding proves superior to Question Rewriting as it can better handle complex long-range dependencies

Product-specific QA – handling diverse data



- When exploring a product website, users have many questions about the given product, or related ones
- A lot of information in multiple formats is available: plain text, key-value pairs, tables, pictures, ...
- **How to build a QA system that gives accurate and natural answers drawing from these diverse data sources?**



Automatic Height Adjustment
Seamlessly move between deep cleaning carpets and hard floors with no loss of suction

Roll over image to zoom in

3 VIDEOS

Eureka FloorRover Bagless Upright Pet Vacuum Cleaner, Suctionseal, Swivel Steering for Carpet and Hard Floor

Visit the EUREKA Store

★★★★☆ 3,093 ratings

List Price: \$164.99 Details
Top Deal: **\$129.99**
Ends in 16h 06m 17s
You Save: **\$35.00 (21%)**

No Import Fees Deposit & \$121.52 Shipping
Coupons: ☐ Save an extra \$10.99 when you buy Eureka products together

Color: Deep Ocean

☐ \$129.99 ☐ \$149.99

Product details

Brand	Eureka
Color	Deep Ocean
Item Dimensions LxWxH	14 x 11.2 x 32.9
Style	Upright
Model Name	FloorRover Upright Pet Vacuum Cleaner
Capacity	1 Liters

- Easy to maneuver: Easy-Glide Wheels and smooth movement of all floor types include
- Automatic height adjustment: Detects floor automatically adjusts for impressive cleaning results
- Multi-cyclone filtration system + HEPA filter: suction and traps 99.97% of dust and allergens

From the manufacturer

EUREKA FLOORROVER DASH
SWIVEL STEERING X MULTI-CYCLONE SUCTION
Move Freely Clean Efficiently

Eureka FloorRover Dash NEU522 Special Features

Easy-Glide Wheels

- Large Easy Glide Wheels utilize innovative Big Wheel technology to ensure smooth movement in every push and pull on all flooring types, including deep carpet.

Multi-Floor Cleaning

- Tiles, Hardwood, Carpet. Cleaning is a breeze with automatic nozzle height adjustment that allows you to move seamlessly between flooring types without constantly having to bend over and manually adjust.
- Additionally, you have the ability to turn the brushroll on or off depending on how deep you want to clean.

Suctionseal

- Suctionseal along fine dust from optimum

Product information

Brand	Eureka
Color	Deep Ocean
Item Dimensions LxWxH	14 x 11.2 x 32.9 inches
Style	Upright
Model Name	FloorRover Upright Pet Vacuum Cleaner
Capacity	1 Liters
Power Source	Electrical cable
Product Dimensions	14 x 11.2 x 32.9 inches
Item Weight	22 pounds
Manufacturer	Eureka
ASIN	B07T5MT5T1
Country of Origin	China
Item model number	NEU522
Customer Reviews	★★★★☆ 3,093 ratings 4.2 out of 5 stars

Customer reviews

★★★★☆ 4.2 out of 5
3,093 global ratings

5 star	63%
4 star	18%
3 star	8%
2 star	4%
1 star	7%

~ How customer reviews and ratings work

By feature

Maneuverability	★★★★☆ 4.2
For cleaning up hair	★★★★☆ 4.2
Suction power	★★★★☆ 4.0

~ See more

Customer questions & answers

Have a question? Search for answers

Question: Do you have to change settings from floor to carpet or is it auto?
Answer: Hello and thank you for your question about the Eureka NEU522 FloorRover Vacuum. This vacuum has Automatic height adjustment. It detects the floor type and automatically adjusts for impressive cleaning results. Have a great day. (BBB A+ Accredited Business)
By Customer Service USA MANUFACTURER on February 9, 2021
~ See more answers (2)

Question: Is this vacuum self-propelled?
Answer: Hello and thank you for your question about the Eureka NEU522 FloorRover Dash. The vacuum is not self-propelled, but it is easy to maneuver and glides easily.
By Customer Service USA MANUFACTURER on July 23, 2020
~ See more answers (2)

Top reviews from the United States

Amazon Customer
★★★★☆ Happy I purchased
Reviewed in the United States on August 17, 2022
Color: Grey and Red | Verified Purchase

I waited an entire month before writing this review, because it's important to me to become familiar with all aspects of a new vacuum cleaner before critiquing. I purchased this primarily because my \$600+ vacuum decided to break a clutch plate and the repair would've cost more than this brand new Eureka. I was also tired of paying \$30 for 4 bags!

I have a shag carpet, and was very nervous because of many reviews stating it didn't work properly on shag carpet....I say (at least for myself) not so at all. Not only did it work, I found that it sucked up dust quite well. I'm not an easy to please type, so I was happily surprised about this. I must say however, that it was most definitely much more strenuous to use than my Riccar vacuum, so you will definitely get a work out using this model. I did appreciate the fairly long cord, didn't really appreciate that it does seem tipsy at times as others have commented.

The suction is very good, the dusting tool definitely would be enhanced by a longer extension. The one thing that did bug me was although I really liked the longer cord, the "peg" area on the vacuum that is meant to hold/wind it doesn't accommodate this length....
~ Read more

2 people found this helpful

Helpful Report abuse

Alida
★★★★☆ Great pet vacuum cleaner
Reviewed in the United States on August 13, 2022
Color: Deep Ocean | Verified Purchase

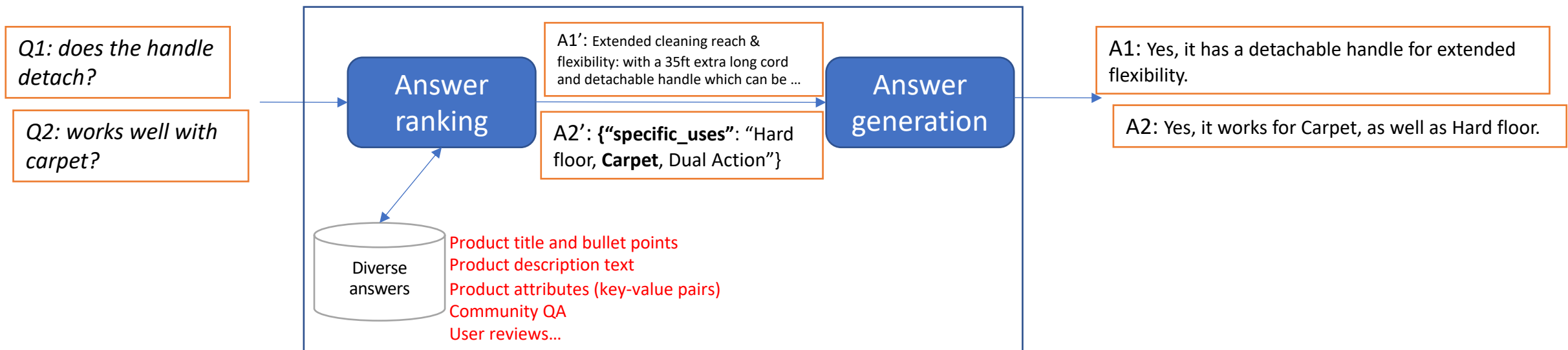
We needed a new vacuum since are last one literally fell apart. The house was in desperate need to be deep cleaning with a vacuum and once we recieved this Eureka it was so easy, simple, and quick to assemble. Once together we went straight to work. We have 4 cats and a dog who shed constantly. This vacuum easily sucked all that up. Plus I was shocked to see how much fine house dust this thing picked up. The suction power is amazing. I used the pet hair brush on the furniture and it picked up the hair without a problem. The vacuum rolls smoothly, easy to control, and does a great job. The only complaint I have for this vacuum is that there is not a longer extension for the hand held suction. Even with that I would still recommend to anyone who is looking for a new vacuum to get this one. It's that good.

Product-specific QA – handling diverse data

Two challenges

1. Select the answer-bearing content from the page:
 - QA model to rank both plain text and semi-structured answer candidates (Transformer LM tuned on small in-domain data set)
 - Apply data augmentation and weak supervision to produce pseudo-labelled QA pairs from product websites
2. Produce a fluent direct answer to the question:
 - Generative model applied to the selected answer-bearing content
 - Challenge: Must be 99% faithful while as natural as possible

Question: works well with carpet?		
Source	Supporting Evidence	Relevance
Attribute	{"specific_uses": "Hard floor, Carpet, Dual Action"}	YES
Bullet Point	On-board accessories include Pet Turbo Brush, multi-angle adapter, crevice tool and dusting brush	NO
Description	Eureka FloorRover Bagless Upright Pet Vacuum Cleaner for Carpet and Hard Floor	YES
Community QA	No there isn't a cover for the handle	NO
Reviews	My dad bought it and was very happy with it.	NO
Annotated answer: Yes, it works well for carpet, as well as hard floor.		



Domain Adaptation for Machine Translation

- **Problem:** Adapt a general MT model to two domains A and B
 - Training first on A and then B will lead to **catastrophic forgetting** (of A)
- **Elastic Weight Consolidation (EWC)** modifies training objective
 - Regularization to keep model parameters close to those for A
 - Makes strong conditional independence assumptions
- **Data Mixing** is a practical way to maintain performance on A
 - Mix in a sample from domain A when training for domain B
 - Size of sample from A determines training set size → hard to balance
- **Combination of EWC and Data Mixing works best**
 - Theoretical analysis shows that Data Mixing is a correction of EWC with weaker independence assumption!

The Devil is in the Details

- **Vocabulary selection** is used to speed up Neural Machine Translation
 - During decoding, only consider union of top-k translations of source tokens
- Minor degradation in BLEU, but noticeable drop in human evaluation!

EN: to swal|low the bitter pill

DE: in den sau|ren ap|fel bei|ßen

- **Solution:** Neural Vocabulary Selection
 - Predict sentence-specific target vocabulary from encoder representations
- **Better quality on idiomatic translations** (at same translation speed)

Jobs at Amazon

- Regularly offer Applied Science internships
 - Aimed at PhD students
 - Duration: 3-6 months, timing flexible
- Openings for permanent positions posted online
- Get in touch!
 - Our recruiter Michal Barabanow: present here at KONVENS!
 - <http://amazon.jobs>