

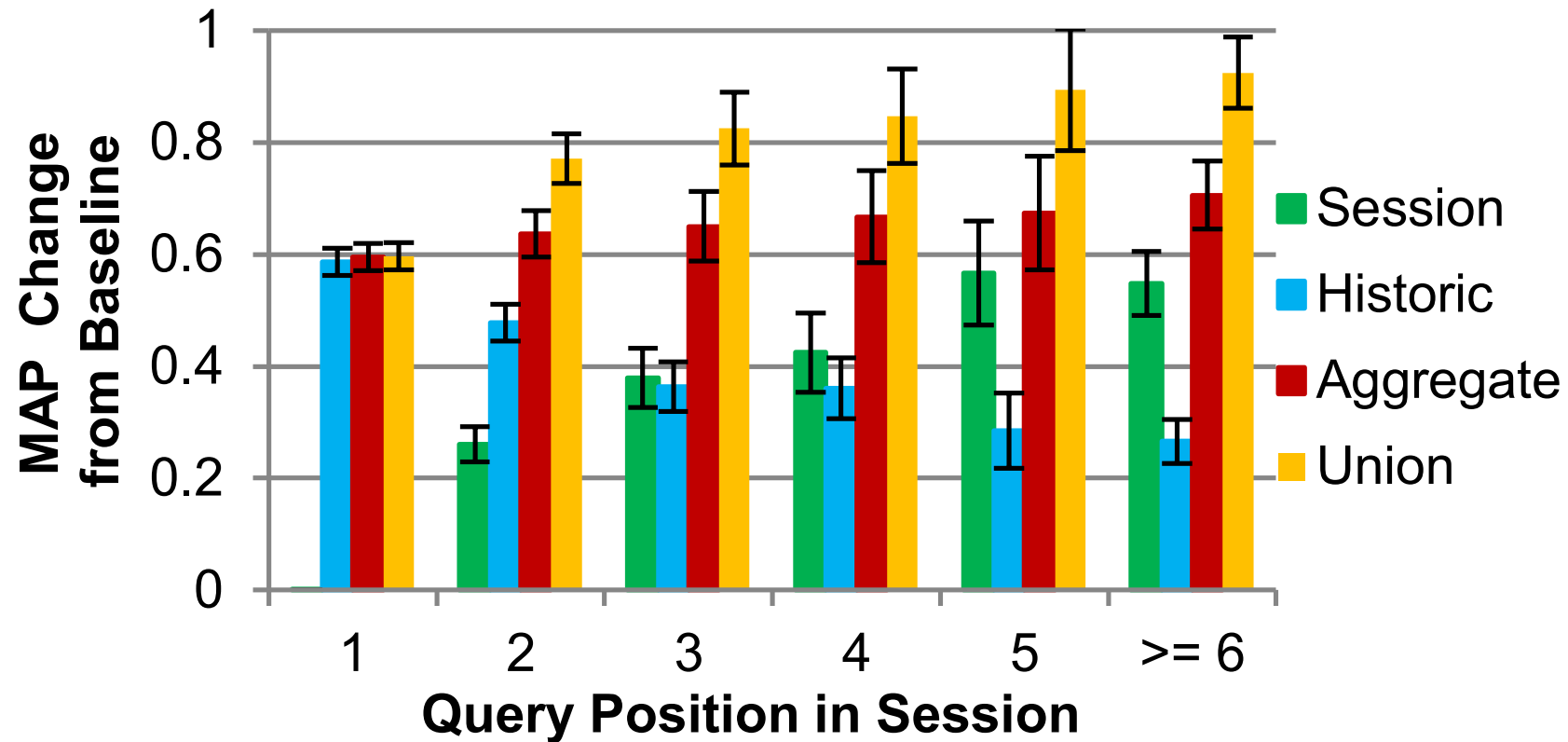
From sessions to conversations and back again

Evangelos Kanoulas

Informatics Institute & Amsterdam Business School
University of Amsterdam

March 19, 2021

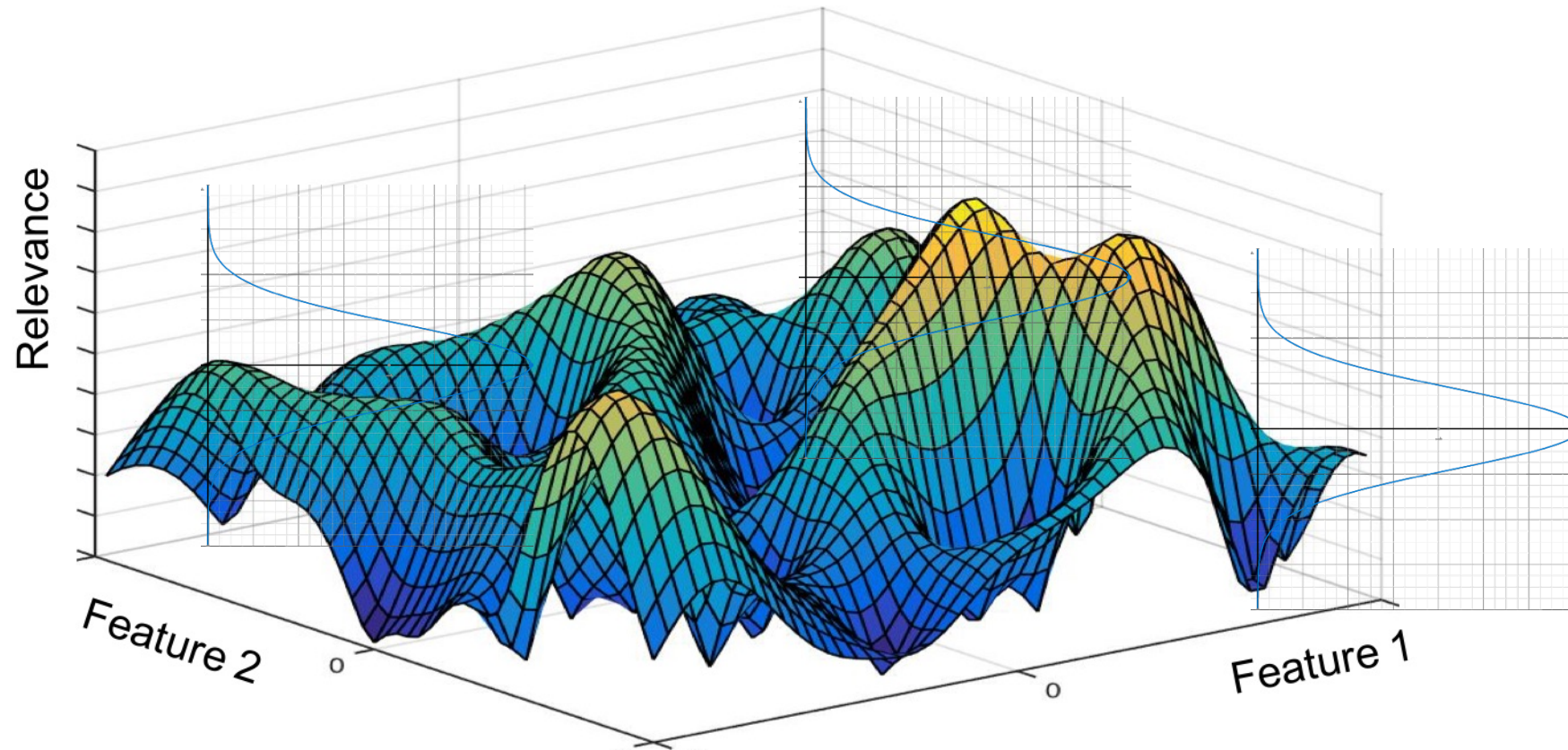
Historic vs Session-based Personalization



- ✓ Bennett, P. N., White, R. W., Chu, W., Dumais, S. T., Bailey, P., Borisyuk, F., & Cui, X. Modeling the impact of short-and long-term behavior on search personalization. SIGIR 2012

What to expect from a search agent?

1. Profile users



2. Account for session history

What is the best available osteoporosis treatment?

Rank List 1

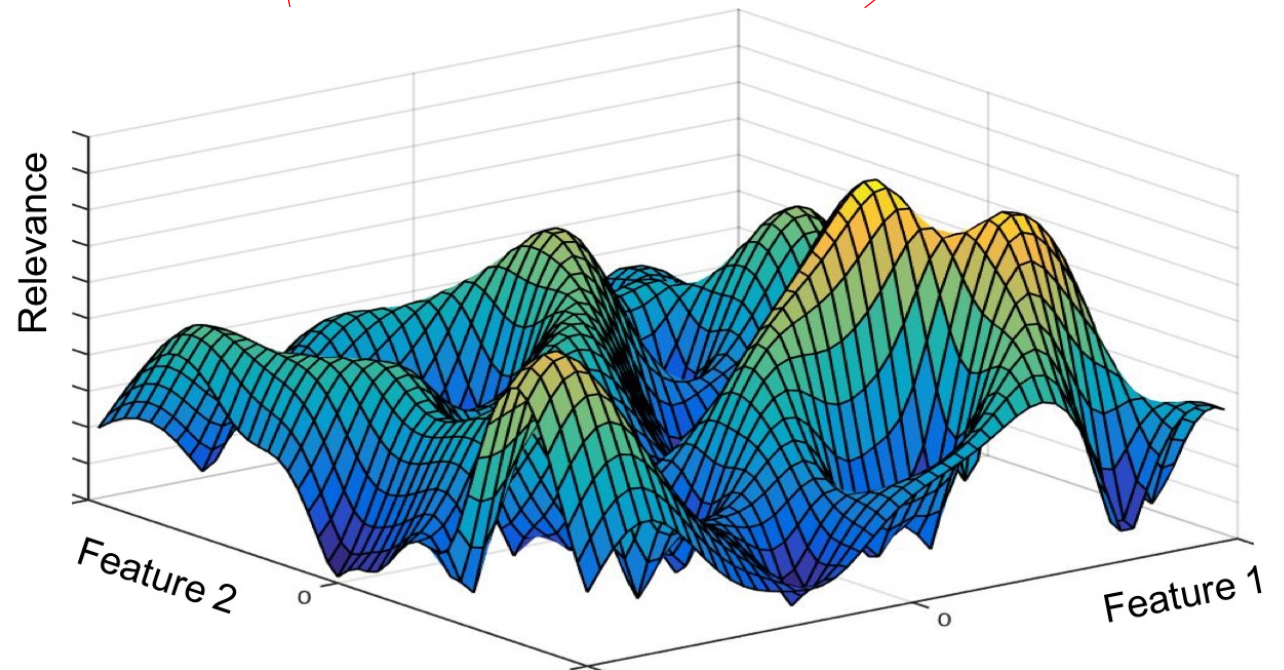
What are the cons of bisphosphonates?

It is an expensive therapy.

Can it rebuild low bone density?

Relate new query to past queries

Update belief



3. Ask clarifying questions

What is the best available osteoporosis treatment?

Are you looking for a pharmaceutical treatment?

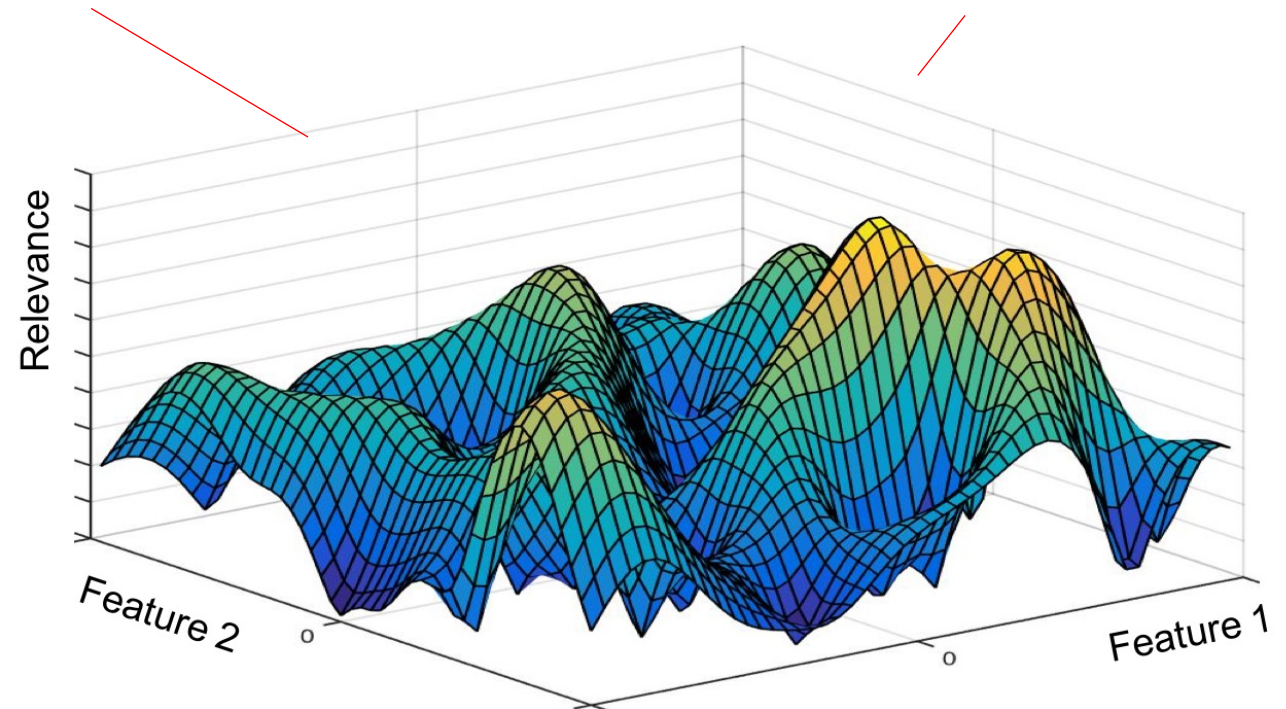
YES

How about combining it with physical exercise?

NO, I am too old for that, except if it is easy to do.

Generate and select the most informative question(s)

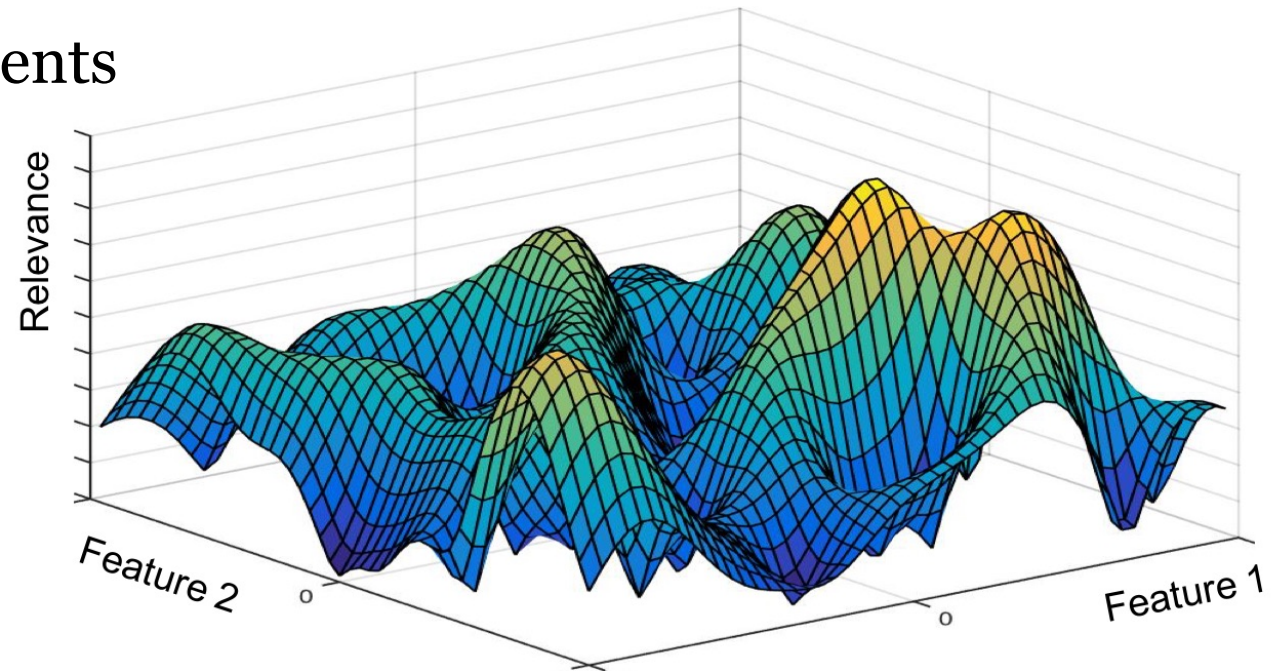
Update belief based on the answer



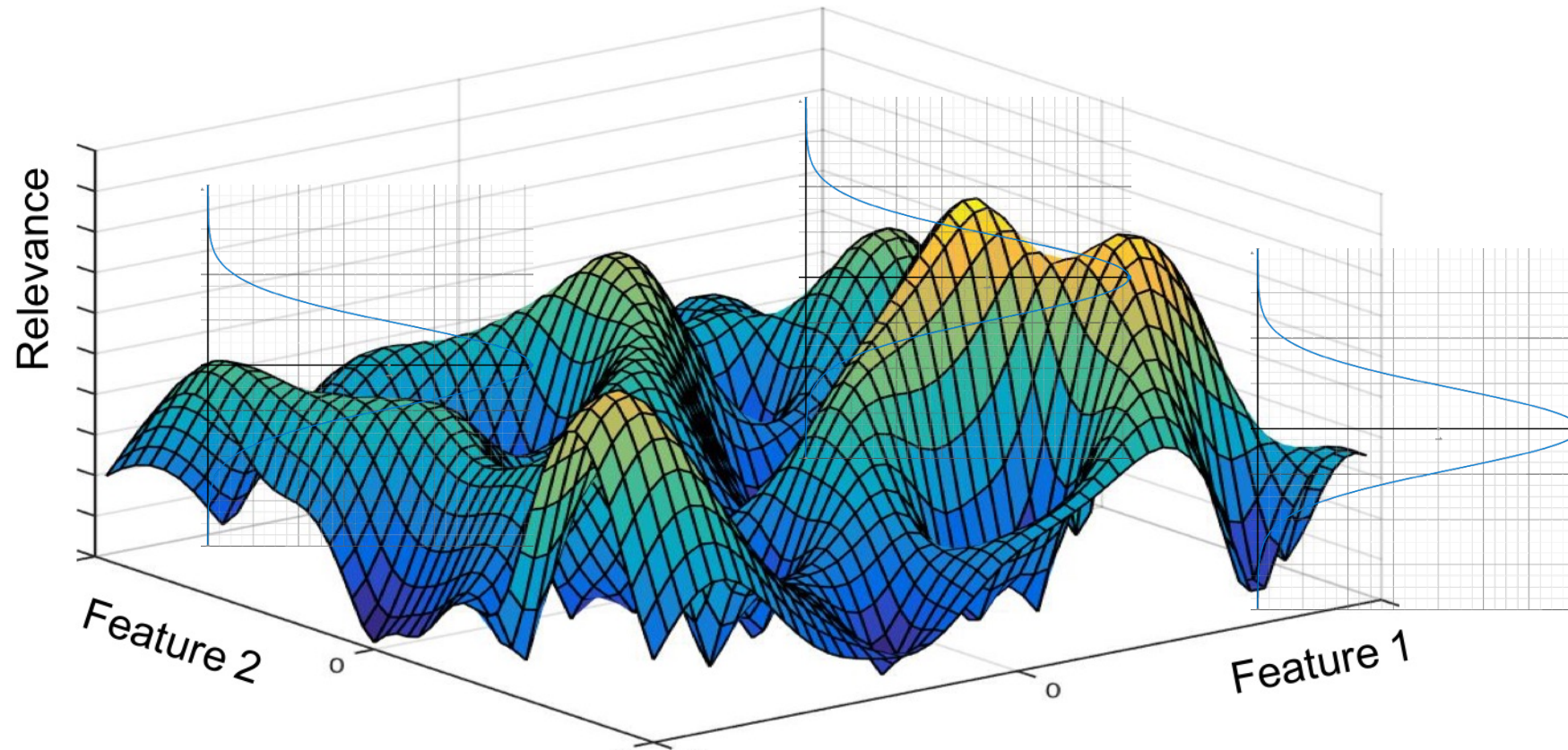
4. Decide Next Action

1. Provide ranking of document/passages
2. Provide direct answer
3. Ask clarifying question
 1. Relevance over items/documents
 2. Natural language questions
 3. Suggested questions

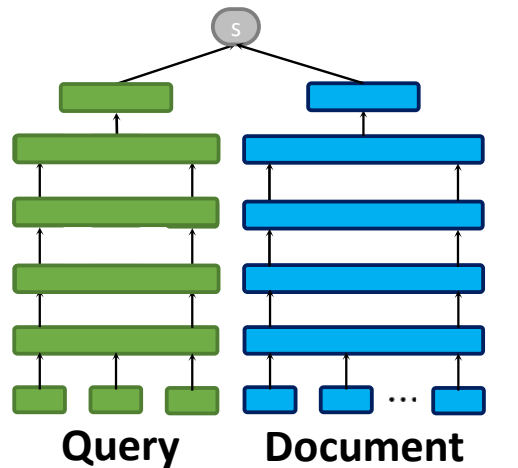
Decide what is the
optimal next action to
take



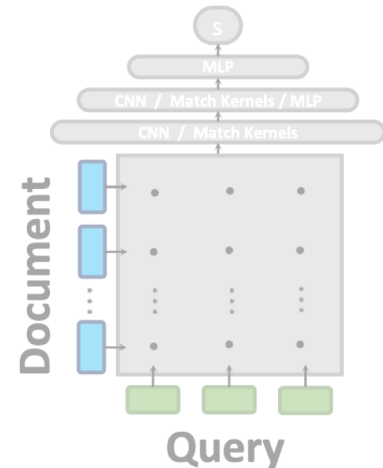
1. Profile users



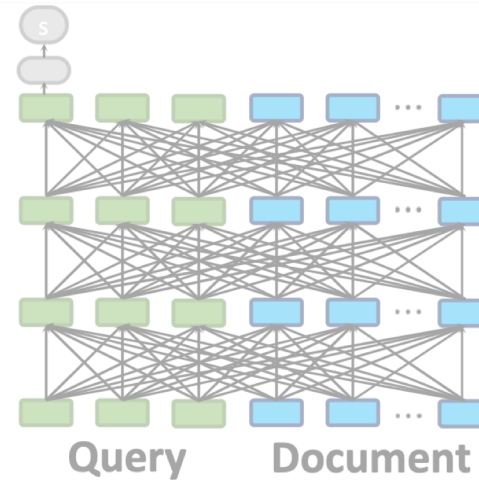
Estimate relevance



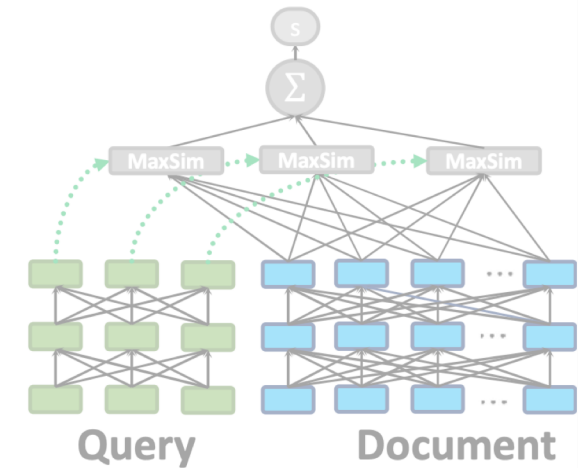
(a) Representation-based Similarity



(b) Query-Document Interaction



(c) All-to-all Interaction



(d) Late Interaction

- ✓ C Van Gysel, M de Rijke, E Kanoulas: Neural Vector Spaces for Unsupervised Information Retrieval. *TOIS* 2018
- ✓ C Van Gysel, M de Rijke, E Kanoulas: Learning Latent Vector Spaces for Product Search. *CIKM* 2016

Challenge: Training complex models

Question:

where is the world's largest ice sheet located today?

Short Answer:

Antarctica

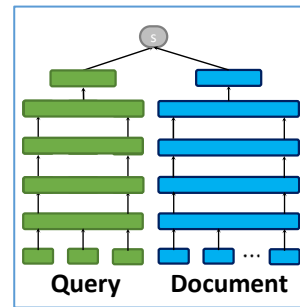
Long Answer:

The Antarctic ice sheet is the largest single mass of ice on Earth . It covers an area of almost 14 million km and contains 30 million km of ice . Around 90 % of the Earth's ice mass is in Antarctica , which , if melted , would cause sea levels to rise by 58 meters . The continent - wide average surface temperature trend of Antarctica is positive and significant at $> 0.05^{\circ}\text{C}$ / decade since 1957 .



question

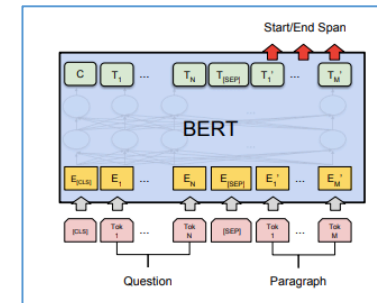
context
Large doc collection



Document list

answer

Extractive phrase
Sentence
Entire Passage

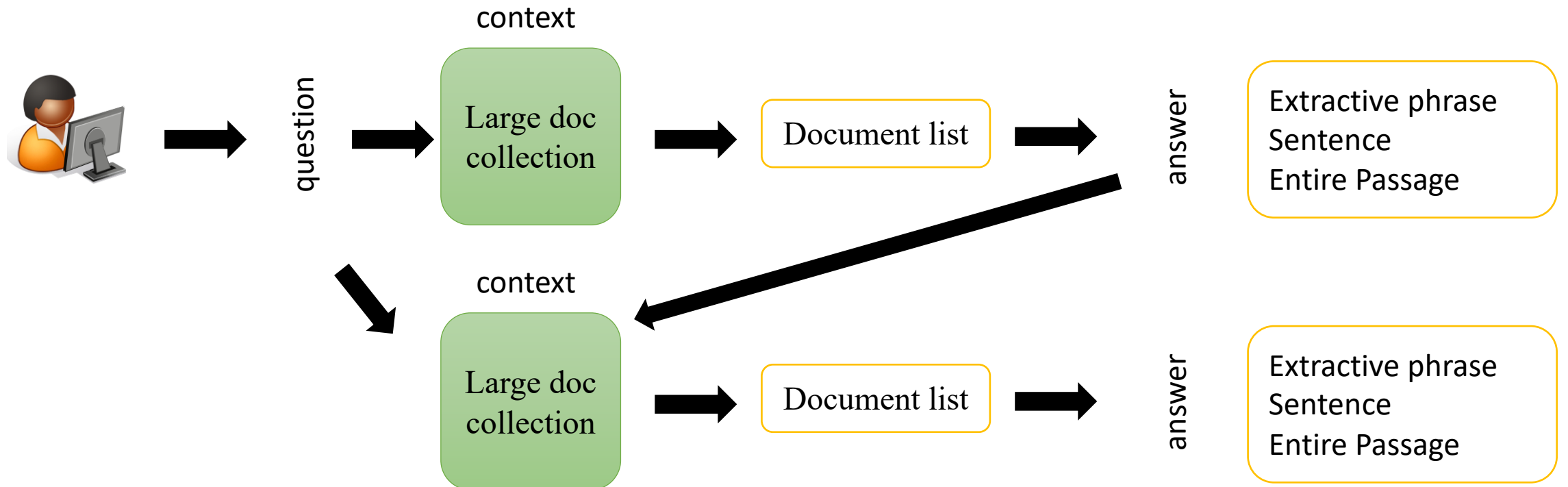


Challenge: Training complex models

Q: What was the nickname of **Judy Lewis**'s father ?

P1: **Judy Lewis** (born Judith Young; November 6, 1935 – November 25, 2011) was an American actress, writer, producer, and therapist. She was the secret biological daughter of actor **Clark Gable** and actress Loretta Young.

P1: **William Clark Gable** (February 1, 1901 – November 16, 1960) was an American film actor, often referred to as "**The King of Hollywood**". He had roles in more than 60 motion pictures in a wide variety of genres during a career that lasted 37 years...



Challenge: Training complex models

Open-domain Complex QA

| Model | Encoder | # GPU | # Epochs | Batch size | # Gradient accumulation steps | TR@2 | TR@10 | TR@20 |
|------------|------------|----------|----------|------------|-------------------------------|-------|-------|-------|
| MDR (full) | RoBERTa | 8 × 32GB | 50 | 150 | 1 | 0.678 | 0.773 | 0.794 |
| MDR | RoBERTa | 8 × 32GB | 50 | 150 | 1 | 0.637 | 0.742 | 0.772 |
| | RoBERTa | 4 × 24GB | 50 | 28 | 1 | 0.607 | 0.711 | 0.736 |
| | RoBERTa | 4 × 24GB | 25 | 28 | 1 | 0.551 | 0.668 | 0.698 |
| | RoBERTa | 4 × 24GB | 20 | 28 | 1 | 0.537 | 0.660 | 0.688 |
| | RoBERTa | 1 × 12GB | 25 | 4 | 32 | 0.441 | 0.581 | 0.620 |
| | BERT | 1 × 12GB | 25 | 4 | 32 | 0.421 | 0.560 | 0.597 |
| | DistilBERT | 1 × 12GB | 25 | 4 | 32 | 0.268 | 0.411 | 0.454 |

2. Account for session history

What is the best available osteoporosis treatment?

Rank List 1

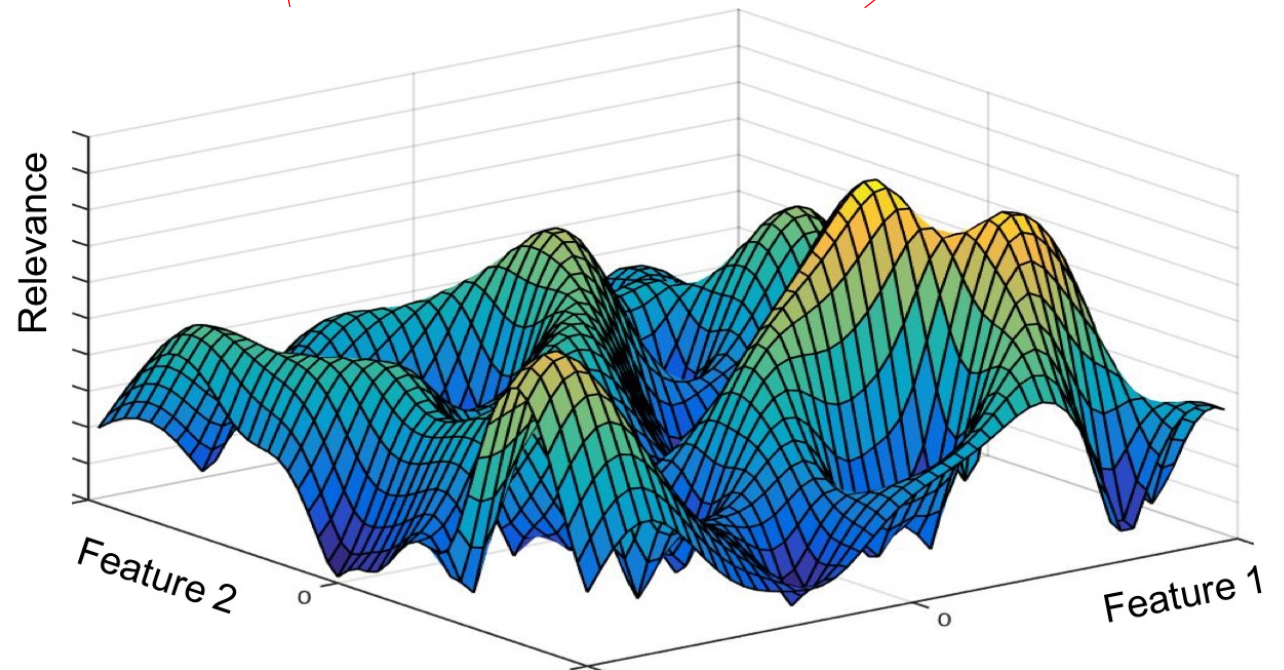
What are the cons of bisphosphonates?

It is an expensive therapy.

Can it rebuild low bone density?

Relate new query to past queries

Update belief



User asking questions

- Session-oriented search
- Task-oriented search
- Conversational search

1. legality alcohol marijuana
2. legality alcohol marijuana history
3. legality alcohol history
4. why is marijuana illegal?
5. is marijuana worse for your health than alcohol?

- ✓ B Carterette, P Clough, M Hall, E Kanoulas, M Sanderson: The TREC Session Track 2011-2014. *SIGIR 2016*
- ✓ C Van Gysel, E Kanoulas, M de Rijke: Lexical Query Modeling in Session Search. *ICTIR 2016*

User asking questions

- Session-oriented search
- Task-oriented search
- Conversational search

1. What are the different types of hydroponic systems to choose from, their pros and cons?
2. What hydroponic system are used in commercial greenhouses?
3. What material is needed to build a hydroponic system myself?
4. Which plants can be grown in a hydroponics system?
5. How long does it take to grow them?
6. What nutrients to test the water for?
7. ...

User asking questions

- Session-oriented search
- Task-oriented search
- **Conversational search**

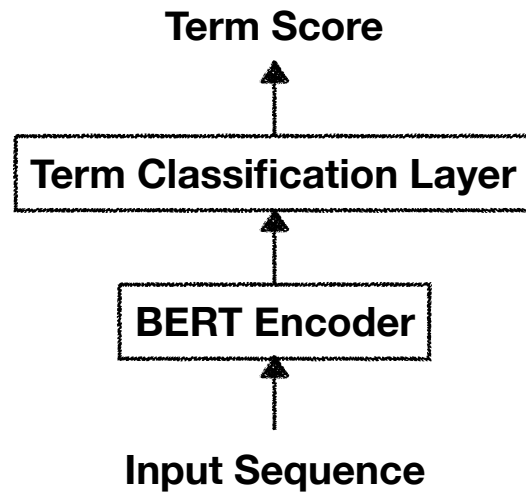
| | |
|-------|------------------------------------|
| 1 | who formed saosin ? |
| 2 | when was the band founded? |
| 3 | what was their first album? |
| <hr/> | |
| 4 | when was the album released? |

Query resolution in conversational search

| Turn | Query |
|------|------------------------------------|
| 1 | who formed saosin ? |
| 2 | when was the band founded? |
| 3 | what was their first album? |
| 4 | when was the album released? |

| Turn | Query |
|------|---|
| 1 | who formed saosin ? |
| 2 | when was the band founded? |
| 3 | what was their first album? |
| 4 | when was the album released? |
| | resolved: when was saosin 's first album released? |

Query Resolution in Conversational Search

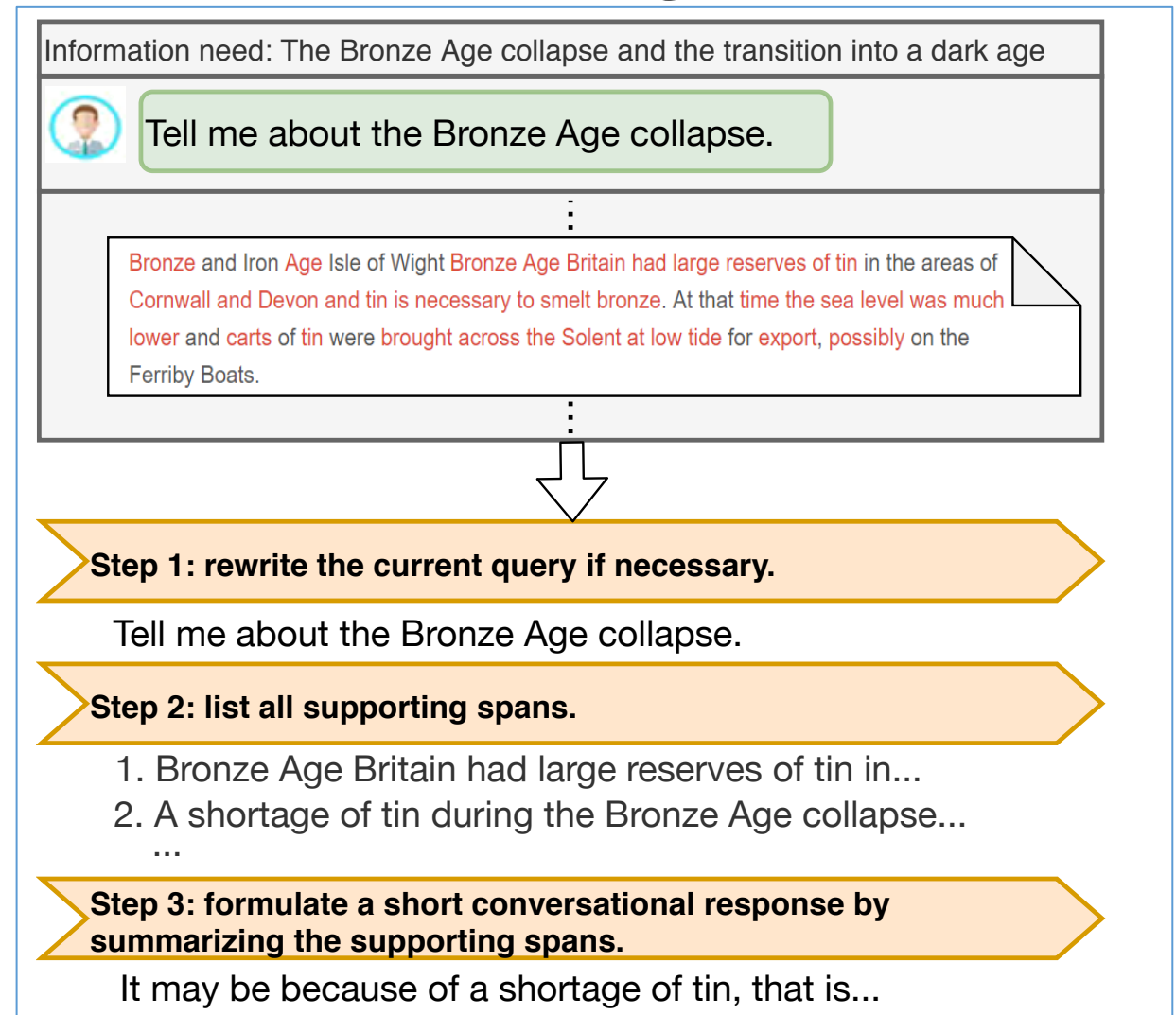


| | | | | | | | | | | | | | | | | | | | | |
|----------------|-------|---------|--------|---------|---------|-----|-----|------|---------|---------|-----|-------|-------|-------------------|-------|------|-----|-----|-------|----------|
| Label | - | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | - | - | - | - | - | - |
| Input Sequence | <CLS> | Who | formed | Saosin? | When | was | the | band | formed? | What | was | their | first | album? | <SEP> | When | was | the | album | released |
| | | Turn #1 | | | Turn #2 | | | | | Turn #3 | | | | Turn #4 (current) | | | | | | |

| Turn | Query |
|------|------------------------------------|
| 1 | who formed saosin ? |
| 2 | when was the band founded? |
| 3 | what was their first album? |
| 4 | when was the album released? |

Relevant passage to turn #4: The original lineup for **Saosin**, consisting of Burchell, Shekoski, Kennedy and Green, was formed in the summer of 2003. On June 17, the **band** released their **first** commercial production, the EP Translating the Name.

Challenge: From document ranking to answer generation



Challenge: From document ranking to answer generation

New SCAI'21 Conversational Question Answering Challenge:
<https://scai.info/>

1. Question rewriting;
2. Document retrieval; and
3. Question answering.

✓ Anantha, R., Vakulenko, S., Tu, Z., Longpre, S., Pulman, S. and Chappidi, S., 2020. Open-Domain Question Answering Goes Conversational via Question Rewriting. arXiv preprint arXiv:2010.04898.

Challenge: Large-scale dynamic collection

- Automate corpus creation (1h / topic)
 - Question generation
 - Answer relevance
- User simulator
 - Question conditioned on previous answer

✓ S Vakulenko et. al, Generating Topic-based Question Sequences for an Automated Evaluation of Conversational Passage Ranking.
(*work in progress*)

Challenge: Large-scale dynamic collection

1. what is the speed of a mako shark
 2. what do mako sharks eat
 3. where does the longfin mako live
 4. where do bull sharks live
 5. what kind of teeth do sharks have
 6. how long do hammerhead sharks live
 7. how long do sandbar sharks live
 8. how long do tiger sharks live
 9. what sharks are endangered
 10. what is the location of the tiger shark
 11. what is the biggest great white shark
1. what are the different types of sharks
 2. are sharks endangered if so which species
 3. tell me more about tiger sharks
 4. what is the largest shark ever to have lived on Earth
 5. what's the biggest shark ever caught
 6. what about for great whites
 7. tell me about mako sharks
 8. what are mako shark adaptations
 9. where do mako sharks live
 10. what do mako sharks eat
 11. how do mako sharks compare with tiger sharks for being dangerous

Challenge: Large-scale dynamic collection

1. what are red blood cells
2. how are red blood cells created
3. how is oxygen transported in blood
4. what is anemia
5. what are the symptoms of anemia
6. can anemia go away
7. what are anemia possible causes
1. does anemia cause fatigue
2. causes of anemia in men
3. can anemia cause death
4. what is normocytic anemia
5. symptoms of anemia in children
6. what is the treatment for iron deficiency anemia
7. red blood cell disorder symptoms
8. what causes low red blood cells
9. causes of low iron levels in blood
10. where are red blood cells found
11. what food is high in iron

3. Asking clarifying questions

What is the best available osteoporosis treatment?

Are you looking for a pharmaceutical treatment?

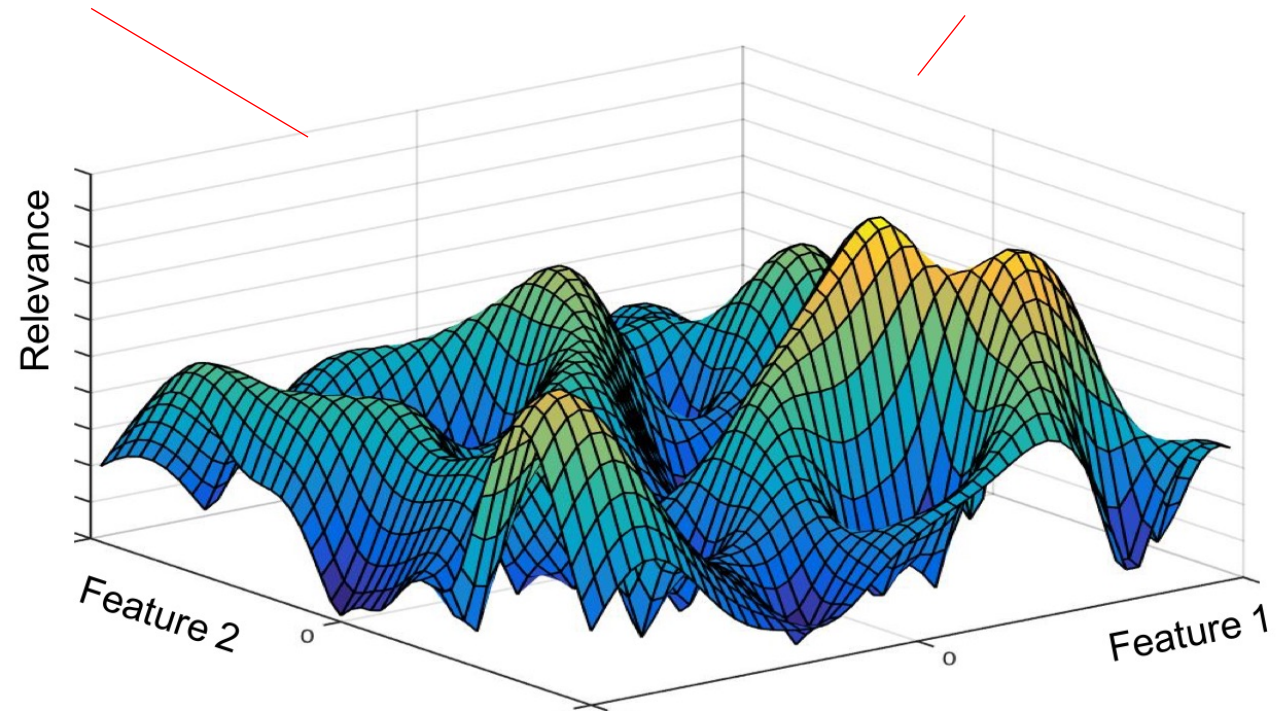
YES

How about combining it with physical exercise?

NO, I am too old for that, except if it is easy to do.

Generate and select the most informative question(s)

Update belief based on the answer



Generate and select clarifying questions



Product works and looks like new. Backed by the 90-day Amazon Renewed Guarantee.

Renewed products work and look like new. These pre-owned products are not Apple certified but have been inspected and tested by Amazon-qualified suppliers. Box and accessories (no headphones included) may be generic. Wireless devices come with the 90-day Amazon Renewed Guarantee. [Learn more](#)

- Fully unlocked and compatible with any carrier of choice (e.g. AT&T, T-Mobile, Sprint, Verizon, US-Cellular, Cricket, Metro, etc.).
- The device does not come with headphones or a SIM card. It does include a charger and charging cable that may be generic, in which case it will be UL or Mfi ("Made for iPhone") Certified.
- Inspected and guaranteed to have minimal cosmetic damage, which is not noticeable when the device is held at arm's length.
- Successfully passed a full diagnostic test which ensures like-new functionality and removal of any prior-user personal information.
- Tested for battery health and guaranteed to have a minimum battery capacity of 80%.

- ✓ J. Zou, Y. Chen, E. Kanoulas, Towards Question-based Recommender Systems. *SIGIR 2020*
- ✓ J. Zou, E. Kanoulas, Towards Question-Based High-Recall Information Retrieval: Locating the Last Few Relevant Documents for Technology Assisted Reviews, *TOIS 2020*
- ✓ J. Zou, E. Kanoulas. Learning to Ask: Question-based Sequential Bayesian Product Search. *CIKM 2019*

The world of entities



Product works and looks like new . Backed by the 90-day **Amazon Renewed Guarantee** . Renewed products work and look like new . These pre - owned products are not **Apple** certified but have been inspected and tested by **Amazon** - qualified suppliers . Box and accessories (no headphones included) may be generic . Wireless devices come with the 90-day **Amazon Renewed Guarantee** . Learn more Fully unlocked and compatible with any carrier of choice (e.g. AT&T , T - Mobile , **Sprint** , **Verizon** , **US** - **Cellular** , Cricket , **Metro** , etc .) . The device does not come with headphones or a **SIM** card . It does include a charger and charging cable that may be generic , in which case it will be **UL** or **Mfi** (" Made for **iPhone** ") **Certified** . Inspected and guaranteed to have minimal cosmetic damage , which is not noticeable when the

- ✓ V Provatorova, S Vakulenko, E Kanoulas, K Dercksen, J Hulst: Named Entity Recognition and Linking on Historical Newspapers, *CLEF 2020*
- ✓ C Wu, E Kanoulas, M de Rijke: Learning entity-centric document representations using an entity facet topic model, *IPM 2020*
- ✓ C Wu, E Kanoulas, M de Rijke: It all starts with entities: A Salient entity topic model. *NLE 2020*
- ✓ C Wu, E Kanoulas, M de Rijke, W Lu: WN-Salience: A Corpus of News Articles with Entity Salience Annotations. *LREC 2020*



Product works and looks like new . Backed by the 90-day **Amazon Renewed Guarantee** . Renewed products work and look like new . These pre - owned products are not **Apple** certified but have been inspected and tested by **Amazon** - qualified suppliers . Box and accessories (no headphones included) may be generic . Wireless devices come with the 90-day **Amazon Renewed Guarantee** . Learn more Fully unlocked and compatible with any carrier of choice (e.g. AT&T , T - Mobile , **Sprint** , **Verizon** , **US** - **Cellular** , Cricket , **Metro** , etc .) . The device does not come with headphones or a **SIM** card . It does include a charger and charging cable that may be generic , in which case it will be **UL** or **Mfi** (“ Made for **iPhone** ”) **Certified** . Inspected and guaranteed to have minimal cosmetic damage , which is not noticeable when the

$$e_l = \arg \min_e \left| \sum_{d \in u_l} (2\mathbb{1}\{e(d) = 1\} - 1) \pi^*(d) \right| - \gamma * R(e)$$

Binary search

Trained question reward

Trained system belief

Update the belief based on the answer



Product works and looks like new . Backed by the 90-day **Amazon Renewed Guarantee** . Renewed products work and look like new . These pre - owned products are not **Apple** certified but have been inspected and tested by **Amazon** - qualified suppliers . Box and accessories (no headphones included) may be generic . Wireless devices come with the 90-day **Amazon Renewed Guarantee** . Learn more Fully unlocked and compatible with any carrier of choice (e.g. AT&T , T - Mobile , **Sprint** , **Verizon** , **US** - **Cellular** , Cricket , **Metro** , etc .) . The device does not come with headphones or a **SIM** card . It does include a charger and charging cable that may be generic , in which case it will be **UL** or **Mfi** (" Made for **iPhone** ") **Certified** . Inspected and guaranteed to have minimal cosmetic damage , which is not noticeable when the

Multinomial distribution
over products

Dirichlet distribution

$$\pi^*(d) \xleftrightarrow{\text{conjugate}} \mathbb{P}_{n+1} = \text{Dir}(\alpha + \sum_{j=0}^n Z_j)$$

Generate ranking list

Example

Salomon Trail IV Tight



amazon.de
All active clothing

Deliver to je Amsterdam 1094GJ
Shop by Department
je's Amazon Today's Deals Gift Cards Sell Help
Fashion SALE bis -60%
Hello, je Your Account Try Prime Your Lists Shopping Basket
Amazon.de SALE Warehouse Deals Fashion-Sale Family Students Subscribe & Save Pantry Gift Central Apps Amazon Assistant
97-112 of over 2,000 results for "active clothing" Sort by Relevance

Show results for

Women's Clothing
Women's Sports Knickers & Bras
Women's Sportswear
Women's Tops & T-Shirts
Women's T-Shirts
Women's Vest Tops

Sports & Outdoors
Men's Exercise & Fitness Tracksuits
Women's Fitness Trousers

Women's Shoes
Women's Athletic & Outdoor Shoes

Automotive
Merchandise Apparel
See All 26 Departments

Refine by

Promotions (What's this?)
☐ Subscribe & Save

Delivery Option (What's this?)
☒ prime
☐ Free shipping for deliveries over EUR 29

Brand
☐ LotMart
☐ XXR
☐ ZEZKT
☐ Banqert

Elten 72752-43 Size 43 ESD S2 "Ambition Boa Low" Safety Shoe - Multi-Colour
by ELTEN
EUR 119.99 prime
Eligible for FREE Delivery.
Only 1 left in stock - order soon.
More buying choices
EUR 96.21 (5 new offers)
★★★★★ 13

Kanpola T-Shirts Men's Short Sleeve Fashion Blouse Tops Basic Fitness T-Shirt
by Kanpola Herren T-Shirts
EUR 3.99 - EUR 8.64
Räumungsverkauf Prime Day and 5 more promotions

Bracco Fulcrum Breathable Back Support
by Bracco
EUR 20.30 prime
Get it by Tuesday, Jun 19
Eligible for FREE Delivery.
★★★★☆ 14
2 kaufen - 10% Rabatt See Details

Rank

| | | |
|--------------------------------|-----|------|
| 1. Is the product for women? | yes | 1753 |
| 2. Is it a slim fit product? | yes | 273 |
| 3. Is it a stretch product? | yes | 8 |
| 4. Is it a technology product? | no | 3 |

User perspective on clarifying questions

Step 1: Category selection


Imagine that you want to buy a product. Please select the product category you are most familiar with (e.g., most frequently purchased category).

Category:

To Step 2

Step 2: Target product assignment

Imagine that you want to buy the target product shown below, **please read the product title & description very carefully**. After you click the "Next step" button, you will interact with our algorithm by answering YES/NO questions:



Product Title: Hog Wild Fish Sticks (Sold Individually)
Product Description: Kids love this pre-scissors skills activity set of 1-piece chopsticks! Use the tongs with oral-motor activities. Simply set up small toys, easy-grip foods or cotton balls for kids to transfer across a midline. ...

Change target Next step

Step 3: Find the target product

Please answer the following algorithmically constructed questions according to the title and description of your target product shown on the last page. After you click the "next" button, the algorithm will select the next question to ask. When you wish to stop answering questions you can click the "stop" button.





Question: Is "rosewood" relevant to the product you are looking for?

- .
- .
- .

☐ Yes
☒ No
☐ Not Sure

Next Stop

Ranking list of search results, from top 1 (left) - top 4(right):

| | | | |
|--|--|--|--|
| Hog Wild Fish Sticks (Sold Individually):  | Bone & Rosewood Chopsticks:  | Fred & Friends Good Fortune Chopsticks:  | 2pk Green Pot Holders/Trivet Set:  |
|--|--|--|--|

Step 4: Questionnaire

Q1: Did you find our question-based system helpful towards locating the target product?

Q2: Will you use such a question-based system for product search or recommendation in the future?

Q3: What was your experience using the question-based system?

Q4: How many questions are you willing to answer for locating your target product?

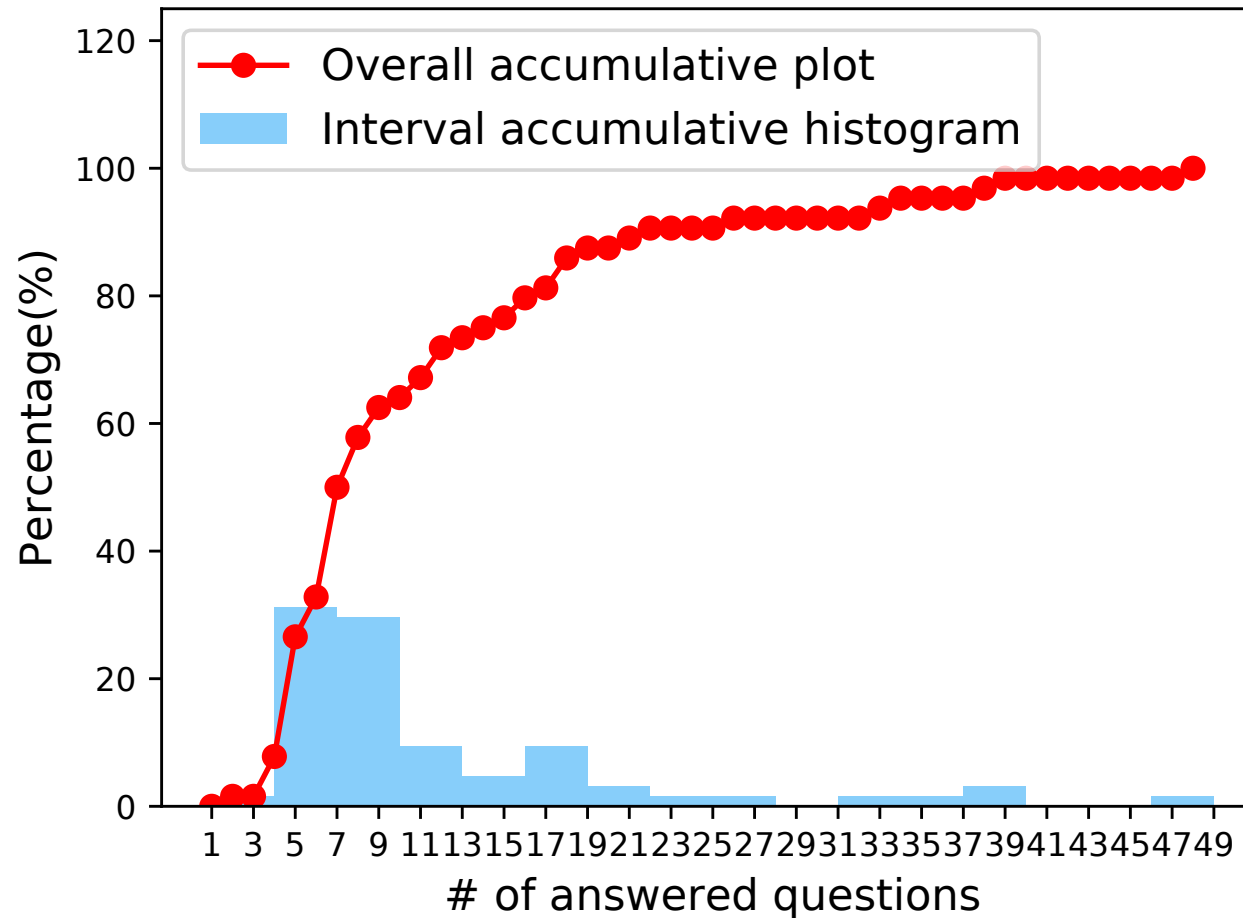
Q5: Why did you click the "Stop" button to stop answering in the last step?

Q6: If selected "other" in Q5, please specify:

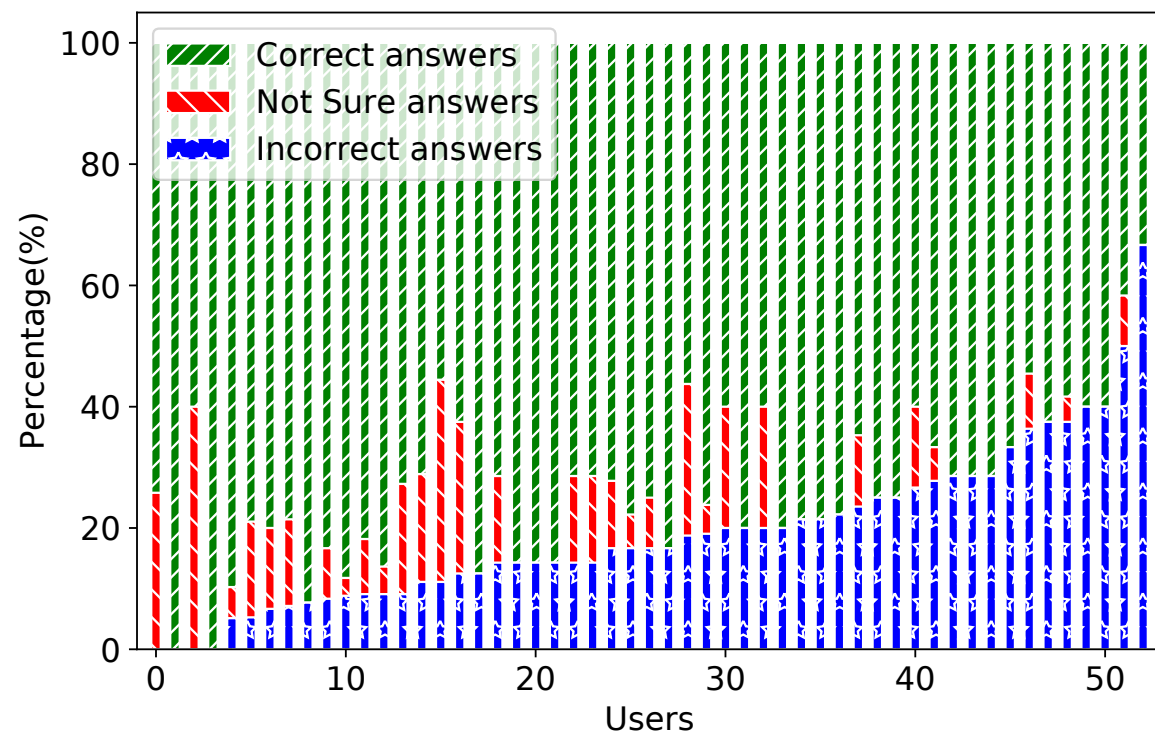
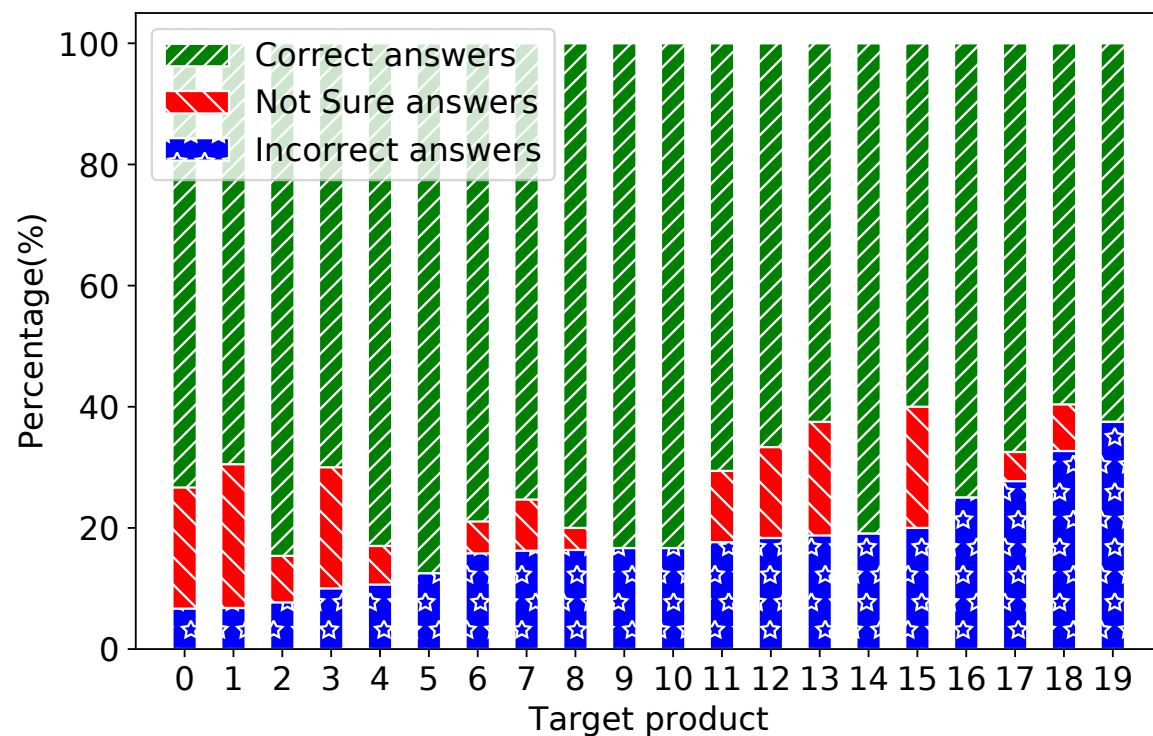
Q7: Are the generated questions easy to answer?

All done

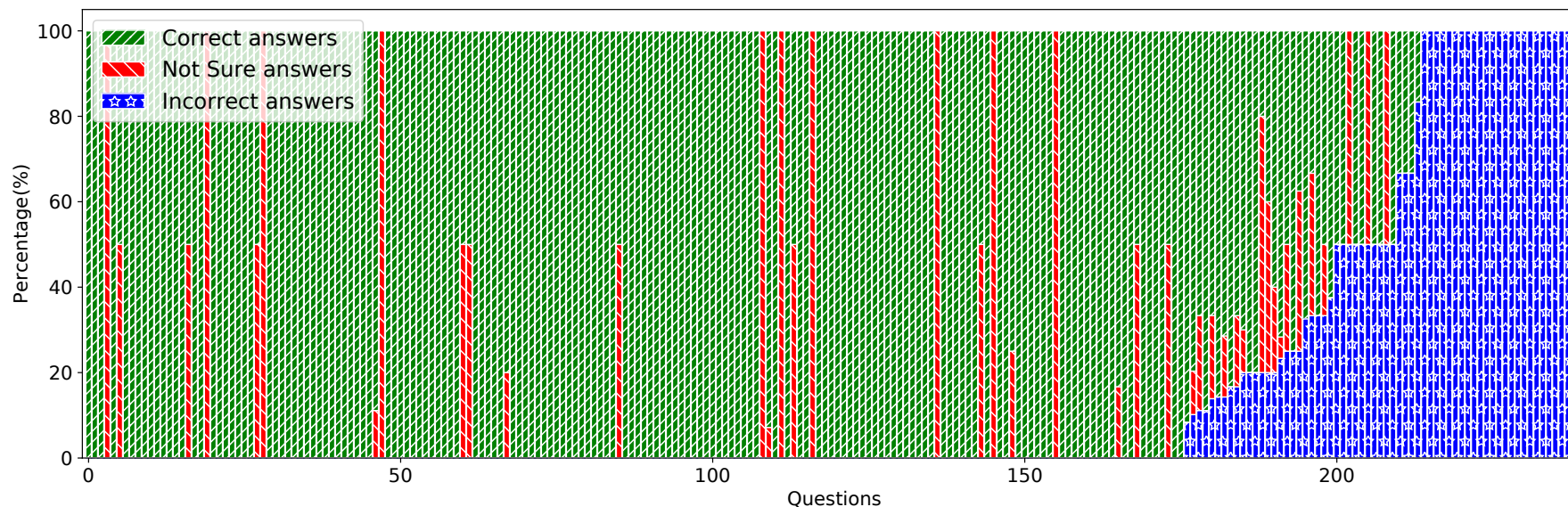
User perspective on clarifying questions



User's ability to answer clarifying questions



User's ability to answer clarifying questions



$$e_l = \arg \min_e \left| \sum_{d \in u_l} (2\mathbb{1}\{e(d) = 1\} - 1)\pi^*(d) \right| - \gamma * R(e) + 2\beta * h(e)$$

Trained question reward

Trained system belief

Noise rate

Challenge: How to generate clarifying questions

- Train constrained question generation
- Identify what discriminates documents and use as constraints

Challenge: Update the belief based on the answer

| | | $\Delta NDCG$ |
|-----------|--|---------------|
| Inf. need | <i>Drafting of Declaration of Independence</i> | |
| Q_0 | all men are created equal | |
| CQ | - Would you like to learn more about Thomas Jefferson ? | +18.53 |
| Answer | - No | |
| Inf. need | <i>Information about Atari arcade games</i> | |
| Q_0 | atari | |
| CQ | - Would you like to play atari arcade games online? | +15.97 |
| Answer | - No | |

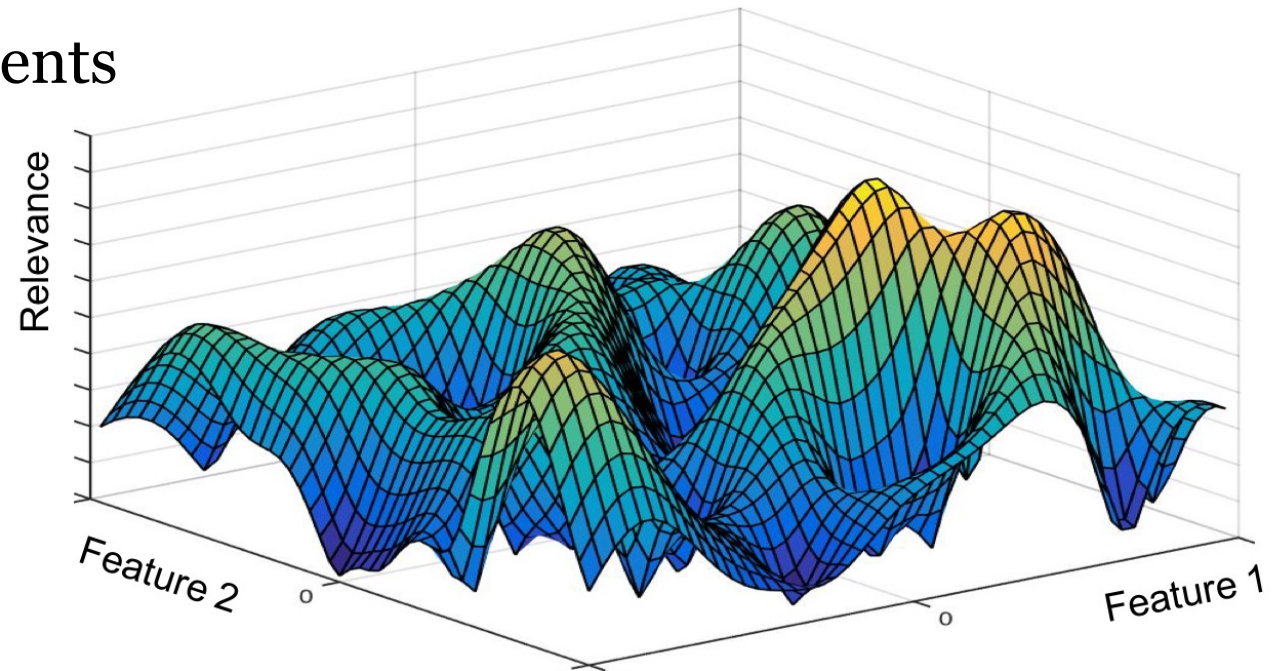
| | | |
|-----------|---|--------|
| Inf. need | <i>How is a credit score determined?</i> | |
| Q_0 | credit report | |
| CQ | - Would you like to know about the process of disputing credit report? | +13.57 |
| Answer | - No | |
| Inf. need | <i>Website of HobbyTown USA</i> | |
| Q_0 | hobby stores | |
| CQ | - Would you like to know the hours of operation? | +11.37 |
| Answer | - No | |
| Inf. need | <i>Find the TSA website, offering air travel tips</i> | |
| Q_0 | air travel information | |
| CQ | - Would you like to book a flight ? | +8.91 |
| Answer | - No | |

- ✓ A Krasakis, M Aliannejadi, N Voskarides, E Kanoulas: Analysing the Effect of Clarifying Questions on Document Ranking in Conversational Search. *ICTIR 2020*

4. Mixed Initiative: Complex Strategies

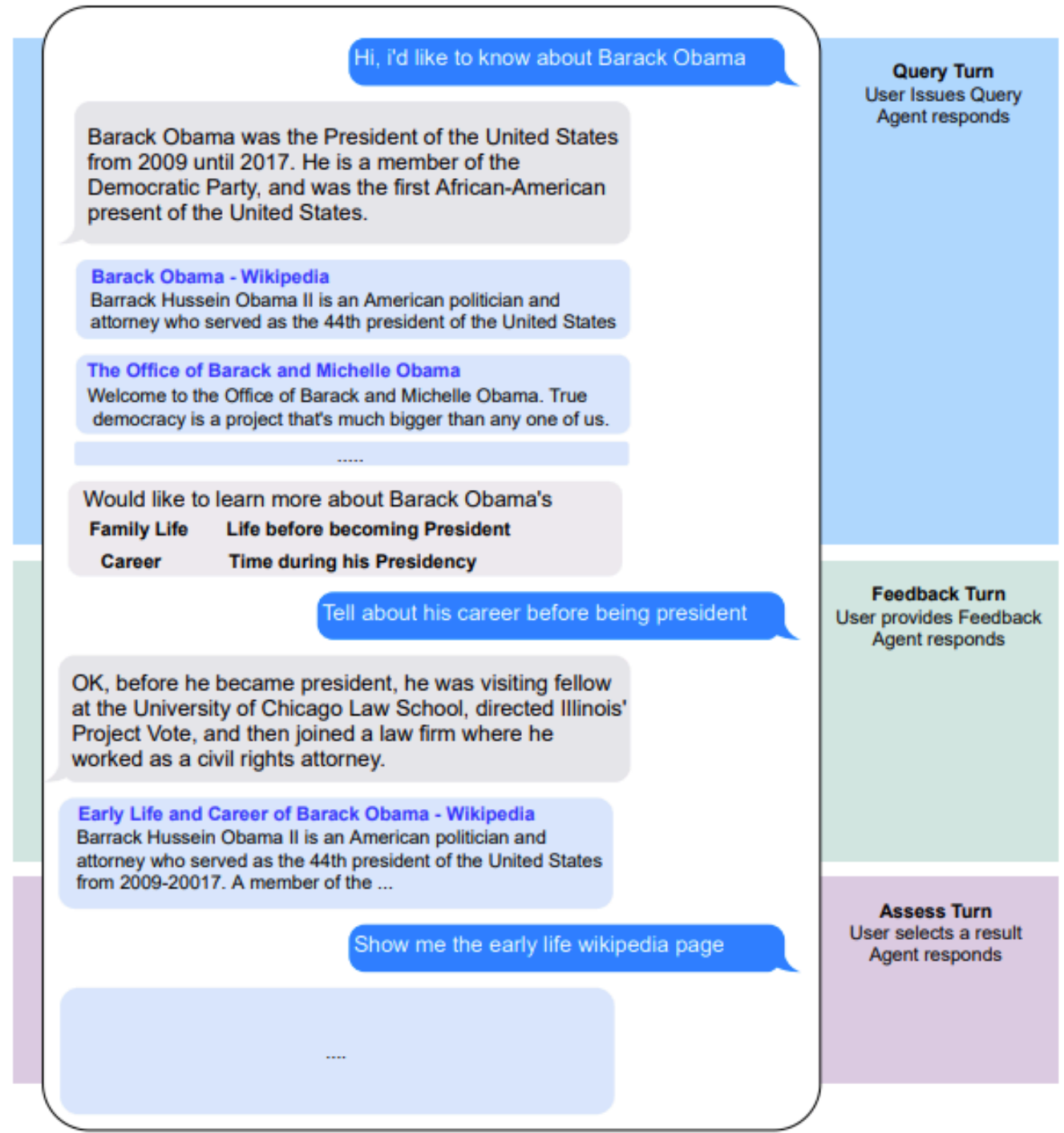
1. Provide ranking of document/passages
2. Provide direct answer
3. Ask clarifying question
 1. Relevance over items/documents
 2. Natural language questions
 3. Suggested questions

Decide what is the
optimal next action to
take



Example of Mixed Initiative

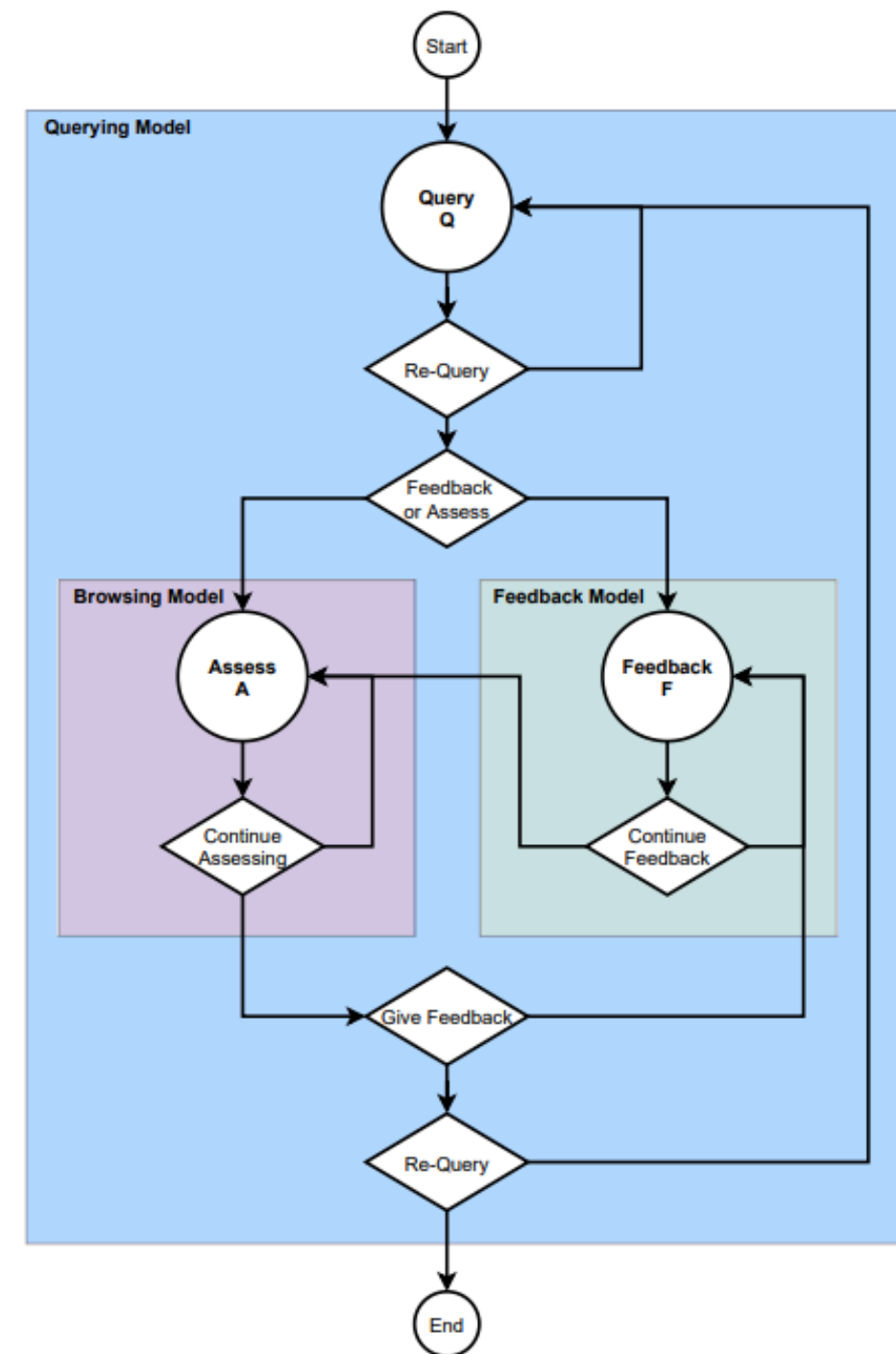
- ✓ M Aliannejadi, et al. Modelling and Measuring Conversational Search: An Analysis of Mixed Initiatives and Conversational Strategies (*under review*)



Conversational search agent interface

1. a voice only CSA often via a virtual assistant
2. a chat based CSA (that is in-situ within a platform like Slack or Telegram)
3. an augmented search engine interface, or
4. a multi-modal virtual assistants.

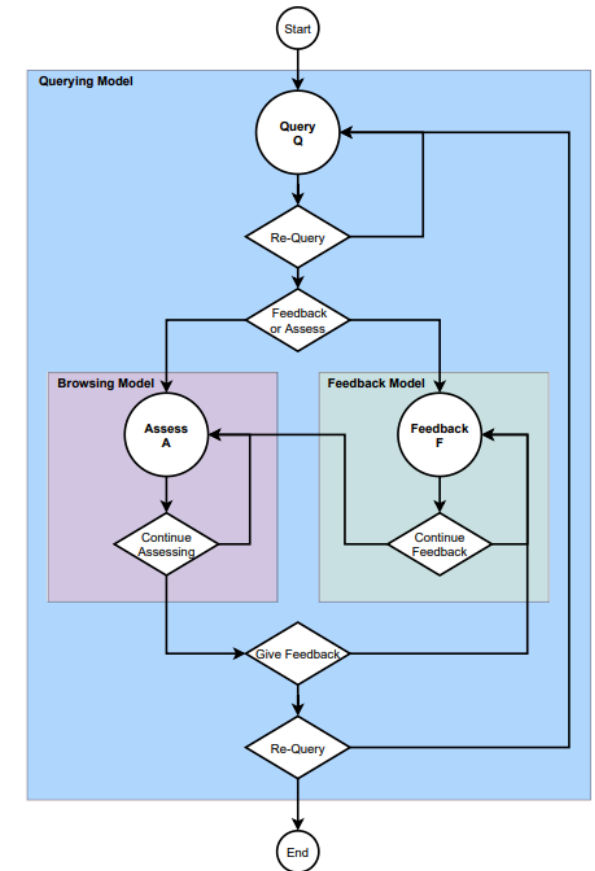
User Interaction Model of Conversational Search



- ✓ M Aliannejadi, et al. Modelling and Measuring Conversational Search: An Analysis of Mixed Initiatives and Conversational Strategies (*under review*)

User Interaction Model of Conversational Search

| SIMULATION | Answer Question | Use Suggested Query |
|-------------------|-----------------|---------------------|
| Feedback First | | |
| Feedback After | | |



- ✓ M Aliannejadi, et al. Modelling and Measuring Conversational Search: An Analysis of Mixed Initiatives and Conversational Strategies (*under review*)

Evaluating the Cost and Benefit of CS

Cumulative gain: $G(t_i, \dots, t_T) = \sum_{i=1}^{i=T} g(t_i)$

Cumulative effort: $C(t_i, \dots, t_T) = \sum_{i=1}^{i=T} c(t_i)$

Rate of gain: $R(t_i, \dots, t_T) = \frac{G(t_i, \dots, t_T)}{C(t_i, \dots, t_T)}$

Grounding Costs

- Time as an estimation of cost.

- Crowdsourcing:

- Feedback

- Assessment

What you do you want to know about COVID-19?

Symptoms

Vaccine

Restrictions

Number of cases

Not interested

Current Information about COVID-19

<https://www.rivm.nl/en/novel-coronavirus-covid-19/current-information>

You are looking for current information about the novel coronavirus (COVID-19) in the Netherlands. Weekly updates with an overview of COVID-19 confirmed cases, number of ...

Looks very useful

Might be useful

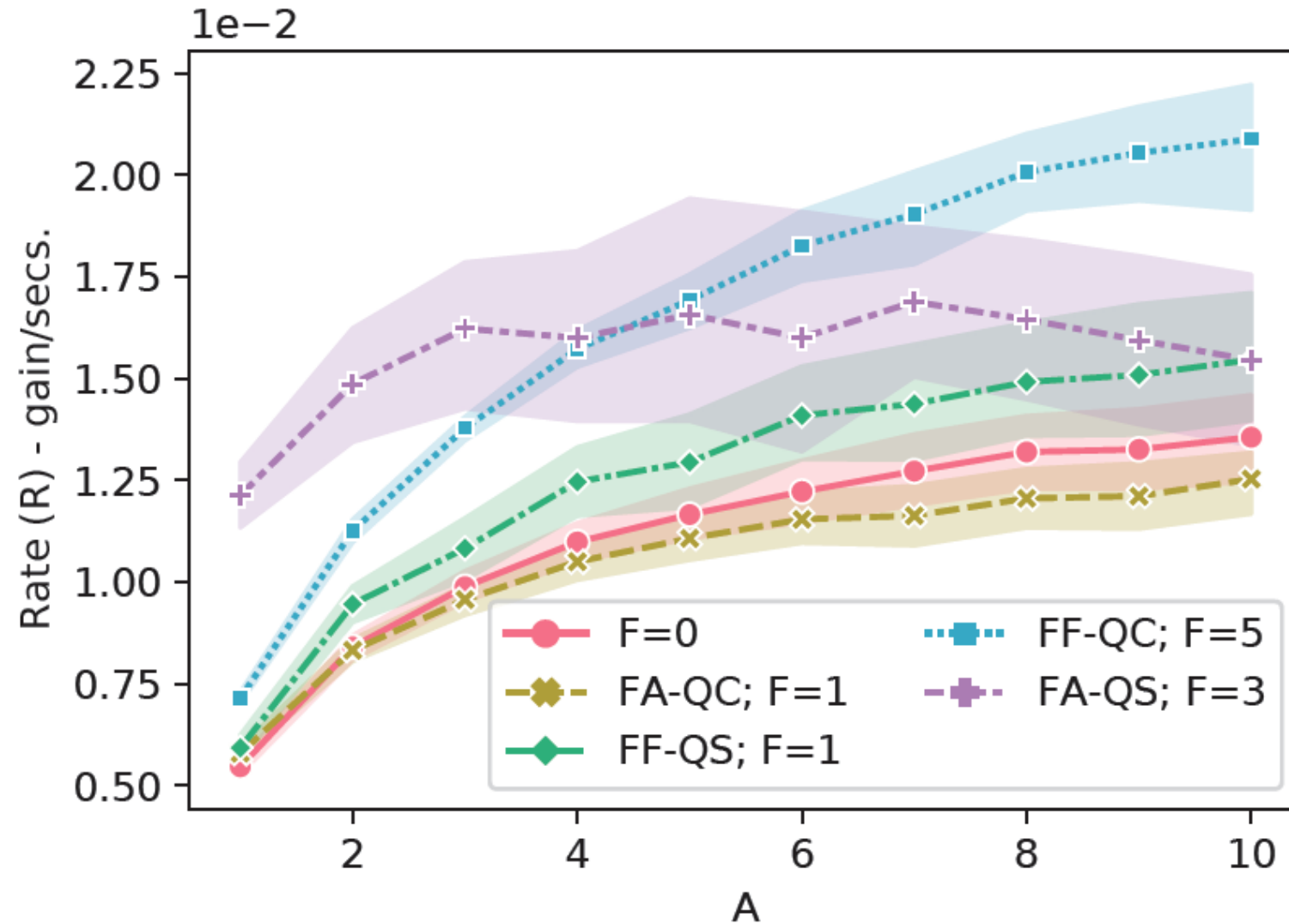
NOT useful

Grounding Costs

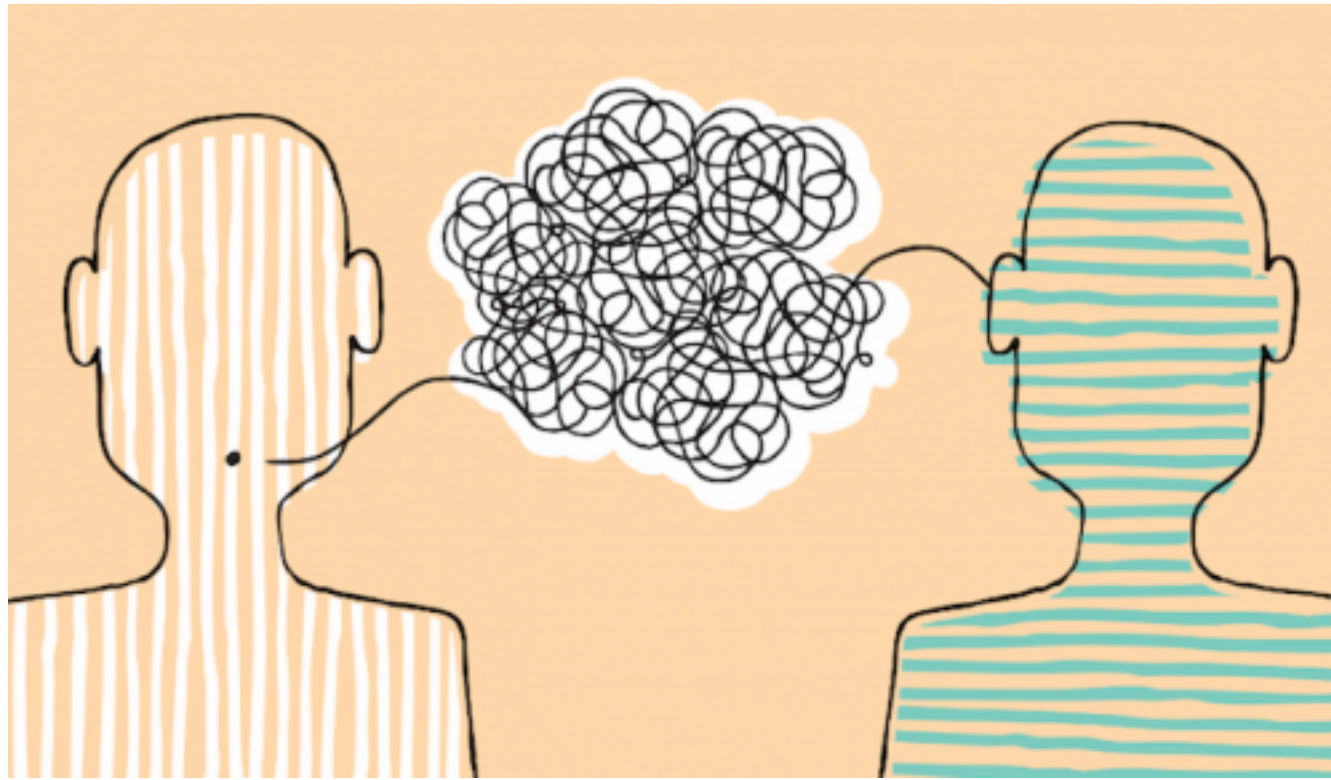
- Time as an estimation of cost.

| | Submit Query | Assess Snippet | Assess Document | Provide Feedback |
|--------|-----------------|-------------------|--------------------|---------------------|
| Mean | 29.33 | 6.26 | 16.96 | 8.33 |
| Median | 22.59 | 3.63 | 10.22 | 5.77 |

Analyzing Strategies



Analyzing Mixed Initiative



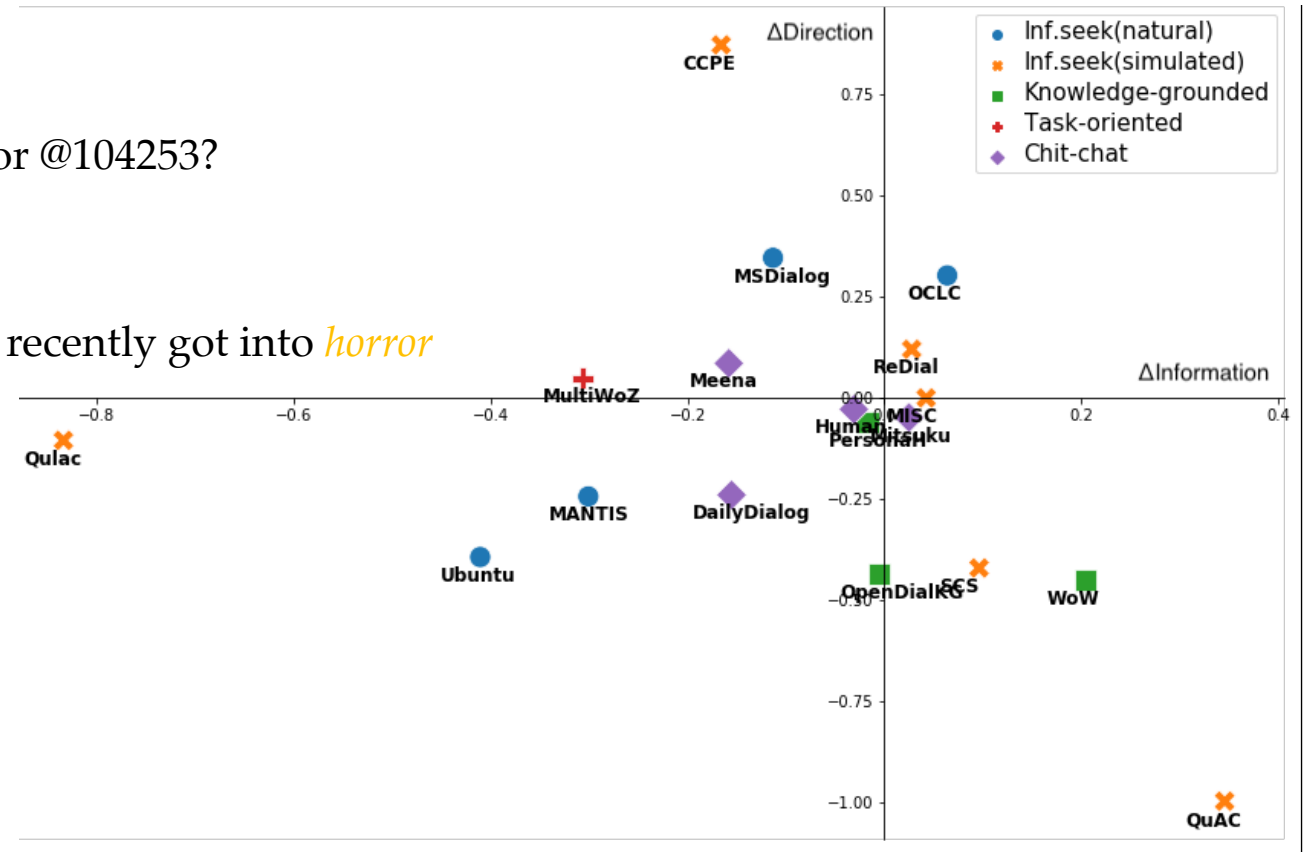
Analyzing Mixed Initiative

| Dataset | Dialogues | Domain | Type | Subtype | Modality | Source | Setup | Dial. len. | Utt. len. | Utt./turn |
|--------------------|-----------|-----------|----------------------|---------|---------------|-------------------------|----------------|-----------------|-----------------|--------------|
| MSDialog [23] | 35,500 | tech | info-seeking | IN | text | forum | natural | 9 (25) | 67 (87) | 1 (3) |
| MANTIS [22] | 1,400 | multi | info-seeking | IN | text | forum | natural | 4 (1) | 98 (160) | 1 (0) |
| OCLC ^a | 560 | library | info-seeking | IN | text | chat | natural | 25 (20) | 12 (15) | 1 (1) |
| Ubuntu [20] | 1,200,000 | tech | info-seeking | IN | text | chat | natural | 6 (8) | 10 (9) | 1 (1) |
| SCSdata [35] | 37 | web | info-seeking | IS | speech | chat | simulated | 27 (21) | 16 (25) | 1 (0) |
| MISC [34] | 110 | web | info-seeking | IS | speech | chat | simulated | 120 (47) | 7 (7) | 2 (2) |
| ReDial [17] | 10,000 | movies | info-seeking | IS | text | chat | simulated | 18 (5) | 6 (5) | 1 (0) |
| CCPE [25] | 502 | movies | info-seeking | IS | text | chat | simulated | 23 (6) | 12 (13) | 1 (0) |
| Qulac [?] | 10,277 | web | info-seeking | IS | text | task^b | simulated | 3 (0) | 8 (3) | 1 (0) |
| QuAC [7] | 11,600 | Wikipedia | info-seeking | IS | text | chat | simulated | 14 (4) | 9 (7) | 1 (0) |
| MultiWOZ [?] | 10,000 | multi | task-oriented | TO | text | chat | simulated | 13 (5) | 13 (6) | 1 (0) |
| Meena/Mitsuku* [1] | 100 | open | social | CC | text | chat | simulated | 18 (4) | 8 (10) | 1 (0) |
| Meena/Meena* [1] | 91 | open | social | CC | text | chat | simulated | 19 (5) | 6 (4) | 1 (0) |
| Meena/Human [1] | 95 | open | social | CC | text | chat | simulated | 15 (2) | 13 (10) | 1 (0) |
| DailyDialog [18] | 11,000 | multi | social | CC | text | book^c | simulated | 7 (4) | 13 (10) | 1 (0) |
| PersonaH [29] | 102 | personal | social | KG | text | chat | simulated | 12 (0) | 9 (4) | 1 (0) |
| WoW [12] | 22,000 | Wikipedia | social | KG | text | chat | simulated | 9 (1) | 16 (7) | 1 (0) |
| OpenDialKG [21] | 13,800 | Freebase | social | KG | text | chat | simulated | 6 (2) | 12 (6) | 1 (0) |

- ✓ S Vakulenko, E Kanoulas, M de Rijke, An Analysis of Mixed Initiative and Collab. in Information-Seeking Dialogues, *SIGIR 2020*
- ✓ S Vakulenko, E Kanoulas, M de Rijke, Large Scale Analysis of Mixed Initiative in Information-Seeking Dialogues for Conversational Search (*under review*)

Analyzing Mixed Initiative

- (R) Hey! What kind of *movies* do you like to watch?
- (H) I'm really big on *indie romance and dramas*
- (R) Ok what's your favorite *movie*?
- (R) Staying with that *genre*, have you seen @88487 or @104253?
- (R) Those are two really good ones
- (H) When I was a kid I liked *horror* like @181097
- (R) @Misery is really creepy but really good. I only recently got into *horror*



- ✓ S Vakulenko, E Kanoulas, M de Rijke, An Analysis of Mixed Initiative and Collab. in Information-Seeking Dialogues, *SIGIR 2020*
- ✓ S Vakulenko, E Kanoulas, M de Rijke, Large Scale Analysis of Mixed Initiative in Information-Seeking Dialogues for Conversational Search, *under review*

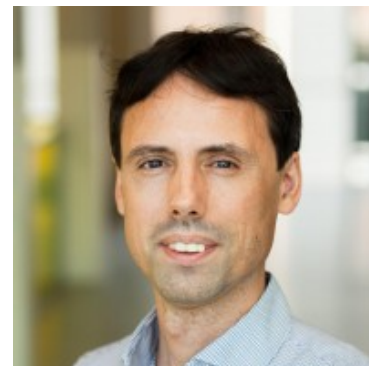
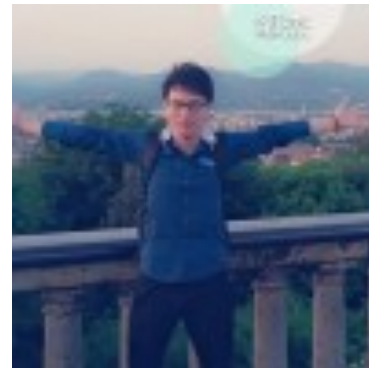
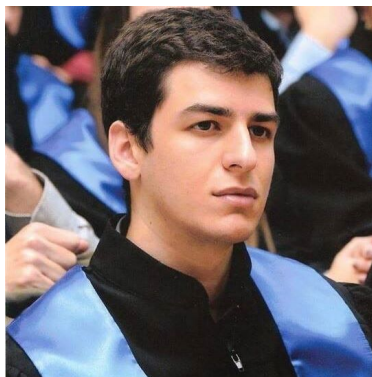
Summing up ...

Agent

1. Provide ranking of document / passages
2. Provide direct answer
3. Ask clarifying question
 1. Relevance over items / documents
 2. Natural language questions
 3. Suggested questions

Evaluation

1. Simulate mixed initiative
2. Decide costs / gains
3. Develop evaluation measures



Thanks!