A Large Scale Analysis of Mixed Initiative in Information-Seeking Dialogues for Conversational Search

Svitlana Vakulenko*, Evangelos Kanoulas and Maarten de Rijke
University of Amsterdam, *Amazon Alexa AI

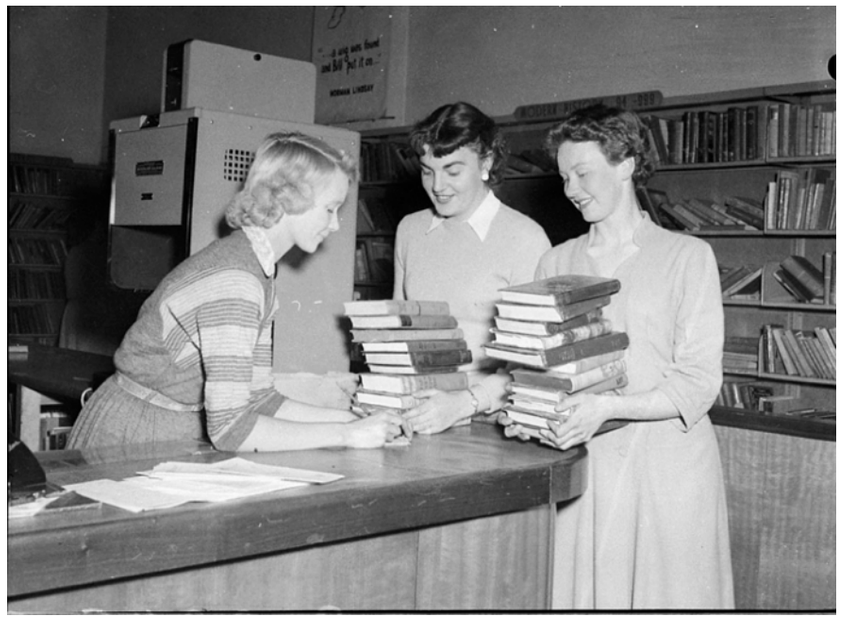
Conversational Search

- conversational information seeking (CIS)
- conversational information retrieval (CIR)
- the task of providing relevant information using a conversational interface
- automating an **information-seeking dialogue**

Hamed Zamani, Johanne R. Trippas, Jeff Dalton, Filip Radlinski: Conversational Information Seeking 2022

Reference Interview

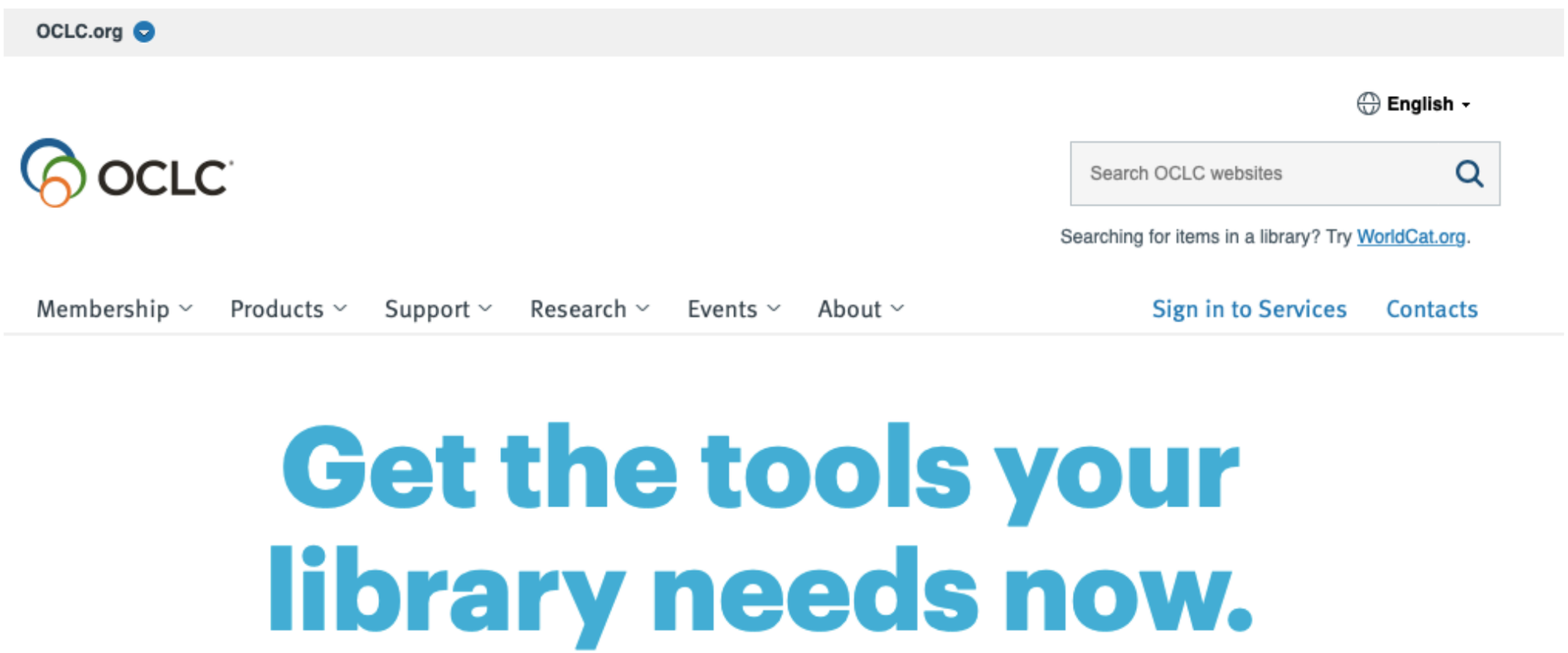
IS: Do you have the book that *they showed* in the filmstrip?
MS: Yes. Uhm. (SILENCE) You're thinking of the one that was on . . . ?
IS: Samurai.
MS: Hhh. The one. It was the picture of the guy swinging the samurai sword.
IS: [Yeah, samurai sword.
MS: Was that Sign of the Chrysanthemum?
I: [Yaa.



Solomon, Paul. "Conversation in information-seeking contexts: A test of an analytical framework." *Library & Information Science Research.* 1997

Data

- global library cooperative (**OCLC**) www.oclc.org
  - 560 anonymized **virtual reference interviews**
  - conducted on-line by professional librarians



Research Questions

RQ1 What are the structural properties of an **information-seeking dialogue**?

RQ2 Which public dialogue **datasets** are similar to a reference interview with a professional librarian?

RQ3 How to **compare dialogues** automatically?

Contributions

* **Framework** for automated dialogue analysis of MI patterns

-> ConversationShape

* **Results** of the large-scale dialogue analysis of MI patterns

-> Dialogue types

Data

- +15 public dialogue datasets
- >150k dialogue transcripts

Dataset	Dialogues	Domain	Type	Subtype	Modality	Source	Setup
MSDialog [27]	35,500	tech	info-seek	IN	text	forum	**natural**
MANTIS [26]	1,400	multi	info-seek	IN	text	forum	**natural**
OCLC[a]	560	library	info-seek	IN	text	chat	**natural**
Ubuntu [24]	1,200,000	tech	info-seek	IN	text	chat	**natural**
SCSdata [42]	37	web	info-seek	IS	**speech**	chat	simulated
MISC [41]	110	web	info-seek	IS	**speech**	chat	simulated
ReDial [21]	10,000	movies	info-seek	IS	text	chat	simulated
CCPE [29]	502	movies	info-seek	IS	text	chat	simulated
Qulac [2]	10,277	web	info-seek	IS	text	**task**[b]	simulated
QuAC [10]	11,600	Wikipedia	info-seek	IS	text	chat	simulated
MultiWOZ [9]	10,000	multi	**task-orient**	TO	text	chat	simulated
Meena/Mitsuku* [1]	100	open	social	CC	text	chat	simulated
Meena/Meena* [1]	91	open	social	CC	text	chat	simulated
Meena/Human [1]	95	open	social	CC	text	chat	simulated
DailyDialog [22]	11,000	multi	social	CC	text	**book**[c]	simulated
PersonaH [34]	102	personal	social	KG	text	chat	simulated
WoW [15]	22,000	Wikipedia	social	KG	text	chat	simulated
OpenDialKG [25]	13,800	Freebase	social	KG	text	chat	simulated

ConversationShape

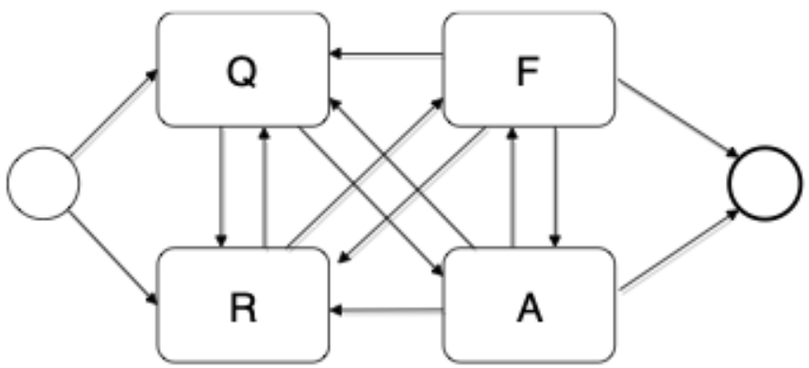
- **fingerprinting** for dialogue representation
- **dialogue flow** model for initiative dynamics
- **asymmetry metrics** for initiative distribution

Dialogue Fingerprint

Fingerprint					Utterance	Terms
Role	Type	Length	Repetitions			
A	H	4	0	0	Hey!	{hey}
A	I	41	1	0	What kind of movies do you like to watch?	{watch, **movi**}
S	N	42	0	0	I'm really big on indie romance and dramas	{romanc, drama, indi}
A	I	30	1	0	Ok what's your favorite movie?	{favorit, **movi**}
A	I	56	0	0	Staying with that genre, have you seen @88487 or @104253	{genr, stay, 88487, 104253}
A	N	30	0	0	Those are two really good ones	{}
S	N	44	0	1	When I was a kid I liked horror like @181097	{181097, kid, **horror**}
A	N	41	0	0	@Misery is really creepy but really good.	{miseri, creepi}
A	N	32	0	1	I only recently got into horror.	{**horror**}

Dialogue Flow

- utterance classification with RoBERTa
  - QuAC, Qulac, SPAADIA and NPS chat
  - 86K samples: 77.5K for training and 8.6K for testing
  - macro-average F1 score of **0.942**

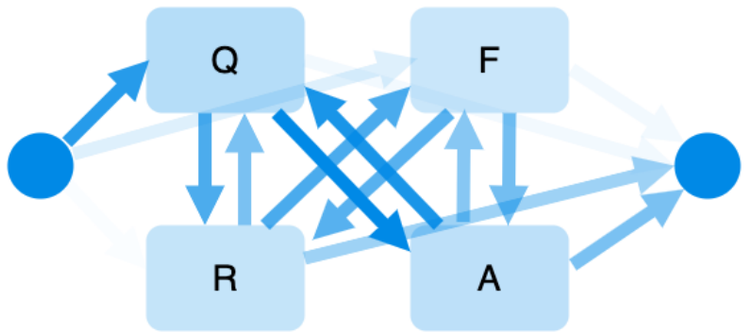


Svitlana Vakulenko, Kate Revoredo, Claudio Di Ciccio, Maarten de Rijke: QRFA: A Data-Driven Model of Information-Seeking Dialogues. ECIR (1) 2019: 541-557

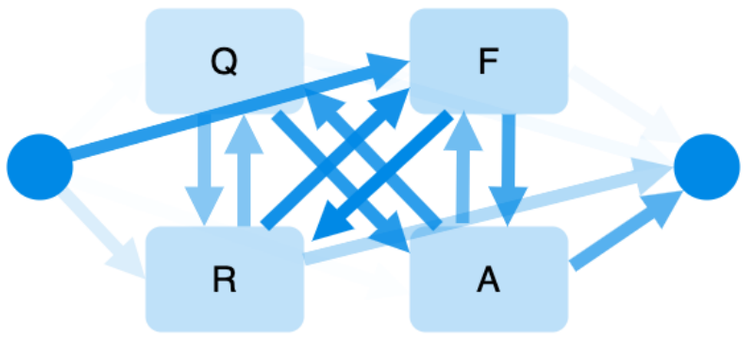
Sharing Dialogues QA~RF

CC

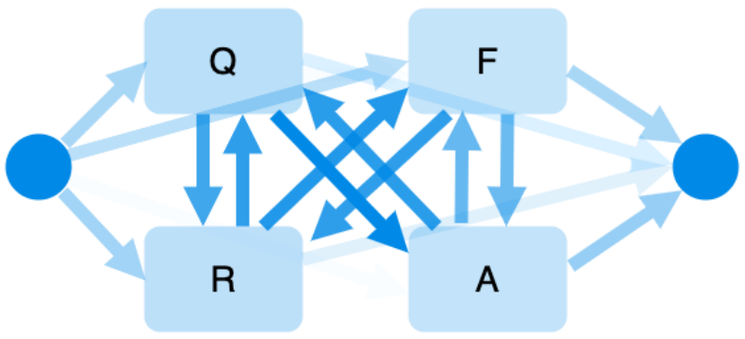
Meena
Mitsuku



Meena
Meena

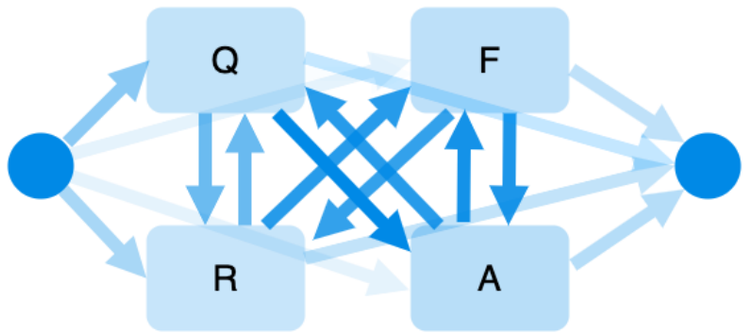


Meena
Human



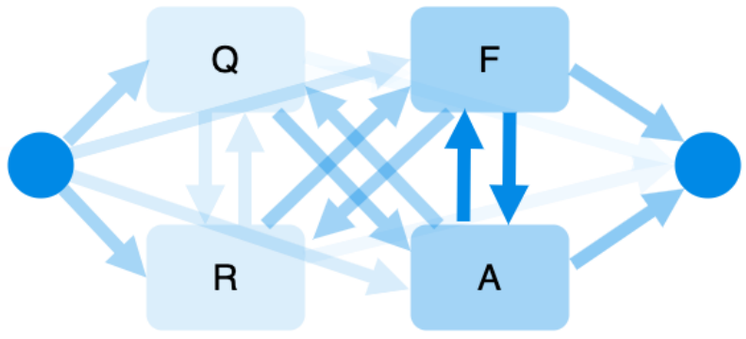
KG

PersonaH

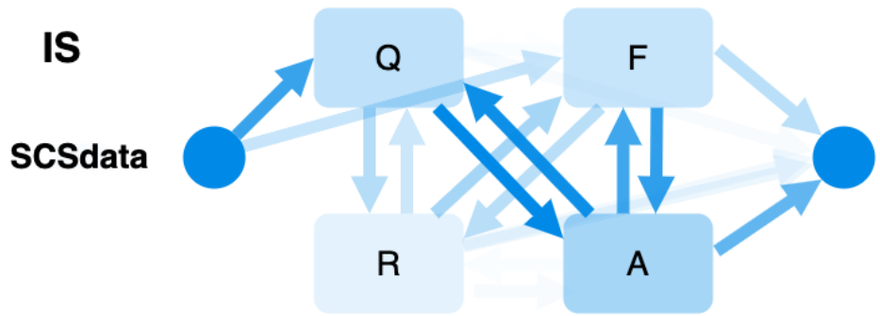


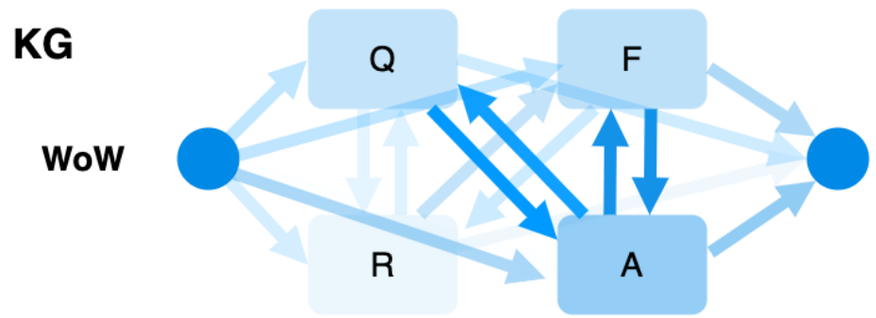
IS

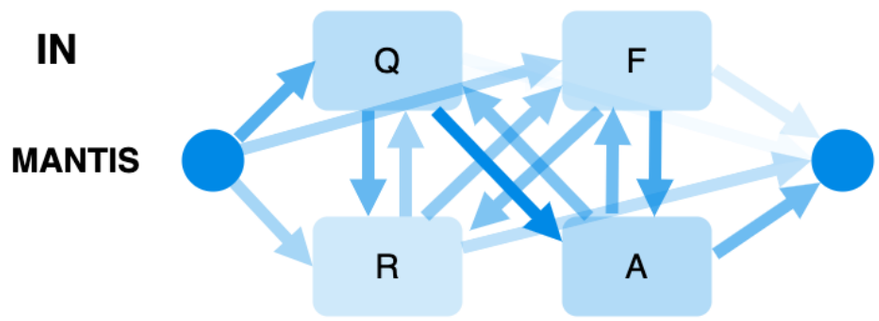
MISC

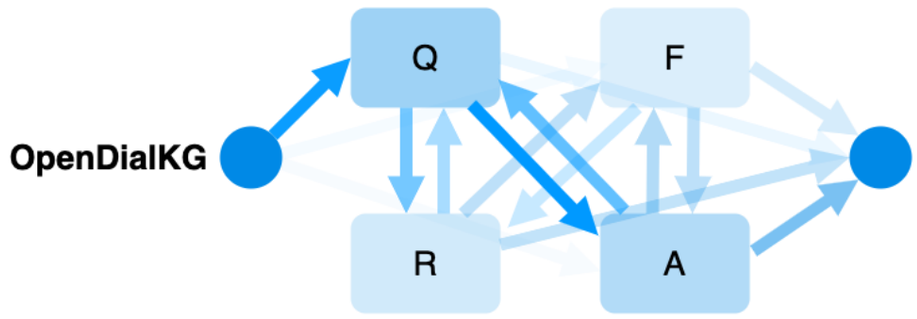


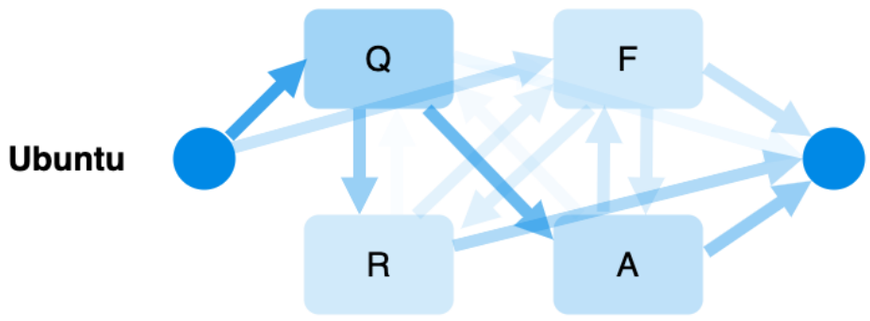
Search Dialogues QA>RF

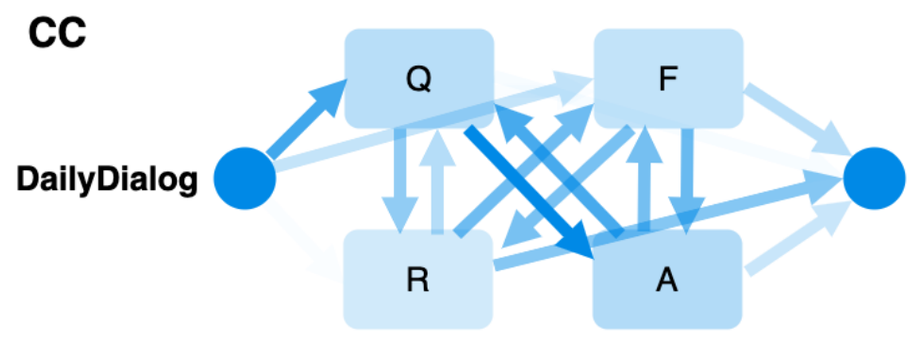




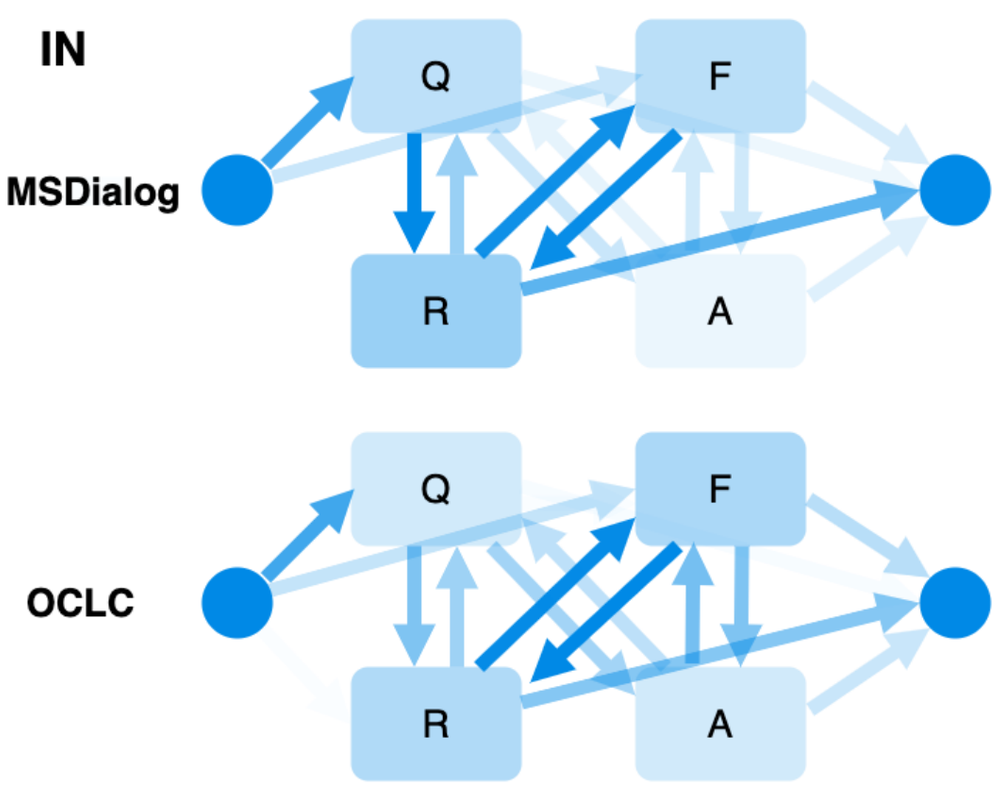


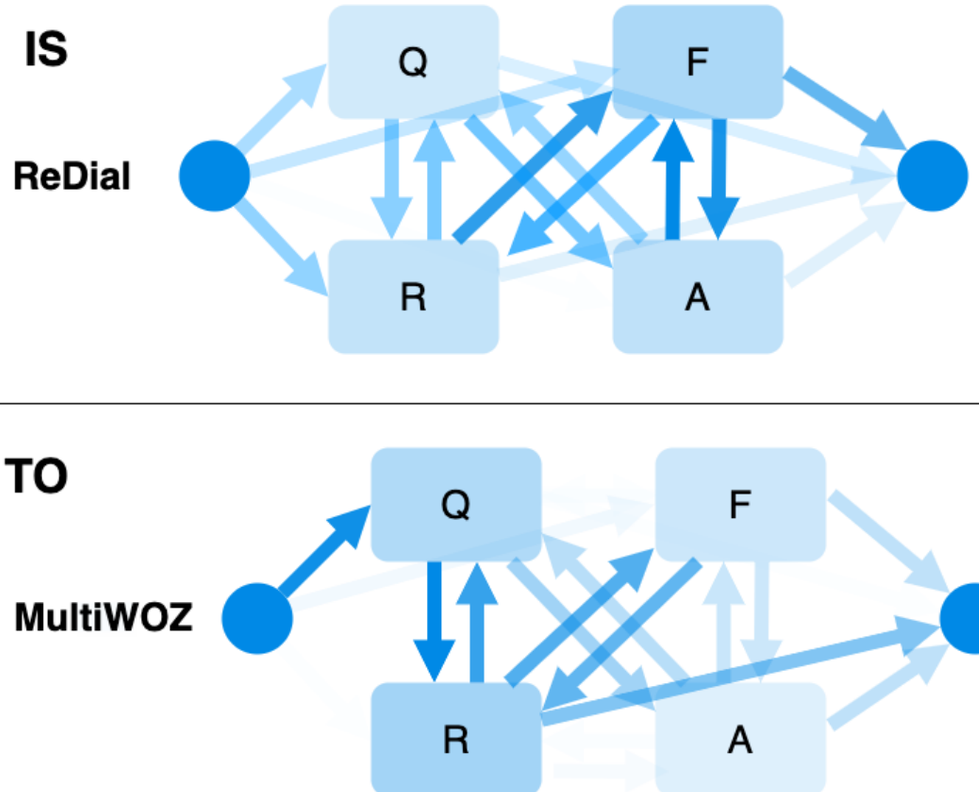






Support Dialogues QA<RF





Asymmetry Metrics

- **Volume** – who talks more in a dialogue?
- **Direction** – who requests information in a dialogue?
- **Information** – who contributes to the dialogue topic?
- **Repetition** – who follows up on the topic?

Svitlana Vakulenko, Evangelos Kanoulas, Maarten de Rijke: An Analysis of Mixed Initiative and Collaboration in Information-Seeking Dialogues. SIGIR 2020: 2085-20

Dialogue Fingerprint

Fingerprint					Utterance	Terms
Role	Type	Length	Repetitions			
A	H	4	0	0	Hey!	{hey}
A	I	41	1	0	What kind of movies do you like to watch?	{watch, **movi**}
S	N	42	0	0	I'm really big on indie romance and dramas	{romanc, drama, indi}
A	I	30	1	0	Ok what's your favorite movie?	{favorit, **movi**}
A	I	56	0	0	Staying with that genre, have you seen @88487 or @104253	{genr, stay, 88487, 104253}
A	N	30	0	0	Those are two really good ones	{}
S	N	44	0	1	When I was a kid I liked horror like @181097	{181097, kid, **horror**}
A	N	41	0	0	@Misery is really creepy but really good.	{miseri, creepi}
A	N	32	0	1	I only recently got into horror.	{**horror**}

Asymmetry Metrics

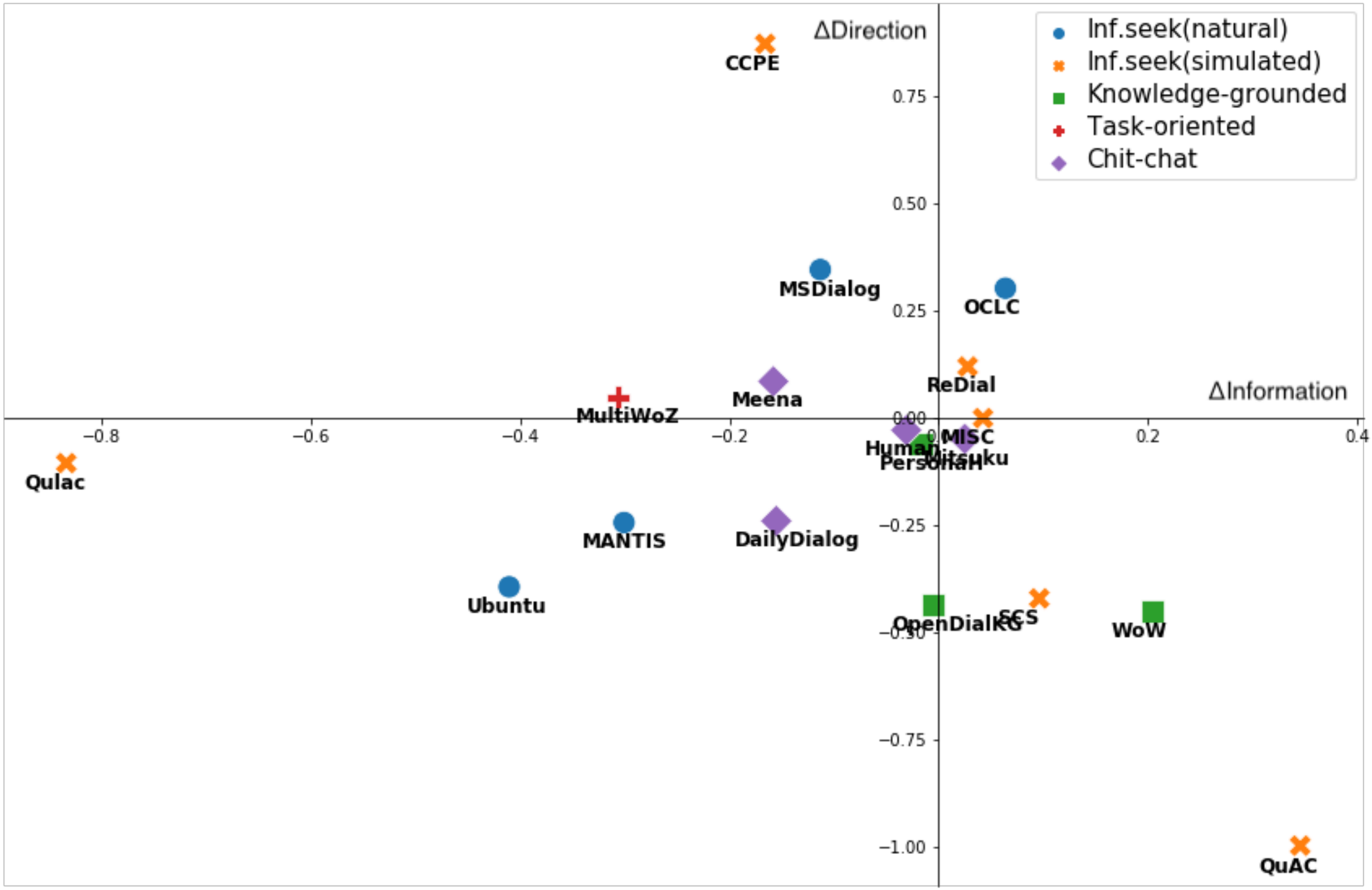
$$Volume_{ir} = \frac{1}{n_i} \sum_{j=1}^{n_i} l_{ij} [r_{ij} = r].$$

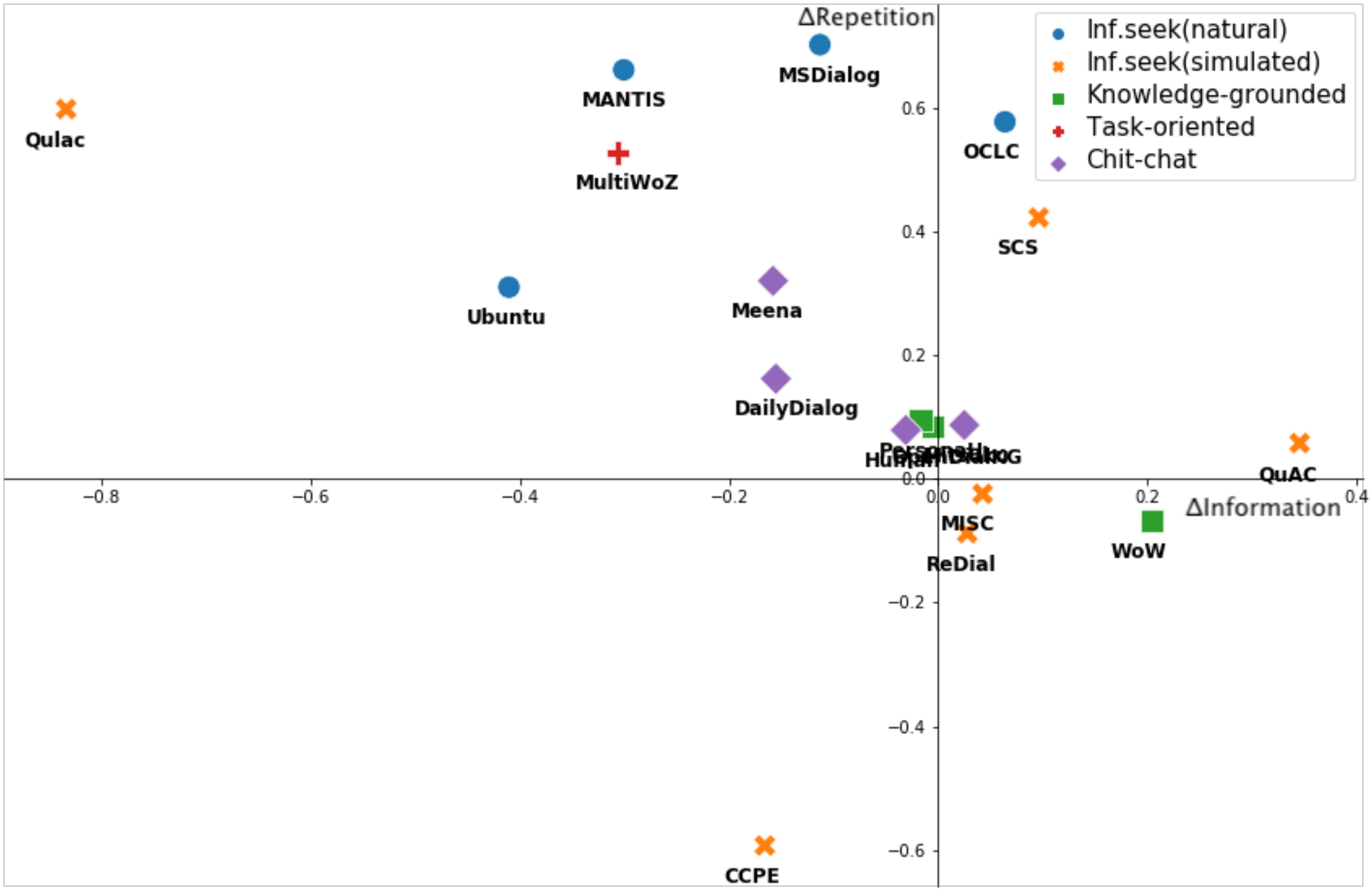
$$Direction_{ir} = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathcal{I}(t_{ij} = I)[r_{ij} = r].$$

$$Information_{ir} = \frac{1}{n_i} \sum_{j=1}^{n_i} \sum_{k=1}^{m_i} \mathcal{I}(v_{ijk} = 1)[r_{ij} = r] \times \mathcal{I}\left(\sum_{g=1}^{j-1} v_{igk} = 0 \right).$$

$$Repetition_{ir} = \frac{1}{n_i} \sum_{j=1}^{n_i} \sum_{k=1}^{m_i} \mathcal{I}(v_{ijk} = 1)[r_{ij} = r] \times \mathcal{I}\left(\sum_{g=1}^{j-1} v_{igk}[r_{ij} \neq r] > 0 \right).$$

$$\Delta Metric = \frac{1}{d} \sum_{i=1}^{d} \frac{Metric_{iA} - Metric_{iS}}{Metric_{iA} + Metric_{iS}}$$


ΔDirection
ΔInformation
Inf.seek(natural)
Inf.seek(simulated)
Knowledge-grounded
Task-oriented
Chit-chat
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−0.8
−0.6
−0.4
−0.2
0.2
0.4
CCPE
MSDialog
OCLC
ReDial
MISC
MultiWoZ
Meena
Human
Mitsuku
PersonaH
Qulac
MANTIS
DailyDialog
Ubuntu
SCS
OpenDialKG
WoW
QuAC


ΔRepetition
ΔInformation
Inf.seek(natural)
Inf.seek(simulated)
Knowledge-grounded
Task-oriented
Chit-chat
Qulac
MANTIS
MSDialog
MultiWoZ
OCLC
SCS
Ubuntu
Meena
DailyDialog
PersonaChat
HumanConvAI2
QuAC
MISC
ReDial
WoW
CCPE
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4

Conclusion

- None of the dialogue datasets mirrors the patterns of mixed initiative in virtual reference interviews from the OCLC dataset.

- **Professional librarians take more initiative in a conversation.** Non-expert intermediaries write less and ask less questions than professional librarians -> We should design better guidelines for data collection.

- Existing datasets collected to inform the conversational search task (MISC, SCSdata, etc.) are not suitable for studying and designing mixed initiative systems. The community should focus on more realistic datasets, such as OCLC, **to better understand the patterns of initiative** from interactions between **skilled interviewers/librarians** and information seekers.