

# Narrative-driven Recommendation as Complex Task

Toine Bogers, Aalborg University Copenhagen, Copenhagen, Denmark  
 Marijn Koolen, Royal Netherlands Academy of Arts and Sciences, Amsterdam, Netherlands

## 1 INTRODUCTION

Narrative-Driven Recommendation (NDR, [1]) is a complex but common scenario:

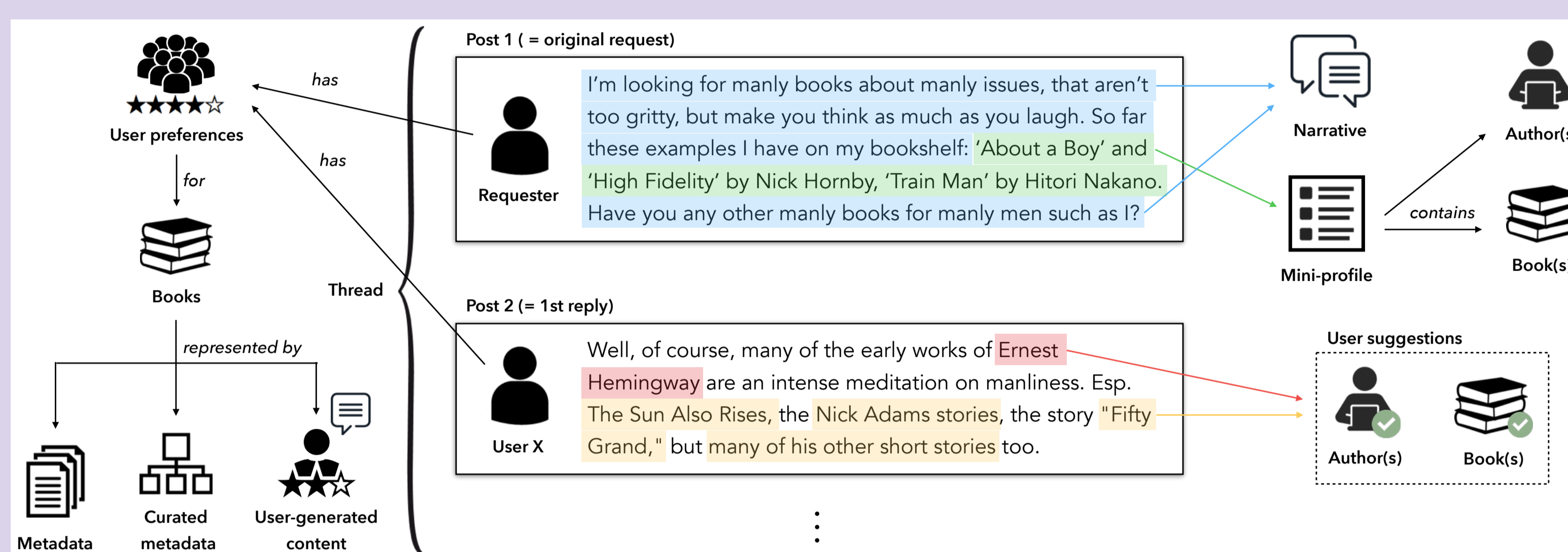
- (1) narrative description of the desired aspects of relevant items
- (2) user preference info., (user profile or example items)

This is related to but different from Conversational Recommendation.

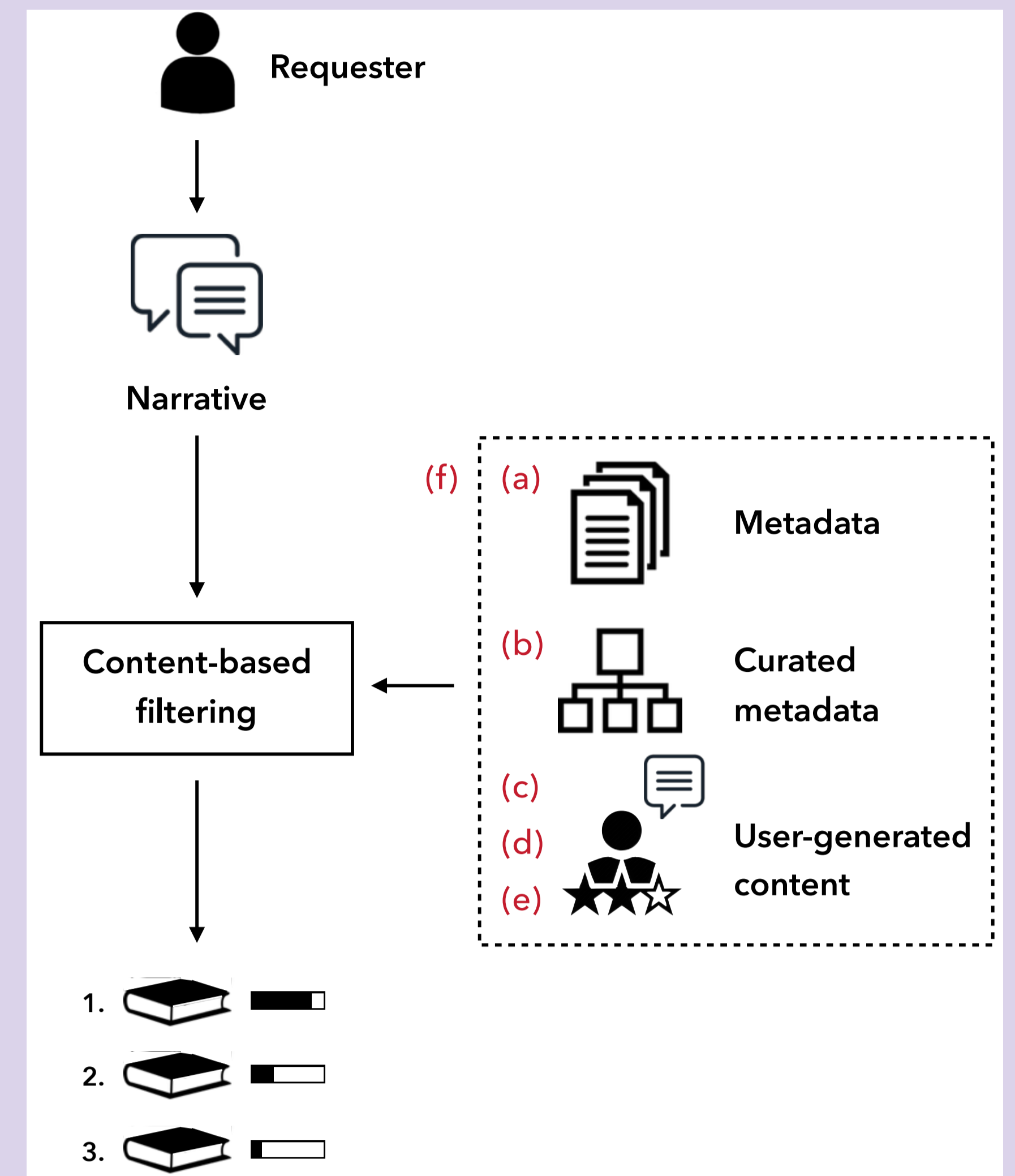
Test NDR algorithms that incorporate textual narratives and example books and authors.

Context: recommendation requests in LibraryThing (LT) discussion forums

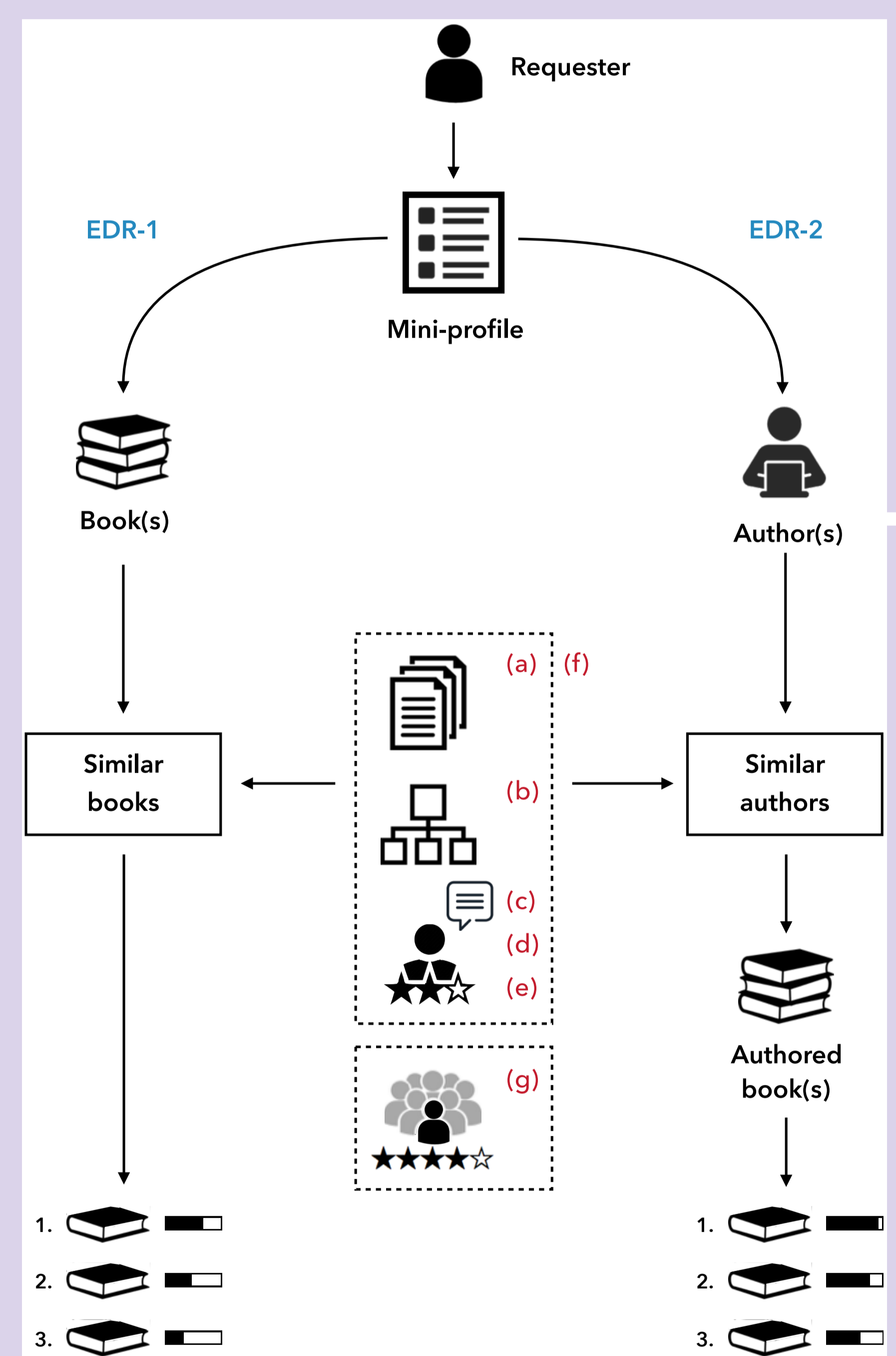
Similar types of needs in domains of films, games, music, television



## 2 SCENARIOS



Narrative-Driven Recommendation



Example-Driven Recommendation

## 3 EXPERIMENTAL SETUP

Datasets:

1. Amazon/Librarything book records: 2.8M (metadata + user-gen. content)
2. LibraryThing crawl: 66K user profiles, 29M transactions, 4.4M ratings

Requests:

- Starting point: 974 forum requests by [1], select requests from user in crawl
- 331 requests, 298 with example items

Baselines: CF: WARP [2], 50/50 train/test, CBR: Indri LM (stopping, JM smoothing)

Distribution of relevance aspects:

	A	C	E	F	M	N	S
Accessibility	137	9%	41	48	28	8	27
Content		598	157	267	176	26	98
Engagement			196	88	40	11	24
Familiarity				326	74	17	45
Metadata					179	11	25
Novelty						34	10
Socio-cultural							133

Content+metadata: **content-based**

Familiarity: **latent factors**

Engagement: **NLP on reviews**

## 4 EVALUATION

Collection	NDR	
	nDCG@10	MRR
(a) Metadata	0.024	0.059 <sup>Δ</sup>
(b) Curated metadata	0.026	0.052
(c) Tags	0.048 <sup>▲</sup>	0.111 <sup>▲</sup>
(d) Reviews	0.067 <sup>▲</sup>	0.164 <sup>▲</sup>
(e) Reviews + Tags	<b>0.074<sup>▲</sup></b>	<b>0.181<sup>▲</sup></b>
(f) All fields	<b>0.074<sup>▲</sup></b>	0.180 <sup>▲</sup>
(g) User Preferences	-	-
Collaborative filtering	0.017	0.031

Collection	Hybrid: NDR+EDR		
	$\lambda_{1,2,3}$	nDCG@10	MRR
Best single run: NDR (e)	-	0.074	0.181
EDR-1 (f) + EDR-1 (g)	.9/.1	0.051 <sup>▽</sup>	0.117 <sup>▽</sup>
EDR-1 (f) + EDR-2 (f)	.8/.2	0.047 <sup>▽</sup>	0.117 <sup>▽</sup>
NDR (e) + EDR-1 (f)	.9/.1	<b>0.088<sup>Δ</sup></b>	0.177
NDR (e) + EDR-1 (g)	.9/.1	<b>0.079</b>	0.193
NDR (e) + EDR-2 (f)	.9/.1	0.065 <sup>▽</sup>	0.162
NDR (e) + EDR-1 (f) + EDR-1 (g)	.8/.1/.1	0.081	0.200
NDR (e) + EDR-1 (f) + EDR-2 (f)	.8/.1/.1	0.082	0.170

Narrative-Driven Recommendation (NDR)

Hybrid: NDR+EDR

Collection	EDR-1						EDR-2					
	k	nDCG@10	MRR	nDCG@10	MRR		k	nDCG@10	MRR	nDCG@10	MRR	
(a) Metadata	200	0.009	0.027	0.010	0.029		100	0.003 <sup>▽</sup>	0.016	0.009	0.045	
(b) Curated metadata	250	0.018	0.035	0.020	0.039		15	0.004 <sup>▽</sup>	0.017	0.012	0.047	
(c) Tags	400	0.045 <sup>▲</sup>	0.091 <sup>▲</sup>	0.050 <sup>▲</sup>	0.100 <sup>▲</sup>		30	0.005 <sup>▽</sup>	0.017	0.014	0.049 <sup>Δ</sup>	
(d) Reviews	200	0.033 <sup>Δ</sup>	0.079 <sup>Δ</sup>	0.036	0.086 <sup>▲</sup>		15	0.006	0.016	0.016	0.045	
(e) Reviews + Tags	200	0.044 <sup>▲</sup>	0.107 <sup>▲</sup>	0.048 <sup>▲</sup>	0.117 <sup>▲</sup>		45	0.006 <sup>▽</sup>	0.017	0.017 <sup>Δ</sup>	0.050 <sup>Δ</sup>	
(f) All fields	800	<b>0.046<sup>▲</sup></b>	<b>0.113<sup>▲</sup></b>	<b>0.051<sup>▲</sup></b>	<b>0.123<sup>▲</sup></b>		15	0.006	0.022	<b>0.018</b>	<b>0.062<sup>Δ</sup></b>	
(g) User preferences	10	0.040 <sup>▲</sup>	0.101 <sup>▲</sup>	0.047 <sup>Δ</sup>	0.108 <sup>▲</sup>		-	-	-	-	-	
Collaborative filtering	-	0.017	0.031	0.019	0.035		-	<b>0.017</b>	<b>0.031</b>	0.007	0.020	

Example-Driven Recommendation (EDR)

- NDR: CF not effective by itself (focus on popular), CBR better
- User-generated content much more effective than metadata
- EDR: reviews effective for matching against example items
- Hybrid: improves NDR and EDR baselines, narrative and examples complementary

## 5 CONCLUSIONS

NDR is a challenging, complex task:

Requires combination of data sources and algorithms to solve.

High variation across users, personalization is key

Future work:

Use NLP to extract structured knowledge from narrative

look at conversational and graph-based models

Experiment with Interface Design (eliciting needs)

[1] Bogers & Koolen (2017). Defining and Supporting Narrative-driven Recommendation. ACM RecSys 2017, Como, Italy.

[2] Jason Weston, Samy Bengio, and Nicolas Usunier. 2011. Wsabie: Scaling up to large vocabulary image annotation. In IJCAI, Vol. 11. 2764–2770.