VIETNAM NATIONAL UNIVERSITY HO CHI MINH CITY
HO CHI MINH CITY UNIVERSITY OF TECHNOLOGY

May 2026

HK252-DATN-070

# Diversity Approach to Knowledge-Graph Based Recommendation Systems

**SUPERVISORS**

Assoc. Prof. Dr. Thoai Nam
Dr. Diep Thanh Dang

**STUDENTS**

Nguyen Ngoc Song Thuong (2252803)
Nguyen Huu Thanh (2552742)

**REVIEWER**

MEng. Nguyen Tuan Khoi

# Table of content


01 **Introduction & Motivation**

02 **Related work**

03 **SRD: Soft-Rank ILD Loss**
*Differentiable diversity objective*

04 **DivKG: Reinforcement Learning Finetune for Diversity**
*Policy optimization with diversity reward*

05 **Conclusion & Future Work**

01

# Introduction & Motivation

# Conversational Recommendation

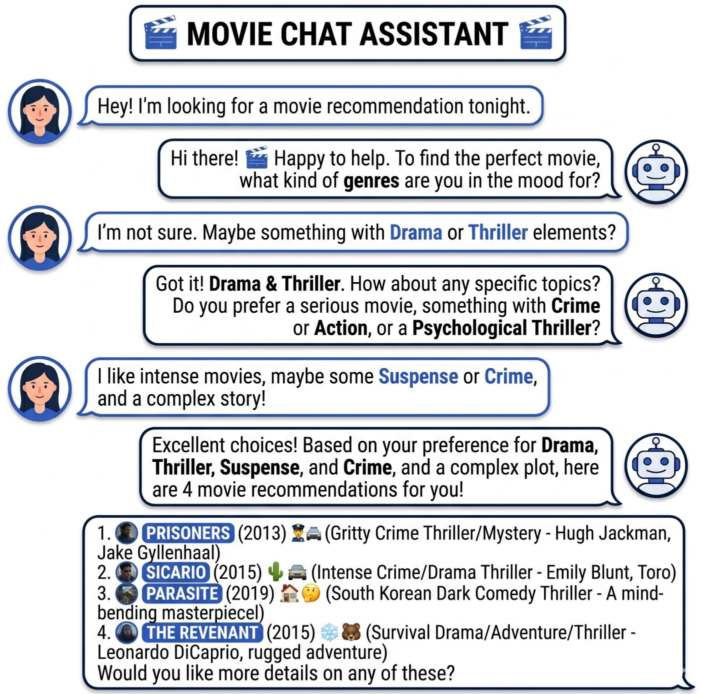



*Image 1.1. Illustration the conversational recommendation system in movie recommendation domain (src: AI-generated)*

## Conversational Recommender Systems (CRS)

suggest relevant items through iterative natural language interactions.

**Interactive Discovery** Explores user preferences through back-and-forth turns.

**Continuous Refinement** Refines recommendations at each step based on immediate feedback.

# Popularity bias: The same movies, every time

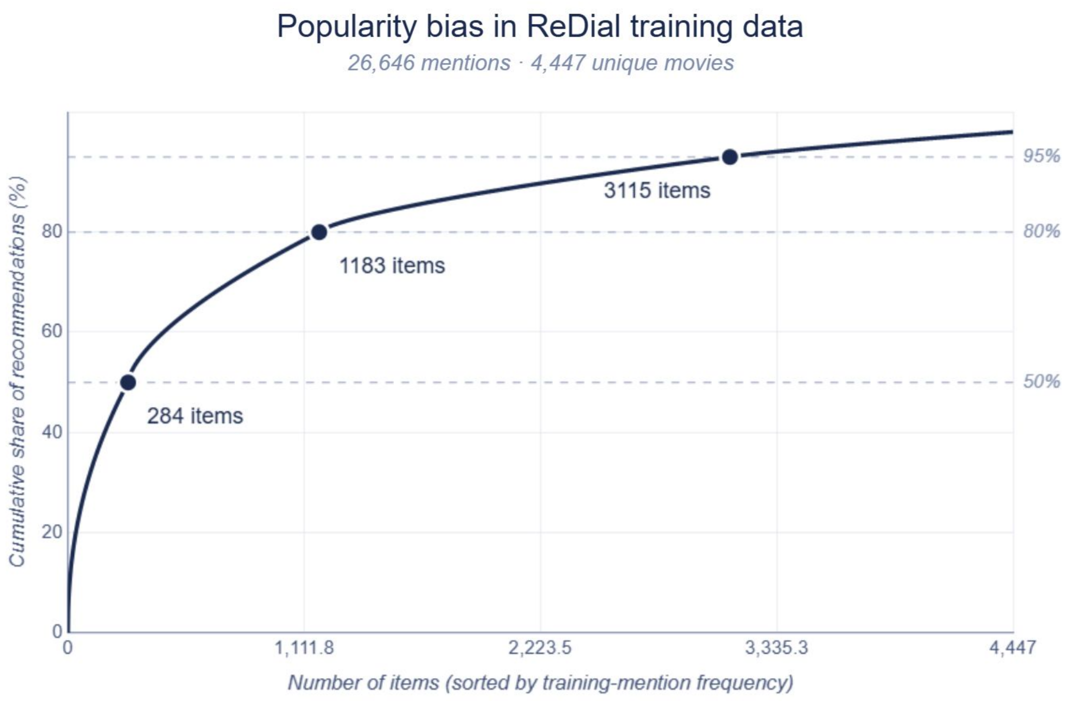



*Figure 1.1. Popularity bias in ReDial training data.*

**4%** of catalog cover **50%** of mentions

**17%** of catalog cover **80%** of mentions

**36%** of catalog never appear in training

# Knowledge Graph integration to CRS

**Knowledge Graph integration (KG-CRS)** solves the fundamental limitations of knowledge connection of entities.

**Entity Grounding** Links unseen entities via shared attributes & semantic relationships

**Multi-hop Reasoning** Traverses knowledge paths to discover relevant candidates.

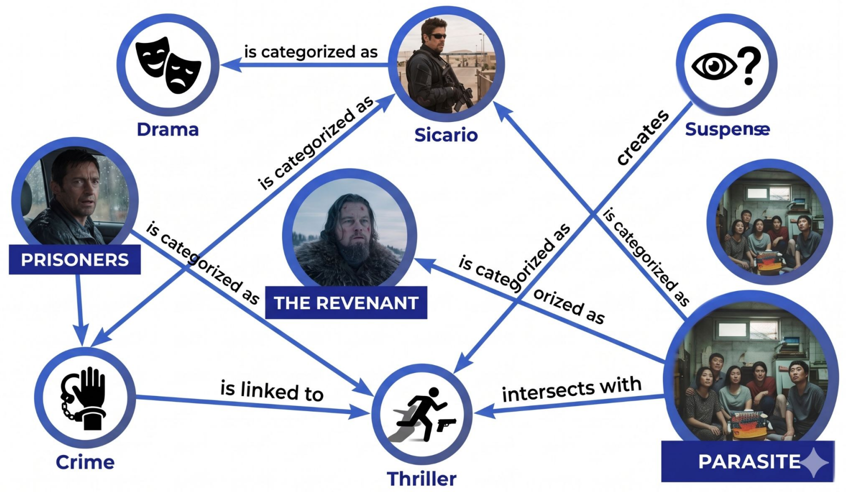



*Image 1.2. Illustration the knowledge graph and show the connection between entities by relationships (src: AI-generated)*



*Image ref: Gemini Nano Banana

02

# Related Work

# Related work: Positioning and the diversity gap

**KBRD** (EMNLP'19)
***Foundational benchmark***
DBpedia + RGCN; the *entry point* of KG-CRS; full-catalog scorer.

**KGSF** (KDD'20)
***Foundational benchmark***
Add ConceptNet word-side KG with gated fusion; *standard* accuracy reference.

**CR-Walker** (EMNLP'21)
***SOTA coverage***
Attribute-restricted hop-2 scoring; *the only baseline that reports coverage*

**TREA** (ACL'23)
***SOTA accuracy***
Tree-structured path-attention; *highest recall* in the survey on both benchmark datasets.

- ✖ None of the four trained against a diversity signal
- ✖ Current Diversity Approaches: Mostly rely on post-hoc reranking methods, while in-training regularizers (such as Determinantal Point Process - DPP), are not config on KG-CRS models yet.

# Benchmark Datasets

Table 2.1. Summary of two benchmark datasets (ReDial and TG-ReDial)

| Datasets | Dialogues | Movie Items | Language |
|---|---|---|---|
| ReDial | 10,006 | 6,924 | English |
| TG-ReDial | 10,000 | 33,834 | Chinese |

# Benchmark Datasets

Table 2.2. Summary of 4 Knowledge Graphs utilized

| Subgraph | Relation Types | Entities | Focus | Used for |
|---|---|---|---|---|
| DBpedia | cast, director, producer, writer, distributor, country | ~30K | Entity-level film facts | ReDial |
| ConceptNet | IsA, HasProperty, RelatedTo, UsedFor | ~29K | Word-level commonsense | ReDial |
| CN-DBpedia | cast, director, country, year, topic links | ~62K | Chinese entity facts + topic-thread connectivity | TG-ReDial |
| HowNet | No types | ~65K | Chinese lexical items | TG-ReDial |

# Benchmark Metrics

## Recall@K

*Measuring accuracy*

$$R@K = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} hit_u(K)$$

*Is the gold item in the top-K list?*

## Coverage@K

*Measuring diversity*

$$Cov@K = \frac{\left|\bigcup_{u \in \mathcal{U}} \mathcal{L}_u\right|}{|\mathcal{I}|}$$

*What fraction of all items get recommended across all users?*

## Harmonic mean

*Measuring trade-off*

$$F1@K = \frac{2 \times R@K \times Cov@K}{R@K + Cov@K}$$

# Baseline comparison

## Accuracy versus diversity, by dataset

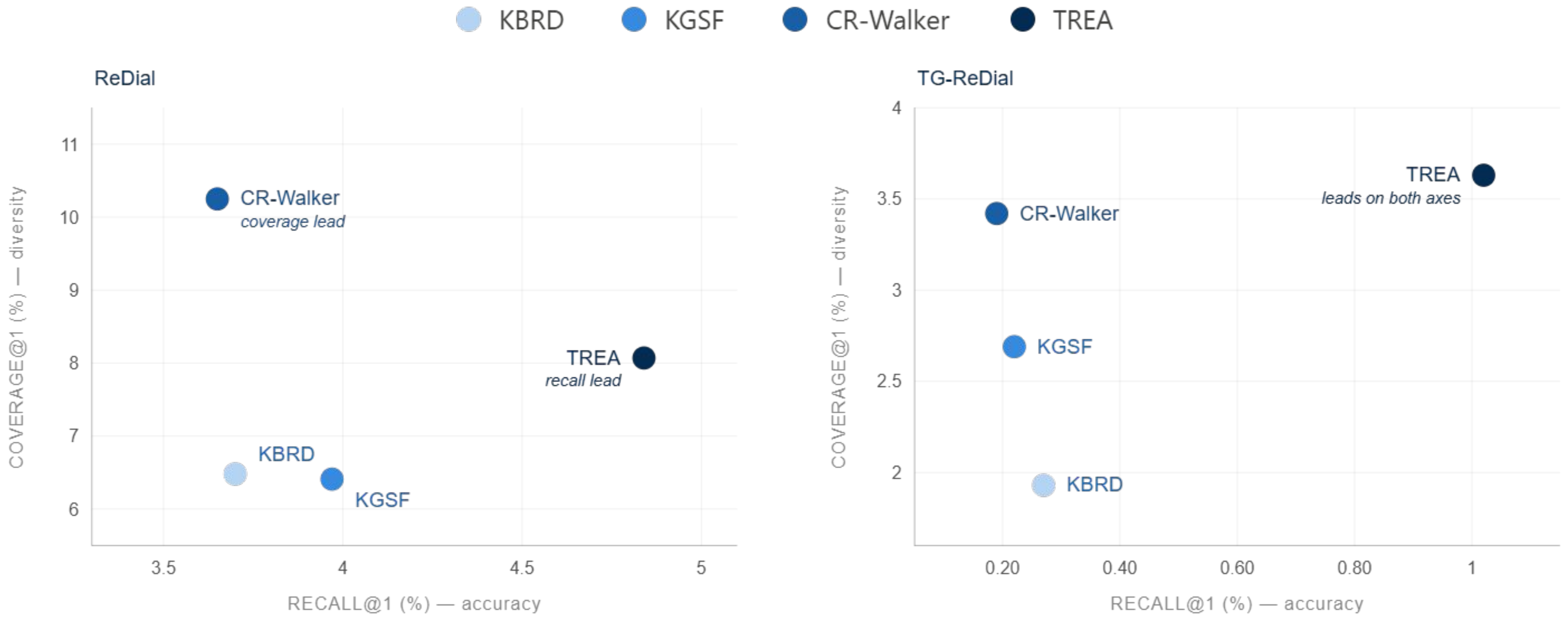



*Figure 2.1. Illustration of the results in recall and coverage of SoTAs on 2 benchmarking datasets*

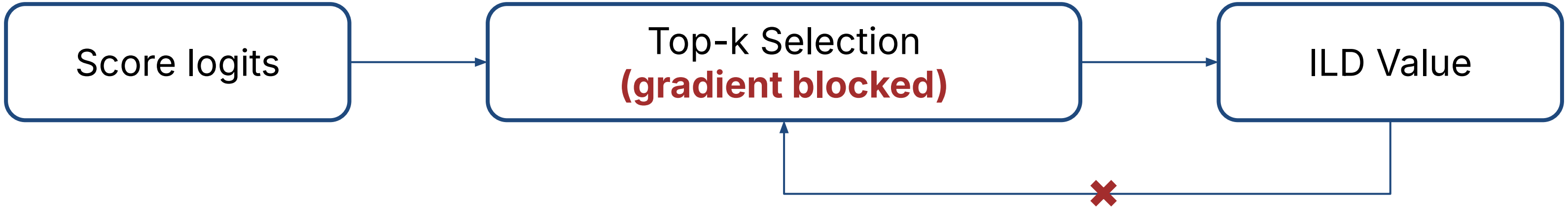
03

First approach • SRD

# Soft-Rank ILD Loss

# Problem statement

- ILD requires selecting top-K items first.
- The top-K operator is piecewise constant → gradient is zero almost everywhere. $\frac{\partial \mathrm{ILD}(S)}{\partial s_i} = 0$




**No gradient reaches the model → diversity cannot be learned**

# Replace hard selection with a soft selection

**Hard top-K**

$$w_i = \begin{cases} 1 \text{ if } i \in \mathcal{S} \\ 0 \text{ otherwise} \end{cases}$$

**Soft selection**

$$w_{b,i} = \frac{\exp(s_{b,i}/\tau)}{\sum_j \exp(s_{b,j}/\tau)}, \quad \tau > 0$$

*Temperature-scaled softmax keeps the gradient path open*

**Role of $\tau$**

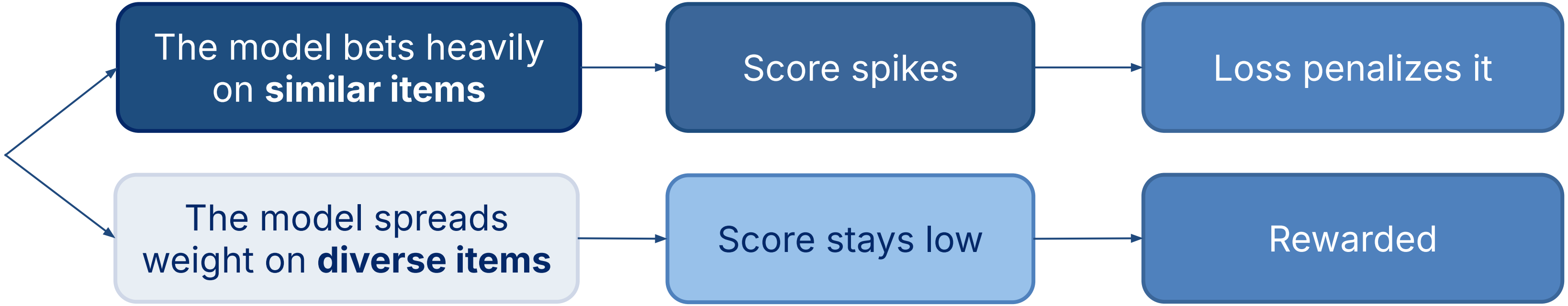
- $\tau \to 0$ : collapses to the top-1 indicator (recovers hard selection).
- $\tau \to \infty$: keeps mass on lower-ranked items, so the diversity gradient has signal to act on.
- For SRD, $\tau$ used deterministically as a differentiable relaxation.

# Measure if recommendations are too similar

We need a single number that gets large when the model recommends similar things, and small when it recommends diverse ones.

Similarity score across all recommended items:

$$\ell_b^{\mathrm{div}} = \mathbf{w}_b^\top S \mathbf{w}_b, \qquad S = \hat{H}\hat{H}^\top$$



# CR-Walker: Attribute-restricted Tree Reasoning

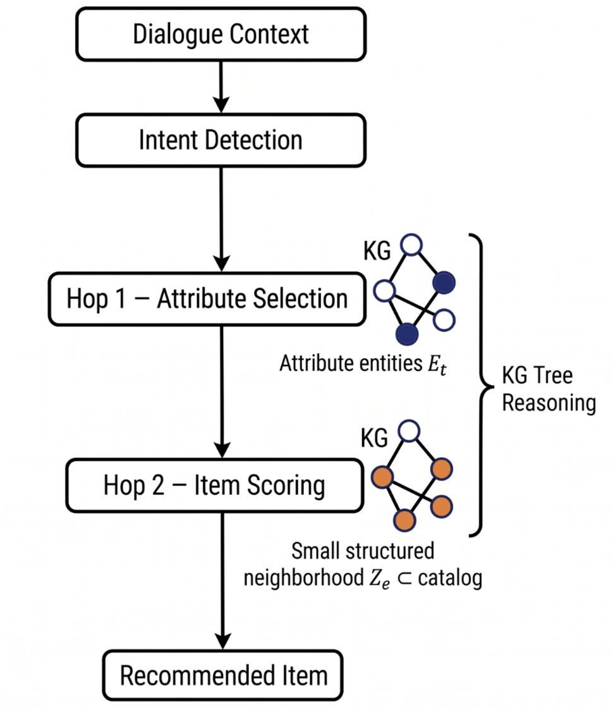



*Figure 3.1. Architecture of CR-Walker*

Candidate set at Hop 2 is **small and structured**
→ Pairwise similarity adds meaningful redundancy signal

# CR-Walker: Architecture after applied SRD

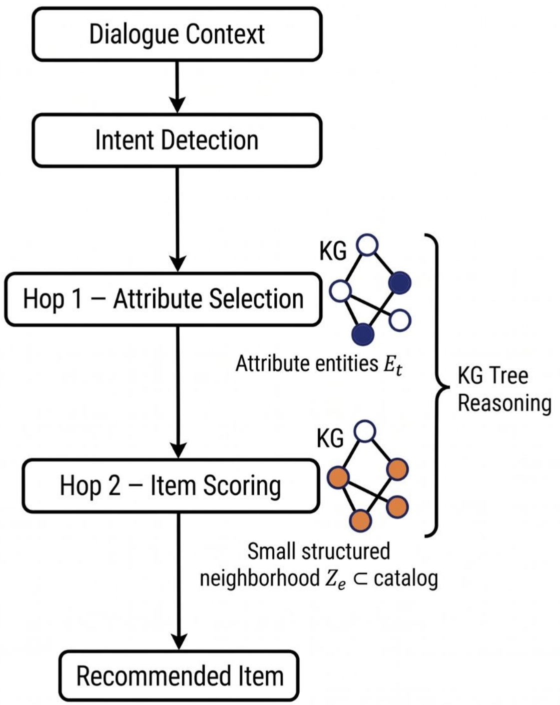



*Figure 3.1. Architecture of CR-Walker*

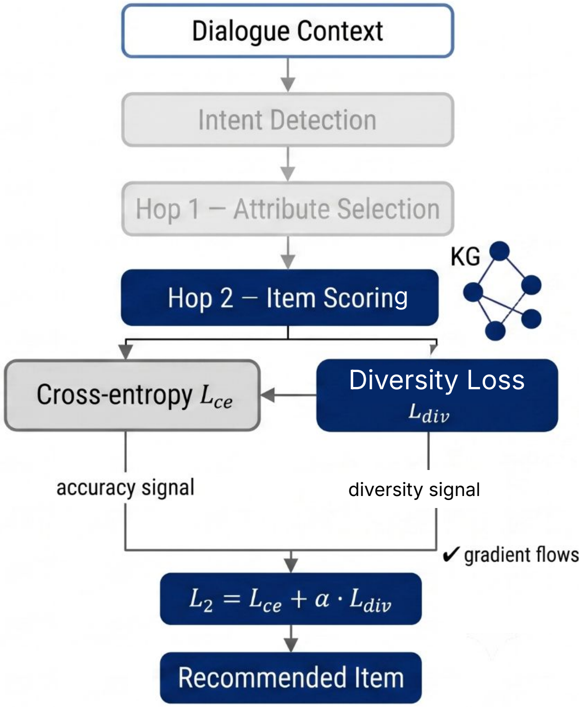



*Figure 3.2. Architecture of CR-Walker after applying SRD*

# Main Results of SRD on CR-Walker backbone

Table 3.1. Results of CR-Walker on dataset ReDial (percentages)

| Method | R@1 | R@10 | R@50 | Cov@1 | Cov@10 | Cov@50 | F1@10 |
|---|---|---|---|---|---|---|---|
| CR-Walker | **3.65** | 15.37 | 33.33 | 10.25 | 31.76 | 62.19 | 20.72 |
| +DPP | 3.19 | 15.57 | **34.33** | 9.60 | 28.96 | 57.12 | 20.25 |
| +SRD (ours) | 3.47 | **16.17** (+5.2%) | 32.44 (-2.6%) | **11.70** | **34.92** (+10%) | **65.38** (+5.1%) | **22.10** (+6.7%) |

Table 3.2. Results of CR-Walker on dataset TG-ReDial (percentages)

| Method | R@1 | R@10 | R@50 | Cov@1 | Cov@10 | Cov@50 | F1@10 |
|---|---|---|---|---|---|---|---|
| CR-Walker | **0.19** | 1.64 | 4.09 | **3.42** | 5.34 | 12.47 | 2.50 |
| +DPP | 0.18 | 1.76 | 3.08 | 3.28 | 5.14 | 12.24 | 1.77 |
| +SRD (ours) | 0.13 | **1.70** (+3.7%) | **4.21** (+2.9%) | 3.25 | **8.98** (+68.2%) | **20.65** (+65.6%) | **2.86** (+14.4%) |

# Limitation of SRD

Table 3.3. Results of KBRD and KGSF on dataset ReDial (percentages)

| Model | R@1 | R@10 | R@50 | Cov@10 | Cov@50 | F1@10 |
|---|---|---|---|---|---|---|
| KBRD | **3.70** | **18.79** | **35.87** | 18.37 | 36.37 | 18.58 |
| **+SRD** | 3.02 | 16.68 (-11.2%) | 35.10 (-2.1%) | **29.10** (+58.4%) | **50.51** (+38.9%) | **21.27** (+14.5%) |
| KGSF | **3.97** | **19.36** | **38.37** | 17.03 | 29.92 | 18.12 |
| **+SRD** | 3.50 | 18.79 (-2.9%) | 37.77 (-1.6%) | **23.83** (+40%) | **39.90** (+33.3%) | **21.01** (+15.9%) |

# Limitation of SRD

Table 3.4. Results of KBRD and KGSF on dataset TG-ReDial (percentages)

| Model | R@1 | R@10 | R@50 | Cov@10 | Cov@50 | F1@10 |
|---|---|---|---|---|---|---|
| KBRD | **0.22** | **1.74** | **6.77** | 14.20 | **29.08** | **3.1** |
| **+SRD** | 0.18 | 1.01 (-42%) | 3.30 (-51.3%) | **14.80** (+4.2%) | 26.81 (-7.8%) | 1.89 (39%) |
| KGSF | **0.27** | **1.20** | **5.12** | 9.58 | 20.58 | **2.12** |
| **+SRD** | 0.04 | 0.94 (-21.7%) | 3.52 (-31.3%) | **13.78** (+43.8%) | **28.56** (+38.8%) | 1.76 (-17%) |

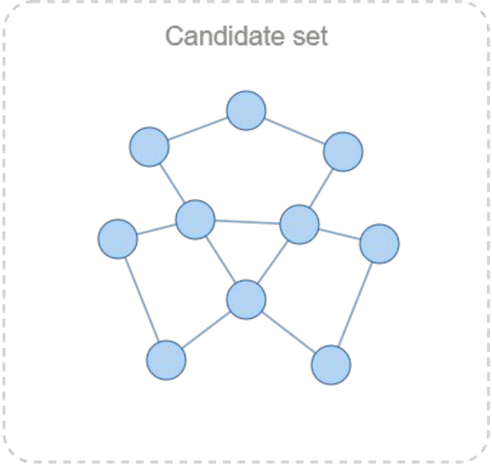
**Mechanism: Why SRD breaks on full-catalog backbones**

- SRD assumes candidate set is small & structurally related
- Full-catalog scorers (~6K ReDial, ~34K TG-ReDial): pairs are mostly unrelated → similarity signal is noisy
- Result: recall degrades, coverage gains shrink
- CR-Walker's neighbor traversal also misses edge-case items outside the KG neighborhood

# Where SRD fall short

### CR-Walker hop-2

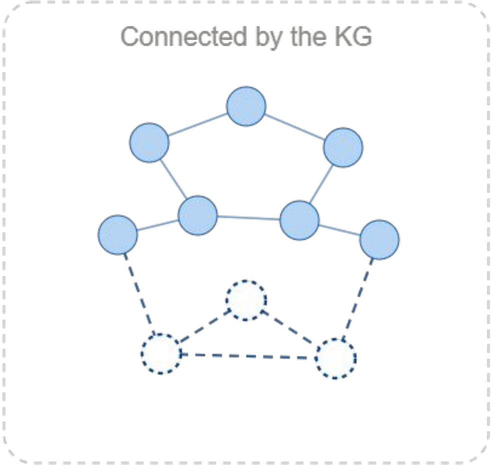
~20 KG-linked items




**Pairs are mostly similar**
in CR-Walker neighborhood

### What the KG misses

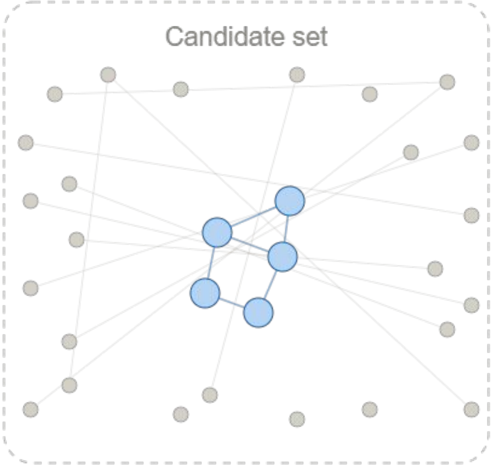
No edge for genre, mood, era




**The KG is incomplete**
Constraint to neighborhood prevent item discovery

### KBRD / KGSF full catalog

~6,924 items, mostly unrelated



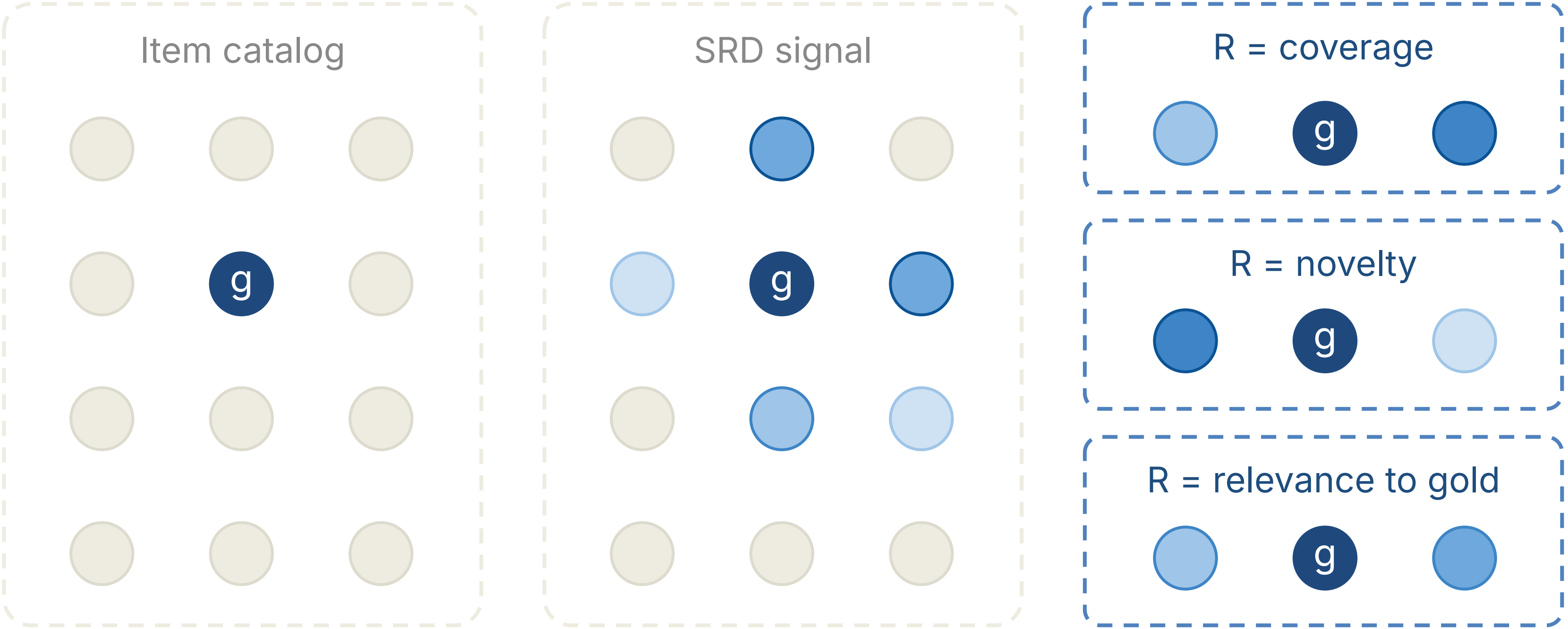

**Pairs are mostly unrelated**
in full catalog scorers

04

Second approach • DivKG

# Reinforcement Learning Finetune for Diversity

# Why reinforcement learning




**Supervised cross-entropy**
Only gold item carry signal

**Supervised SRD**
Add embedding similarity signal

**Reinforcement learning**
Items scored by any reward we plug in

# DivKG: Adapting the RLHF template to KG-CRS

Same skeleton: supervised pretrain → KL-anchored policy gradient

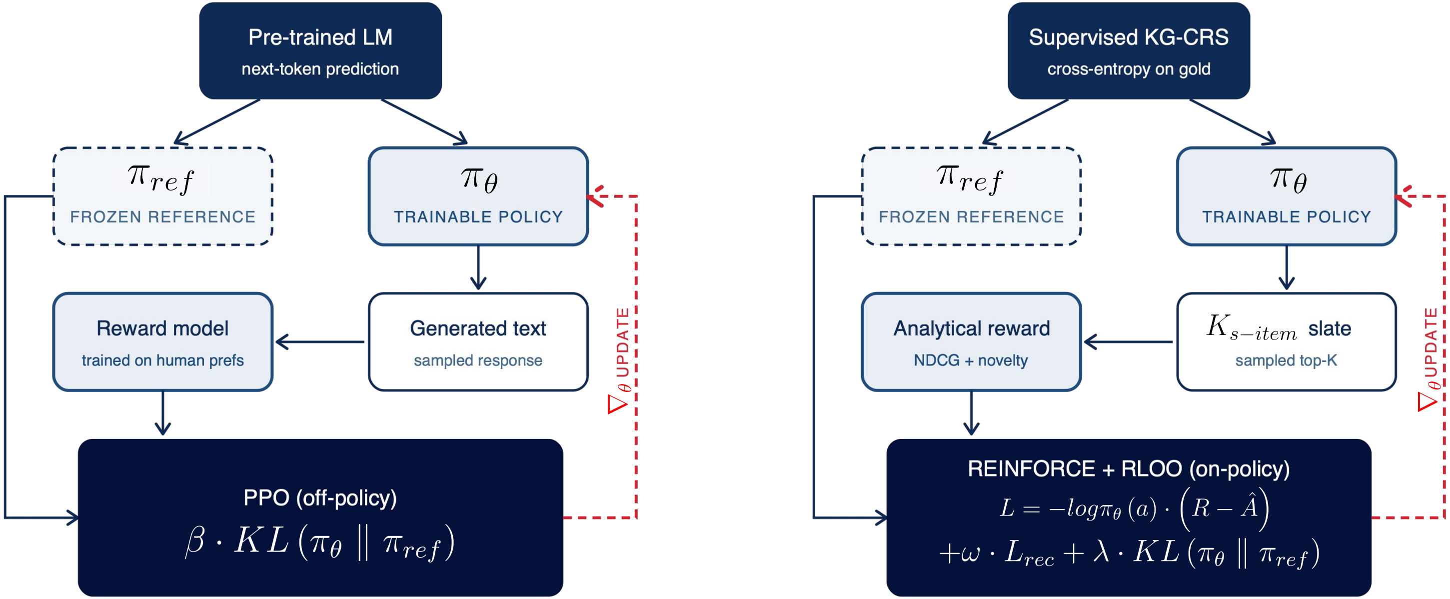



*Image 4.1. RLHF pipeline and Our DivKG pipeline in comparison.DivKG inherits RLHF's stability mechanism (KL anchor against a frozen reference) and discards the parts that don't apply to logged dialogue corpora: no value function, no horizon, no learned reward model.*

# DivKG pipeline

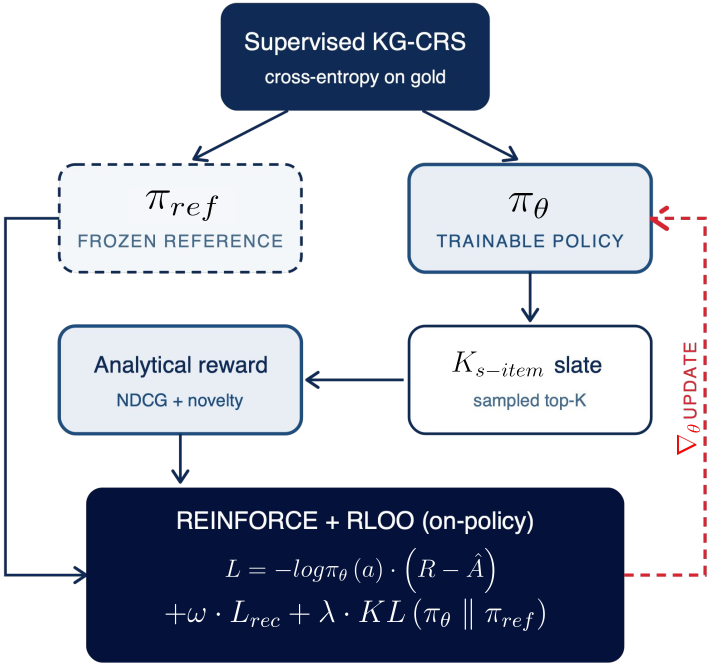



*Image 4.2. DivKG fine-tuning loop with frozen reference.*

**Policy**. Temperature-scaled softmax over logits scores

$$\pi = \frac{\exp(s/\tau)}{\sum_{i \in \mathcal{C}} \exp(s_i/\tau)}$$

**1.** Sampling G slate of K item

$$\mathbf{a} \sim \mathrm{Multinomial}_{K_s}(\pi_\theta)$$

**2.** For each slate, calculate the composite reward and the advantage with **RLOO**

$$\hat{A}^{(m)} = R^{(m)} - \frac{1}{G-1} \sum_{j \neq m} R^{(j)}$$

**3.** REINFORCE loss

$$\mathcal{L}_{\mathrm{RL}} = -\mathbb{E}\left[\hat{A} \log \pi_\theta(a \mid c)\right]$$

# Reward composition

$$R = \alpha_1 R_{ndcg} + \alpha_2 R_{nov} + \alpha_3 R_{rel-nov}$$

### NDCG

$$\frac{1}{log_2(r_{gold} + 2)}$$

*Rewards placing gold item near the top*

### Novelty

$$\left(1 - \frac{\text{pop}}{\text{pop}_{max}}\right) \cdot \lambda$$

*Rewards low-popularity items, DCG-discounted.*

### Relevance-Novelty

$$R_{nov} \cdot [\cos(e_i, e_g)]_+$$

*Reward items that are novel and semantically close to gold.*

# Fine-tune loss

$$L = \mathcal{L}_{rec} + \mathcal{L}_{\mathrm{RL}} + \lambda \mathrm{KL}(\pi_\theta || \pi_{ref})$$

supervised anchor to preserve accuracy

Allow exploration with diversity rewards

KL regularization to prevent drift from reference policy

# Main results - ReDial


*Figure 4.2. Illustration of the results in recall and coverage of SoTAs on Dataset ReDial*

# Main results - ReDial

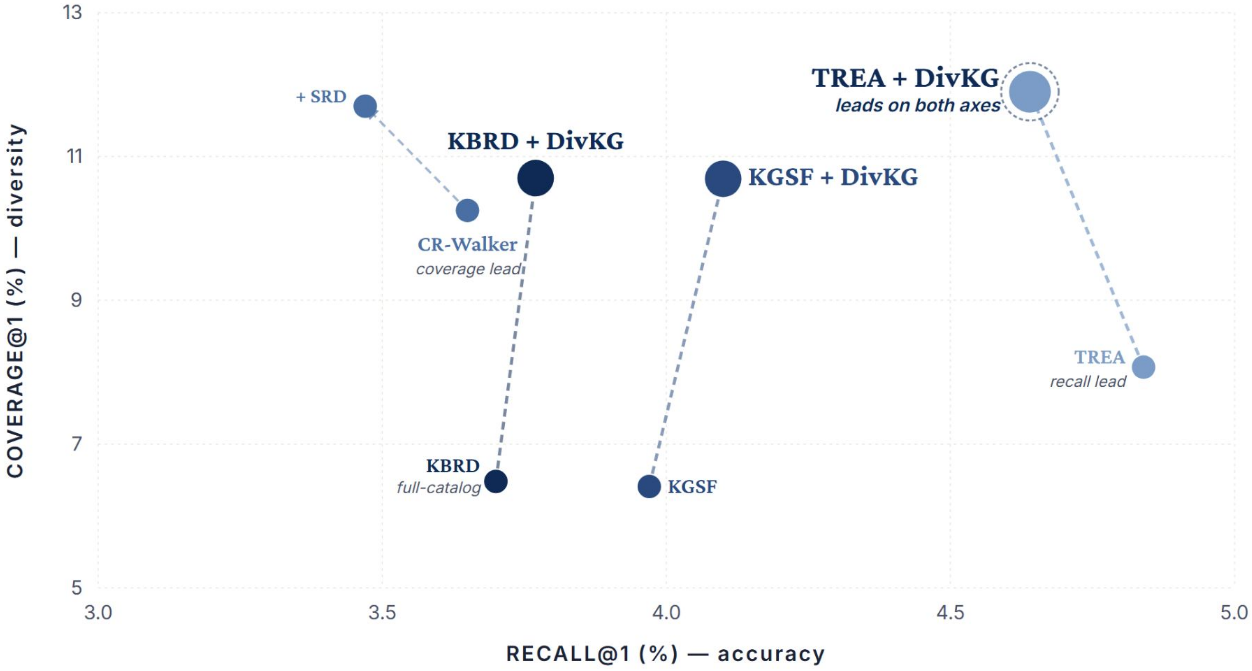



*Figure 4.3. Illustration of the results in recall and coverage of SoTAs when applying our techniques on Dataset ReDial*

# Main results - ReDial

Table 4.1. Results of TREA when applying DivKG compare to other baseline on Redial (percentages)
*Bold denotes best value, underline denotes second best value.*

| Model | R@1 | R@10 | R@50 | Cov@1 | Cov@10 | Cov@50 | F1@10 |
|---|---|---|---|---|---|---|---|
| CR-Walker | 3.65 | 15.37 | 33.33 | 10.25 | 31.76 | 62.19 | 20.72 |
| + SRD | 3.47 | 16.17 | 32.44 | <u>11.70</u> | <u>34.92</u> | <u>65.38</u> | 22.10 |
| TREA | **4.84** | **21.37** | **41.88** | 8.07 | 23.87 | 45.26 | 22.55 |
| **+ DivKG** | <u>4.64</u> | <u>20.10</u> | <u>39.45</u> | **11.90** | **41.75** | **71.16** | **27.14** |

**TREA + DivKG** wins Cov@1, Cov@10, Cov@50, F1@10 simultaneously.  beat CRWalker on every metrics and while still maintain high recall compare to TREA baseline.

# Main results - ReDial

Table 4.2. Results of other backbone when applying DivKG on dataset ReDial (percentages)

| Model | R@1 | R@10 | R@50 | Cov@1 | Cov@10 | Cov@50 | F1@10 |
|---|---|---|---|---|---|---|---|
| KBRD | 3.70 | 18.79 | 35.87 | 6.48 | 18.37 | 36.37 | 18.58 |
| **+DivKG** | 3.77 | 18.79 | 35.67 | **10.70** (+65%) | **32.74** (+78%) | **56.09** (+54%) | **23.87** |
| KGSF | 3.97 | 19.36 | 38.37 | 6.41 | 17.03 | 29.92 | 18.12 |
| **+DivKG** | 4.10 | 21.16 | 38.47 | **10.69** (+66%) | **30.88** (+81%) | **51.21** (+71%) | **25.11** |

DivKG help improves KBRD and KGSF **substantially on coverage**, while maintaining or even slightly increasing recall.

# Main results - TG-ReDial

Table 4.3. Results of DivKG on SoTA backbones and compare with SRD on TG-ReDial(percentages)

| Model | R@1 | R@10 | R@50 | Cov@1 | Cov@10 | Cov@50 | F1@10 |
|---|---|---|---|---|---|---|---|
| CR-Walker | 0.19 | 1.64 | 4.09 | 3.42 | 5.34 | 12.47 | 2.50 |
| +SRD | ↓ 0.13 | 1.70 | 4.21 | ↑ 3.25 | ↑ 8.98 | ↑ 20.65 | ↑ 2.86 |
| KBRD | 0.27 | 1.20 | 5.12 | 1.93 | 9.58 | 20.58 | 2.13 |
| +DivKG | ↓ 0.22 | 1.25 | 5.21 | ↑ 2.23 | ↑ 11.19 | ↑ 23.51 | ↑ 2.25 |
| KGSF | 0.22 | 1.74 | 6.77 | 2.69 | 14.20 | 29.08 | 3.10 |
| +DivKG | 0.40 | 1.74 | ↓ 6.11 | ↑ 3.19 | ↑ 17.15 | ↑ 34.00 | ↑ 3.16 |
| TREA | 1.02 | 3.21 | 9.05 | 3.63 | 16.40 | 28.92 | 5.37 |
| +DivKG | **1.07** | **4.01** | **9.98** | ↑ **4.11** | ↑ **19.32** | ↑ **31.50** | ↑ **6.64** |

DivKG improve coverages across all models with minimal recall drop. **TREA + DivKG** continues to dominate all metrics.

# Qualitative - KGSF + ReDial

**KGSF + ReDial · test index 3803.** Seed = *A Perfect Getaway, Orphan, The Purge*; concept tokens includes *thrillers, purge, family, girl*.
Gold ⭐ = *Children of the Corn* (training popularity **6**, long-tail).

## KGSF baseline

AvgPop 82.1

346


It (2017)

100


The Conjuring

24


It Follows

6


Haunting in Connecticut

34


The Conjuring 2

*Returns head-of-distribution horror. Gold buried at rank* ***9****; no seed surfaces in top-5.*

## KGSF + DivKG

AvgPop 53.2

4


Orphan

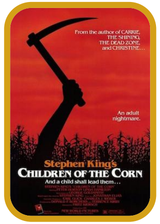
4


The Fourth Kind

20


Don't Breathe

6




Children of the Corn

34


The Conjuring 2

*Surfaces seed* ***Orphan*** *at rank 1 and* **recovers the gold *Children of the Corn*** *at rank 4.*



*Image ref: Google.com*

05

# Conclusion & Future work

# Conclusion & Future Work

CONTRIBUTION 1

## SOFT-RANK ILD LOSS

A differentiable diversity objective trained end-to-end alongside the accuracy loss.

| | |
|---|---|
| **ReDial** | advances the published coverage SOTA. |
| **TG-ReDial** | generalizes to prior works on enhancement of coverage but trade-off on accuracy. |

SUBMITTED

**SOMET 2026**

Manuscript under review

CONTRIBUTION 2

## RL FINETUNE FOR DIVERSITY

A framework that fine-tunes a pretrained KG-CRS toward diversity-aware policies.

| | |
|---|---|
| **ReDial** | overtakes coverage SOTA while keeping the recall lead. |
| **TG-ReDial** | improves SOTA on both recall and coverage as once. |

TARGET

**SIGIR / RecSys 2027**

Remaining work: additional CRS backbone

# Conclusion & Future Work

## FUTURE WORK - THE NEXT OPEN QUESTIONS

**User study.** Does catalog-level diversity translate to perceived recommendation quality in live conversation?

**Beyond CRS.** Methods generalises to any ranker? testing on sequential-rec and search benchmarks.

**Pareto reward design.** Replace the linear accuracy/diversity mix with explicit Pareto-front exploration.

*SRD source code*

*DivKG source code*

# References


[1] F. Ricci et al., Recommender Systems Handbook, Springer, 2011.

[2] M. Kaminskas & D. Bridge, Diversity, serendipity, novelty, and coverage: a survey, TiiS, 2017.

[3] H. Abdollahpouri et al., Controlling popularity bias in learning-to-rank recommendation, RecSys, 2017.

[4] Y.-J. Park & A. Tuzhilin, The long tail of recommender systems, RecSys, 2008.

[5] C. Gao et al., Advances and challenges in conversational recommender systems, AI Open, 2021.

**[6] Q. Chen et al., Towards knowledge-based recommender dialog system (KBRD), EMNLP, 2019.**

**[7] K. Zhou et al., Improving CRS via knowledge graph based semantic fusion (KGSF), KDD, 2020.**

**[8] W. Ma, R. Takanobu & M. Huang, CR-Walker: tree-structured graph reasoning and dialog acts, EMNLP, 2021.**

**[9] W. Li et al., TREA: tree-structure reasoning schema for conversational recommendation, ACL, 2023.**

[10] R. S. Sutton & A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed., MIT Press, 2018.

[11] D. M. Ziegler et al., Fine-tuning language models from human preferences, arXiv:1909.08593, 2019.

[12] R. Li et al., Towards deep conversational recommendations (ReDial), NeurIPS, 2018.

[13] K. Zhou et al., Towards topic-guided conversational recommender system (TG-ReDial), COLING, 2020.

[14] Q. Guo et al., A survey on knowledge graph-based recommender systems, IEEE TKDE, 2022.

[15] K. Papineni et al., BLEU: a method for automatic evaluation of machine translation, ACL, 2002.

[16] C.-Y. Lin, ROUGE: a package for automatic evaluation of summaries, ACL Workshop, 2004.

[17] J. Li et al., A diversity-promoting objective function for neural conversation models, NAACL-HLT, 2016.

[18] A. Hogan et al., Knowledge graphs, ACM Computing Surveys, 2021.

[19] M. Schlichtkrull et al., Modeling relational data with graph convolutional networks (R-GCN), ESWC, 2018.

[20] R. J. Williams, Simple statistical gradient-following algorithms for connectionist RL, Machine Learning, 1992.

[21] W. Kool, H. van Hoof & M. Welling, Buy 4 REINFORCE samples, get a baseline for free! ICLR Workshop, 2019.

[22] R. Speer, J. Chin & C. Havasi, ConceptNet 5.5: an open multilingual graph of general knowledge, AAAI, 2017.

[23] J. Carbonell & J. Goldstein, The use of MMR, diversity-based reranking, SIGIR, 1998.

[24] A. Kulesza & B. Taskar, Determinantal point processes for machine learning, Found. Trends ML, 2012.

[25] L. Chen, G. Zhang & E. Zhou, Fast greedy MAP inference for DPP, NeurIPS, 2018.

[26] H. Steck, Calibrated recommendations, RecSys, 2018.

[27] Y. Yue & C. Guestrin, Linear submodular bandits and their application to diversified retrieval, NeurIPS, 2011.

[28] D. Stamenkovic et al., Diverse and novel recommendations through multi-objective RL, WSDM, 2022.

[29] A. Ahmadian et al., Back to basics: revisiting REINFORCE-style optimization for LLMs, ACL, 2024.

[30] M. Chen et al., Top-K off-policy correction for a REINFORCE recommender system, WSDM, 2019.

# References

## Related work references

[6] Q. Chen et al., Towards knowledge-based recommender dialog system (KBRD), EMNLP, 2019.

[7] K. Zhou et al., Improving CRS via knowledge graph based semantic fusion (KGSF), KDD, 2020.

[8] W. Ma, R. Takanobu & M. Huang, CR-Walker: tree-structured graph reasoning and dialog acts, EMNLP, 2021.

[9] W. Li et al., TREA: tree-structure reasoning schema for conversational recommendation, ACL, 2023.

⭐⭐⭐

# Thank you for your attention!

*We extend our sincere gratitude to Assoc. Prof. Dr. Nam Thoai and Dr. Diep Thanh Dang for their invaluable guidance and direction throughout the development of this work.*

# Appendix: Experiments Hardware Setup

Table a.1. Details of Hardware running for our experiments

| Workstation | CPU | GPU | VRAM | RAM | Run |
|---|---|---|---|---|---|
| 1 | Intel Core i5-12600KF | NVIDIA RTX 3060 | 12 GB | 32 GB | SRD |
| 2 | AMD Ryzen 9 7900X | NVIDIA GeForce RTX 4070 Ti SUPER | 16 GB | 64 GB | DivKG |

# Appendix: Experiments Hyperparameters Setup

Table a.2. Details of Hyperparameters running for SRD

| Dataset | Embed Size | Learning Rate | Weight Decay | Epochs | Diversity Weight | Temperature |
|---|---|---|---|---|---|---|
| ReDial | 128 | $5 \times 10^{-4}$ | 0.01 | 1.23 | 0.1 | 0.1 |
| TG-ReDial | 128 | $5 \times 10^{-4}$ | 0.01 | 2.00 | 0.01 | 0.01 |

# Appendix: Experiments Hyperparameters Setup

Table a.3. Details of Hyperparameters running for DivKG

| Category | HP | Value | Context/Description |
| --- | --- | --- | --- |
| Temperature | $\tau$ | 3.0 | RLOO Baseline Setting |
| KL Weight | $\lambda$ | 0.01 | Kullback-Leibler Regularization |
| Output Size | $K_5$ | 50 | Slate Size per recommendation |
| Sampling | $G$ | 4 | Slates per context (RLOO) |
| Optimizer | LR / Adam | $10^{-4}$ / Clipping 0.1 | Gradient management with Adam |
| Anchoring | $\omega$ | {0.5, 1.0} | Supervised-anchor weight sweep |

# Appendix: Convergence speed of SRD compare to DPP

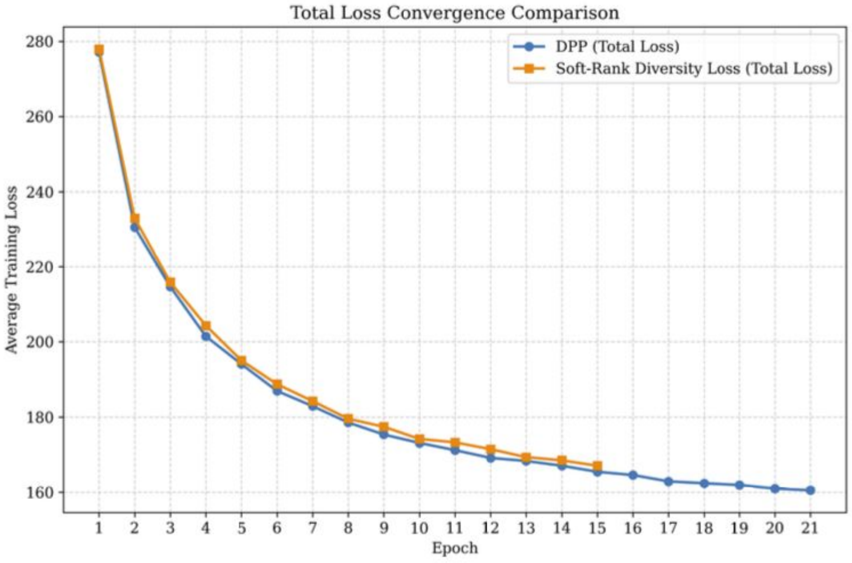



*Figure a.1. The figure shows convergence curves on ReDial. Training runs for 60 epochs with early stopping at patience 5.*

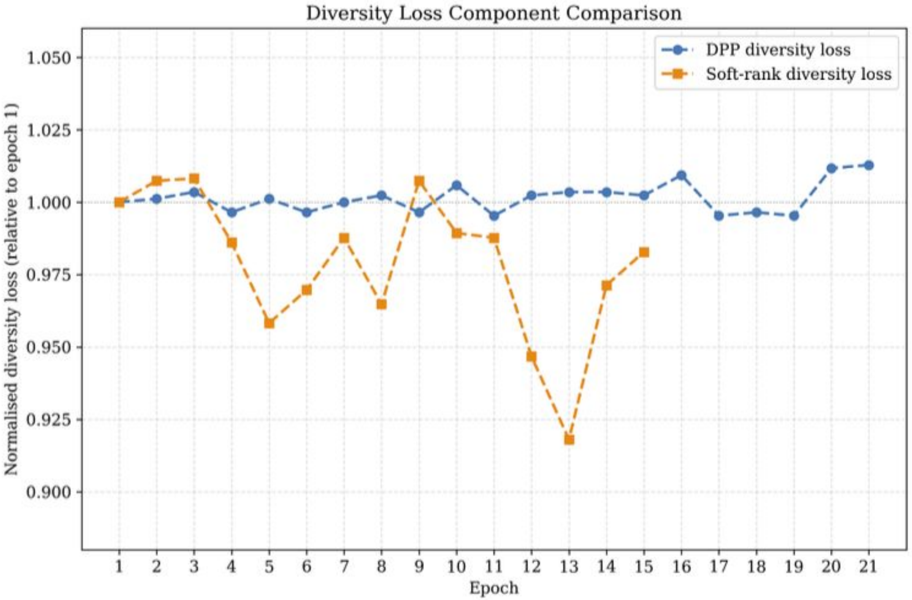
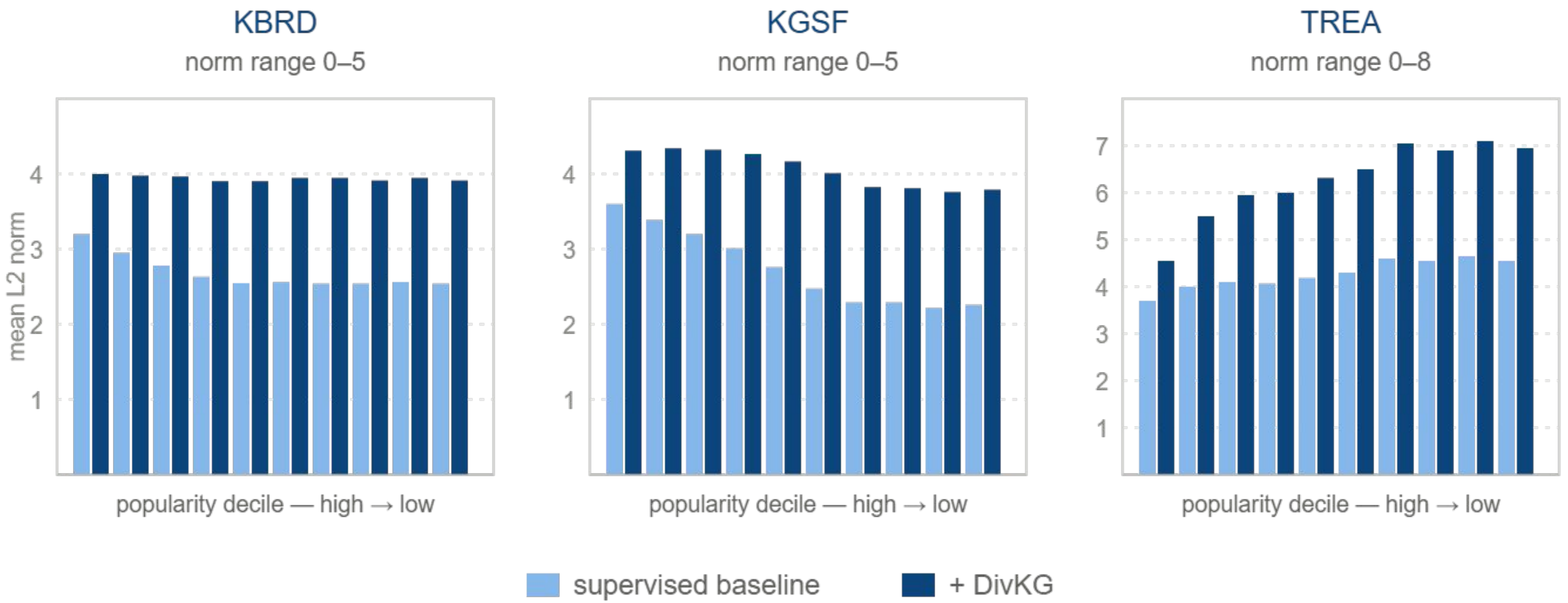



*Figure a.2. The figure shows the diversity series are normalized to their epoch-1 value, where the DPP diversity loss (blue) shows no downward trend across 21 epochs, while the SRD value (orange) decreases steadily*

# Average Embedding norms in different popularity deciles



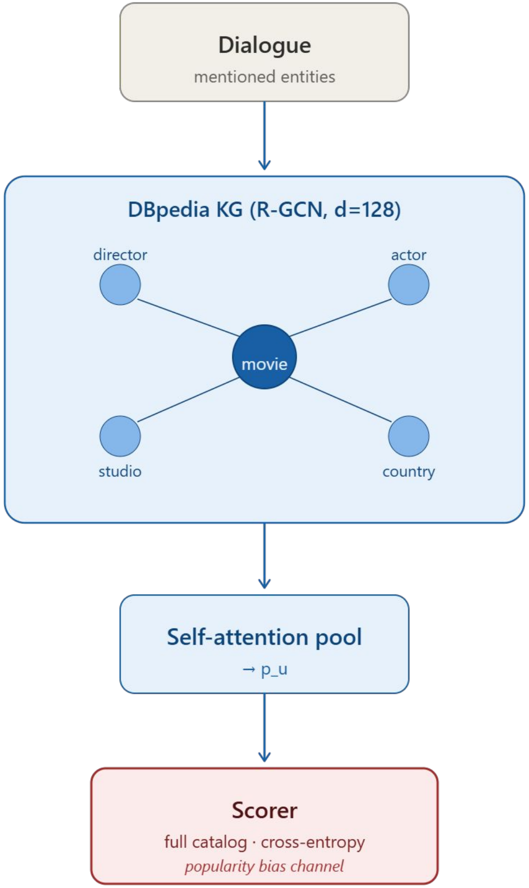
# On TREA, DivKG operates through the gate, not the embeddings

DivKG routes through **whichever architectural knob each backbone exposes**: embedding norms on KBRD/KGSF, the path/current gate on TREA.

| Component | Baseline | + DivKG | Change | Verdict |
|---|---|---|---|---|
| Mean entity-embedding norm | 4.60 | 6.71 | +46% | Norms move, but don't predict ranks (Spearman 0.14) |
| Across-path attention distribution | — | — | unchanged | Flat across all 24 paths |
| Across-path attention entropy | 2.75 bits | 2.69 bits | −2% | Noise |
| Path / current fusion gate | 0.72 | 0.66 | −8% (mean) | Shifts toward the less popularity-biased channel |

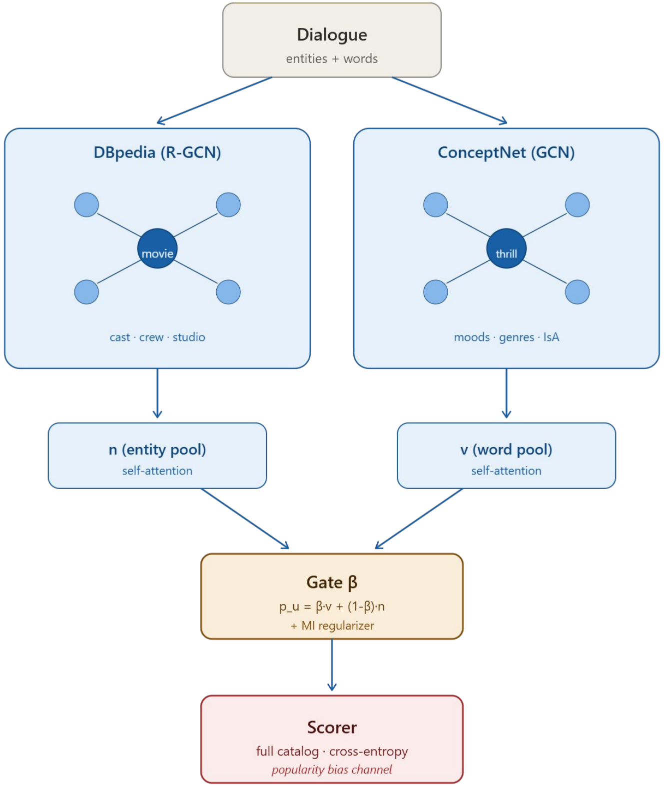
Path-history channel carries items with mean popularity **43.2**; current-turn channel mean popularity **40.2**. Paired Wilcoxon: path > current in 57.1% of contexts, $p \approx 4.4 \times 10^{-16}$. DivKG dials TREA toward the less popularity-biased of two channels.

## KBRD — DBpedia R-GCN + self-attention pool



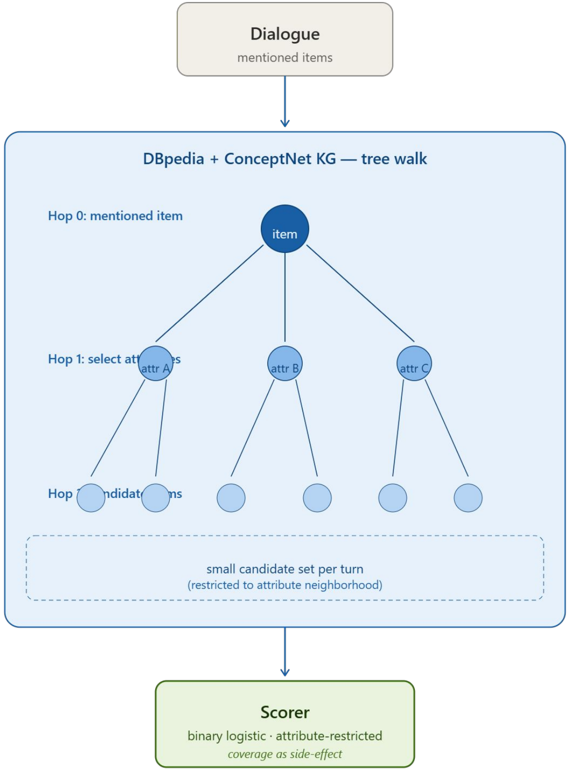

*User rep. built only from mentioned entities — no dialogue text signal*

## KGSF — gated fusion of DBpedia and ConceptNet



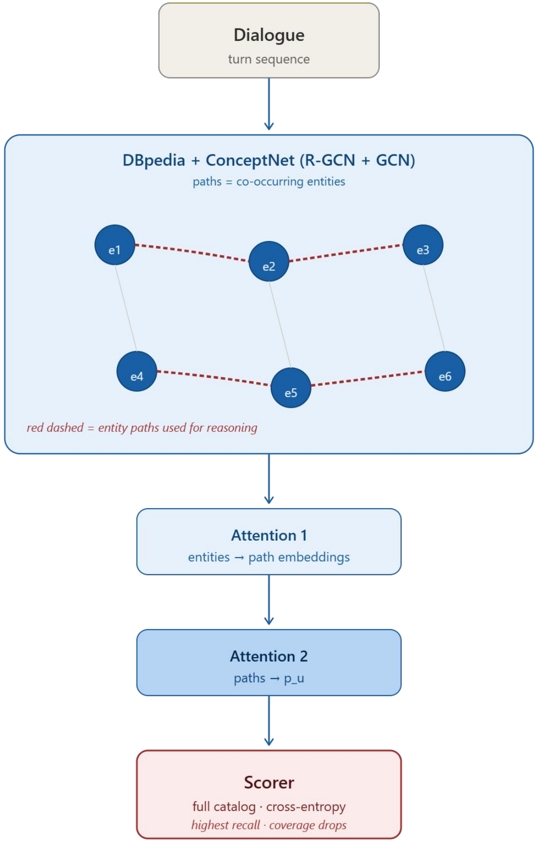

*Two KGs fused via gate — but still full-catalog cross-entropy*

## CR-Walker — tree walk over the KG (attribute-restricted)




*Higher baseline coverage — side-effect of restricted scoring, not optimized diversity*

## TREA — multi-hierarchical reasoning over entity paths




*Best recall in survey — but full-catalog CE returns the coverage problem*

# Problem Statement: Existing Diversity Loss

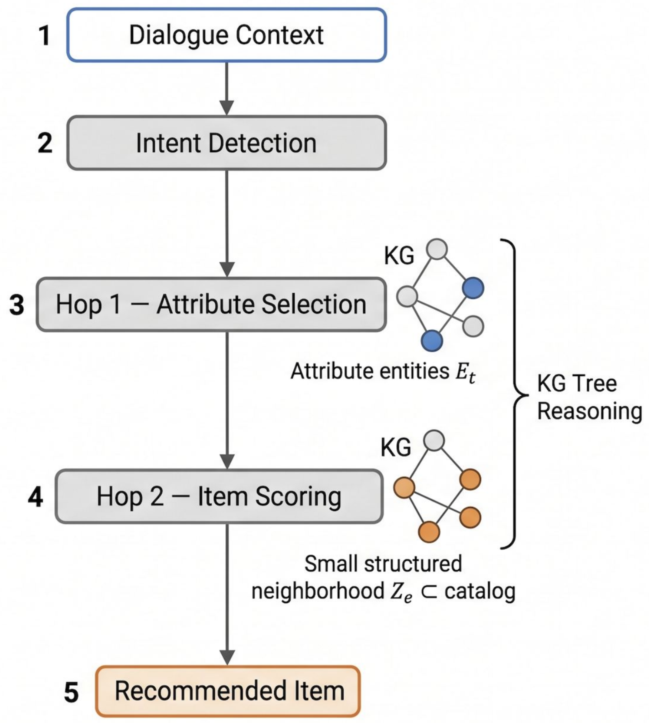
- DPP (Determinantal Point Process) is the standard diversity-aware subset selection method in recommender systems
- But in practice, computing diversity requires first selecting a discrete subset → that selection step is non-differentiable

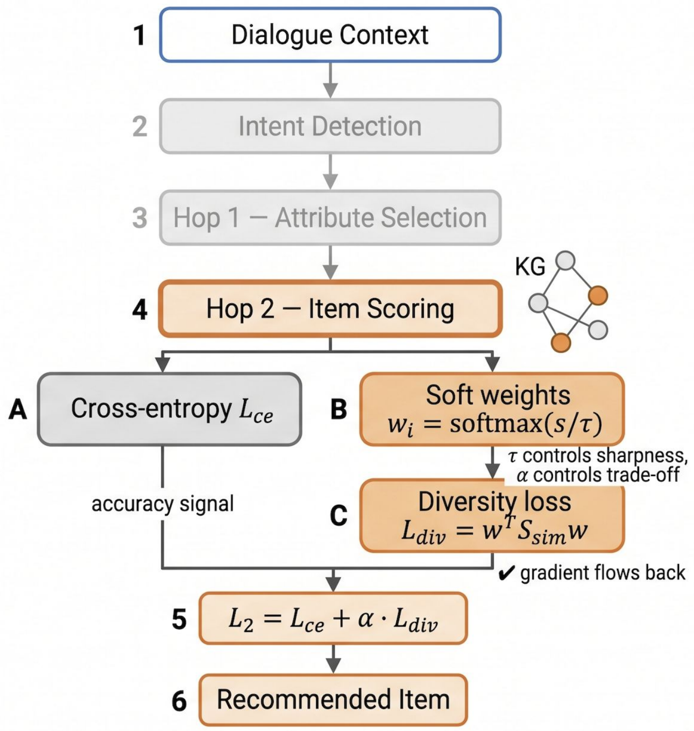
- Motivation: relax the hard top-K with a temperature-scaled softmax:

$$w_{b,i} = \frac{\exp(s_{b,i}/\tau)}{\sum_j \exp(s_{b,j}/\tau)}, \quad \tau > 0$$

- $\tau \to 0$: collapses to the top-1 indicator (recovers hard selection).
- $\tau$ large: keeps mass on lower-ranked items, so the diversity gradient has signal to act on.
- For SRD, $\tau$ used deterministically as a differentiable relaxation.

# CR-Walker: SRD Applied



# Benchmark Datasets

Table 2.1. Summary of two benchmark datasets (ReDial and TG-ReDial)

| Datasets | Dialogues | Movie Items | Language |
|---|---|---|---|
| ReDial | 10,006 | 6,924 | English |
| TG-ReDial | 10,000 | 33,834 | Chinese |

Table 3: Sample conversation from validation set.

| | |
|---|---|
| SEEKER: | hi ! i 'm looking for an action filled movie similar to jurassic park |
| HUMAN: | ok , have you seen jurassic world ? |
| HRED: | hi ! i love that one ! have you seen the conjuring or the conjuring 2 ? |
| OURS: | hi what kind of movies do you like ? |
| SEEKER: | yes i loved that movie as well . are there any good movies without dinosaurs ? |
| HUMAN: | have you seen king kong ? |
| HRED: | no , i have n't seen it yet . have you seen the avengers ? it 's about a couple of friends who are their own . |
| OURS: | i have n't seen that one . i will have to check it out . have you seen jurassic park |
| SEEKER: | yes ! that was another great giant monster movie : ) |
| HUMAN: | what about jaws & amp ; jaws 2 ? oldies but goodies scary action ? suspense gives me the rumblies |
| HRED: | i have n't seen that one yet but i heard it was good . i have n't seen it yet . i have n't seen it yet . i have n't seen it yet . |
| OURS: | i have n't seen that one yet . i will have to check it out . what about walking with dinosaurs ? |

*Image 2.1. Example of a part of conversation from dataset ReDial*

| Role | Topic | Utterance |
|---|---|---|
| Seeker | greeting | Hello, do you have any film recommendations for me tonight? |
| Recommender | genre: romance | Sure, are you in the mood for something romantic? |
| Seeker | mood | Something light and warm, please. |
| Recommender | *Love Letter* | Then I would recommend *Love Letter (1995)*. |
| Seeker | feedback | I have seen that one already. |
| Recommender | *Beijing Love Story* | How about *Beijing Love Story (2014)*? |
| Seeker | accept | That sounds good, I will watch it. |

*Image 2.2. Example of a conversation from dataset TG-ReDial*

**Image 2.1: https://redialdata.github.io/website/assets/img/ex1.png*

# Benchmark Datasets

Table 2.1. Summary of two benchmark datasets (ReDial and TG-ReDial)

| Datasets | Dialogues | Movie Items | Language |
|---|---|---|---|
| ReDial | 10,006 | 6,924 | English |
| TG-ReDial | 10,000 | 33,834 | Chinese |

Table 2.2. Summary of 3 Knowledge Graphs utilized in this thesis

| Subgraph | Key relations | Entities | Focus | Used by |
|---|---|---|---|---|
| DBpedia | cast, director, producer, writer, distributor, country | ~30K | Entity-level film facts | All backbones |
| ConceptNet | IsA, HasProperty, RelatedTo, UsedFor | ~29K | Word-level commonsense | KGSF |
| CN-DBpedia | cast, director, country, year, topic links | ~54K | Chinese entity facts + topic-thread connectivity | TG-Redial |

# Related Work: Metrics

## Accuracy

***Recall@K:*** Is the gold item in the top-K list

$$Recall@K = \frac{1}{|\mathcal{U}|}\sum_{u\in\mathcal{U}} hit_u(K)$$

$\mathcal{U}$ represents the set of all users, $u$ is an individual user, and $hit_u(K)$ is an indicator that equals 1 if the rank $r_u$ of the ground-truth item is $\leq K$, and 0 otherwise

***NDCG@K*** (Normalized Discounted Cumulative Gain): Does it rank higher up the list

$$NDCG@K = \frac{1}{|\mathcal{U}|}\sum_{u\in\mathcal{U}} \frac{hit_u(K)}{log_2(r_u+1)} / Z_K$$

$\mathcal{U}$ is the user set, $u$ is the user, $hit_u(K)$ indicates a top-K hit, $r_u$ is the rank of the ground-truth item, and $Z_K$ is a normalization constant to ensure the score falls within the range $[0, 1]$

# Related Work: Metrics

## Diversity

***Coverage@K:*** What fraction of all items get recommended across all users?

$$Coverage@K = \frac{|\bigcup_{u \in \mathcal{U}} \mathcal{L}_u|}{|\mathcal{I}|}$$

***Intra-list Diversity (ILD):*** Are items within one user's list different from each other?

$$ILD(\mathcal{L}_u) = \frac{2}{K(K-1)} \sum_{1 \leq a < b \leq K} d(i_a, i_b)$$

***Novelty:*** Are recommended items from the long tail?

$$Novelty(\mathcal{L}_u) = \frac{1}{K} \sum_{i \in \mathcal{L}_u} -log_2 p(i)$$

Annotations:

- $\mathcal{L}_u$ is the top-K list recommended to a specific user $u$
- $\mathcal{U}$ represents the entire user set
- $\mathcal{I}$ denotes the complete item catalog
- $K$ is the number of items in that list
- $d(\cdot,\cdot)$ represents a dissimilarity function (such as $1-\cos(e_a,e_b)$)and $i_a, i_b$ are the items located at list positions $a$ and $b$
- $p(i)$ represents the empirical interaction frequency of item $i$ in the dataset

# Related Work: Metrics

## Harmonic Mean

$$F1@K = \frac{2 \cdot M_1@K \cdot M_2@K}{M_1@K + M_2@K}$$

$M_1@K$ and $M_2@K$ represent the respective values of the two metrics being combined, such as Recall and Coverage

# Related Work: SoTA Results

## ReDial (test set, percentages)

| Model | R@1 | R@10 | R@50 | Cov@1 | Cov@10 | Cov@50 | F1@10 |
|---|---|---|---|---|---|---|---|
| *KBRD* | 3.70 | 18.79 | 35.87 | 6.48 | 18.37 | 36.37 | 18.58 |
| *KGSF* | 3.97 | 19.36 | 38.37 | 6.41 | 17.03 | 29.92 | 18.12 |
| *CR-Walker* | 3.65 | 15.37 | 33.33 | **10.25** | **31.76** | **62.19** | 20.72 |
| *TREA* | **4.84** | **21.37** | **41.88** | 8.07 | 23.87 | 45.26 | **22.55** |

## TG-ReDial (test set, percentages)

| Model | R@1 | R@10 | R@50 | Cov@1 | Cov@10 | Cov@50 | F1@10 |
|---|---|---|---|---|---|---|---|
| *KBRD* | 0.27 | 1.20 | 5.12 | 1.93 | 9.58 | 20.58 | 2.13 |
| *KGSF* | 0.22 | 1.74 | 6.77 | 2.69 | 14.20 | **29.08** | 3.10 |
| *CR-Walker* | 0.19 | 1.64 | 4.09 | 3.42 | 5.34 | 12.47 | 2.50 |
| *TREA* | **1.02** | **3.21** | **9.05** | **3.63** | **16.40** | 28.92 | **5.37** |

# Related Work: SoTA Results

## ReDial (test set, percentages)

| Model | R@1 | R@10 | R@50 | Cov@1 | Cov@10 | Cov@50 | F1@10 |
|---|---|---|---|---|---|---|---|
| ***KBRD*** | 3.70 | 18.79 | 35.87 | 6.48 | 18.37 | 36.37 | 18.58 |
| ***KGSF*** | 3.97 | 19.36 | 38.37 | 6.41 | 17.03 | 29.92 | 18.12 |
| ***CR-Walker*** | 3.65 | 15.37 | 33.33 | **10.25** | **31.76** | **62.19** | 20.72 |
| ***TREA*** | **4.84** | **21.37** | **41.88** | 8.07 | 23.87 | 45.26 | **22.55** |

***Problem***: All four optimise the ***same target***: cross-entropy against the gold item from a long-tailed catalog.

***Goal:*** lift catalog coverage substantially while keeping recall close to the best-in-class supervised baseline with one training-time fix that generalises to any KG-CRS backbone.

# Building the soft loss step-by-step

*Step 1: Normalize item embeddings*

$$\hat{h}_i = \frac{h_i}{\|h_i\|_2}$$

where $h_i$ is vector representation of item $i$

*Step 2: Form item-item cosine similarity matrix*

$$S_{\text{sim}} = \hat{H}\hat{H}^T \quad // \; (S_{\text{sim}})_{ij} = \cos\left(\hat{h}_i, \hat{h}_j\right) \in [-1, 1]$$

*Step 3: Weighted mean pairwise cosine similarity for turn $b$ is*

$$\ell_b^{\text{div}} = \sum_{i=1}^{M}\sum_{j=1}^{M} w_{b,i}\, w_{b,j} \cos(\hat{\mathbf{h}}_i, \hat{\mathbf{h}}_j) = \mathbf{w}_b^{\top}\mathbf{S}_{\text{sim}}\,\mathbf{w}_b$$

*Step 4: Finalize the batch-level diversity loss averages over all dialogue turns*

$$\mathcal{L}_{\text{div}} = \frac{1}{B}\sum_{b=1}^{B} \ell_b^{\text{div}}$$

High model bets on similar items → penalized

Low model spreads across diverse items → rewarded

# Can we optimize the formula ?

With M candidate items (6.9K for ReDial, 33.8K for TG-ReDial), the similarity matrix S has $M^2$entries
$\rightarrow$ recomputing it every batch is prohibitive.

- We have $SW = (\hat{H}\hat{H}^T)W = \hat{H}(\hat{H}^T W) \rightarrow O(dMB)$ instead of $O(M^2 d)$

→ *Same result. Scales linearly with catalog size.*

The pipeline become:

1. First, project weights into embedding space
2. Then, score all items → S is never materialized

# Ablation: Removing the supervised anchor breaks the policy

Reducing the weight of $L_{rec}$ in the final loss produce just **subtle decrease** in recall. However, if remove $L_{rec}$ completely (w=0), **coverage spikes** but **recall collapses**.

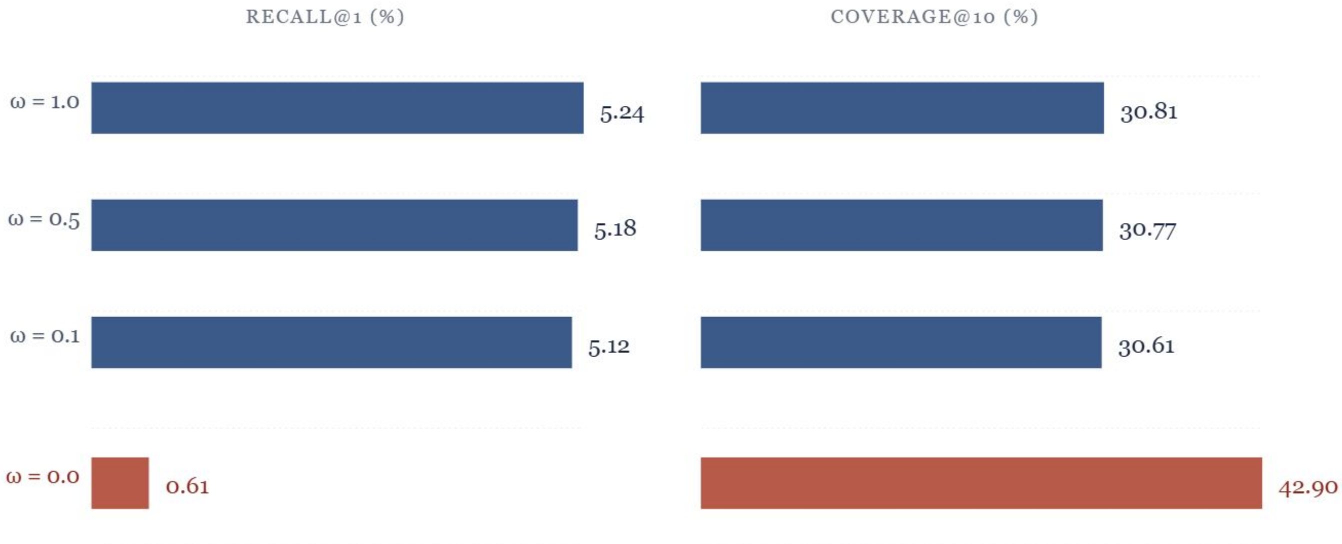



*Figure 4.4. Effect of the supervised-anchor weight ω on TREA/ReDial.*
*All other hyperparameters held fixed ($\alpha_1 = \alpha_3 = 1$, $\alpha_2 = 0$)*

# Ablation: No directional pattern in the effect of the reward terms

No directional pattern in the effect of NDCG or novelty variant. Toggling NDCG or swapping novelty type shifts no metric consistently.


*Figure 4.5. Measuring the effect of NDCG across three metrics, with relevance-weight novelty and plain novelty*

# Appendix: Annotations for generic formula

Table a.4. List of annotations explanation

| Annotation | Context/Description |
|---|---|
| $\mathcal{L}_u$ | is the top-K list recommended to a specific user $u$ |
| $\mathcal{U}$ | represents the entire user set |
| $u$ | an individual user |
| $\mathcal{I}$ | denotes the complete item catalog |
| $K$ | The number of items in the recommendation list |
| $d(\cdot,\cdot)$ | Supervised-anchor weight sweep |