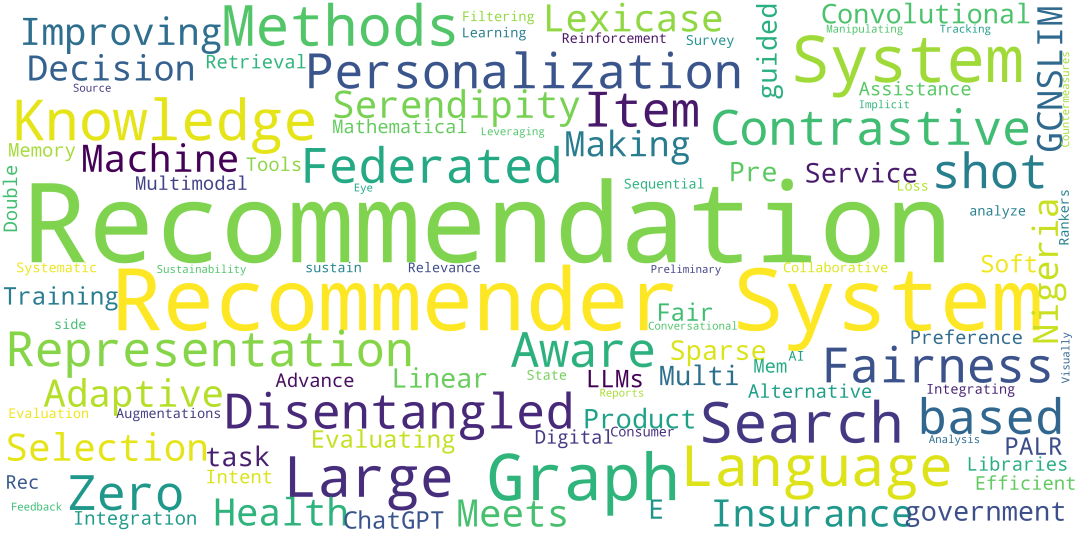
# 论文周报 | 推荐系统领域最新研究进展

原创 ML\_RSer [机器学习与推荐算法](javascript:void(0);) 2023-05-22 08:10 发表于北京

收录于合集#推荐系统干货分享239个

**嘿，记得给“机器学习与推荐算法”添加星标**

本文精选了上周（0515-0521）最新发布的16篇推荐系统相关论文，主要研究方向包括大型语言模型赋能推荐系统、推荐中的公平性、搜索助力推荐、图推荐系统、隐私保护推荐系统、对话推荐系统（来自谷歌）等。



以下整理了论文标题以及摘要，如感兴趣可移步原文精读。

1. Improving Recommendation System Serendipity Through Lexicase Selection

2. Adaptive Graph Contrastive Learning for Recommendation

3. When Search Meets Recommendation: Learning Disentangled Search  Representation for Recommendation, SIGIR2023

4. GCNSLIM: Graph Convolutional Network with Sparse Linear Methods for  E-government Service Recommendation

5. Graph-guided Personalization for Federated Recommendation

6. Zero-shot Item-based Recommendation via Multi-task Product Knowledge  Graph Pre-Training

7. Is ChatGPT Fair for Recommendation? Evaluating Fairness in Large  Language Model Recommendation

8. Knowledge Soft Integration for Multimodal Recommendation

9. Mem-Rec: Memory Efficient Recommendation System using Alternative  Representation

10. Preference or Intent? Double Disentangled Collaborative Filtering

11. Contrastive State Augmentations for Reinforcement Learning-Based  Recommender Systems, SIGIR2023

12. Integrating Item Relevance in Training Loss for Sequential Recommender Systems

13. Consumer-side Fairness in Recommender Systems: A Systematic Survey of  Methods and Evaluation

14. Large Language Models are Zero-Shot Rankers for Recommender Systems

15. Manipulating Visually-aware Federated Recommender Systems and Its  Countermeasures

16. Leveraging Large Language Models in Conversational Recommender Systems, from Google

### **1. Improving Recommendation System Serendipity Through Lexicase Selection**

Ryan Boldi, Aadam Lokhandwala, Edward Annatone, Yuval Schechter, Alexander Lavrenenko, Cooper Sigrist

https://arxiv.org/abs/2305.11044

Recommender systems influence almost every aspect of our digital lives. Unfortunately, in striving to give us what we want, they end up restricting our open-mindedness. Current recommender systems promote echo chambers, where people only see the information they want to see, and homophily, where users of similar background see similar content. We propose a new serendipity metric to measure the presence of echo chambers and homophily in recommendation systems using cluster analysis. We then attempt to improve the diversity-preservation qualities of well known recommendation techniques by adopting a parent selection algorithm from the evolutionary computation literature known as lexicase selection. Our results show that lexicase selection, or a mixture of lexicase selection and ranking, outperforms its purely ranked counterparts in terms of personalization, coverage and our specifically designed serendipity benchmark, while only slightly under-performing in terms of accuracy (hit rate). We verify these results across a variety of recommendation list sizes. In this work we show that lexicase selection is able to maintain multiple diverse clusters of item recommendations that are each relevant for the specific user, while still maintaining a high hit-rate accuracy, a trade off that is not achieved by other methods.

### **2. Adaptive Graph Contrastive Learning for Recommendation**

Yangqin Jiang, Chao Huang, Lianghao Xia

https://arxiv.org/abs/2305.10837

Recently, graph neural networks (GNNs) have been successfully applied to recommender systems as an effective collaborative filtering (CF) approach. The key idea of GNN-based recommender system is to recursively perform the message passing along the user-item interaction edge for refining the encoded embeddings, relying on sufficient and high-quality training data. Since user behavior data in practical recommendation scenarios is often noisy and exhibits skewed distribution, some recommendation approaches, e.g., SGL and SimGCL, leverage self-supervised learning to improve user representations against the above issues. Despite their effectiveness, however, they conduct self-supervised learning through creating contrastvie views, depending on the exploration of data augmentations with the problem of tedious trial-and-error selection of augmentation methods. In this paper, we propose a novel Adaptive Graph Contrastive Learning (AdaptiveGCL) framework which conducts graph contrastive learning with two adaptive contrastive view generators to better empower CF paradigm. Specifically, we use two trainable view generators, which are a graph generative model and a graph denoising model respectively, to create contrastive views. Two generators are able to create adaptive contrastive views, addressing the problem of model collapse and achieving adaptive contrastive learning. With two adaptive contrasive views, more additionally high-quality training signals will be introduced into the CF paradigm and help to alleviate the data sparsity and noise issues. Extensive experiments on three benchmark datasets demonstrate the superiority of our model over various state-of-the-art recommendation methods. Further visual analysis intuitively explains why our AdaptiveGCL outperforms existing contrastive learning approaches based on selected data augmentation methods.

### **3. When Search Meets Recommendation: Learning Disentangled Search  Representation for Recommendation, SIGIR2023**

Zihua Si, Zhongxiang Sun, Xiao Zhang, Jun Xu, Xiaoxue Zang, Yang Song, Kun Gai, Ji-Rong Wen

https://arxiv.org/abs/2305.10822

Modern online service providers such as online shopping platforms often provide both search and recommendation (S&R) services to meet different user needs. Rarely has there been any effective means of incorporating user behavior data from both S&R services. Most existing approaches either simply treat S&R behaviors separately, or jointly optimize them by aggregating data from both services, ignoring the fact that user intents in S&R can be distinctively different. In our paper, we propose a Search-Enhanced framework for the Sequential Recommendation (SESRec) that leverages users' search interests for recommendation, by disentangling similar and dissimilar representations within S&R behaviors. Specifically, SESRec first aligns query and item embeddings based on users' query-item interactions for the computations of their similarities. Two transformer encoders are used to learn the contextual representations of S&R behaviors independently. Then a contrastive learning task is designed to supervise the disentanglement of similar and dissimilar representations from behavior sequences of S&R. Finally, we extract user interests by the attention mechanism from three perspectives, i.e., the contextual representations, the two separated behaviors containing similar and dissimilar interests. Extensive experiments on both industrial and public datasets demonstrate that SESRec consistently outperforms state-of-the-art models. Empirical studies further validate that SESRec successfully disentangle similar and dissimilar user interests from their S&R behaviors.

### **4. GCNSLIM: Graph Convolutional Network with Sparse Linear Methods for  E-government Service Recommendation**

Lingyuan Kong, Hao Ding, Guangwei Hu

https://arxiv.org/abs/2305.08586

Graph Convolutional Networks have made significant strides in Collabora-tive Filtering recommendations. However, existing GCN-based CF methods are mainly based on matrix factorization and incorporate some optimization tech-niques to enhance performance, which are not enough to handle the complexities of diverse real-world recommendation scenarios. E-government service recommendation is a crucial area for recommendation re-search as it involves rigid aspects of people's lives. However, it has not received ad-equate attention in comparison to other recommendation scenarios like news and music recommendation. We empirically find that when existing GCN-based CF methods are directly applied to e-government service recommendation, they are limited by the MF framework and showing poor performance. This is because MF's equal treatment of users and items is not appropriate for scenarios where the number of users and items is unbalanced. In this work, we propose a new model, GCNSLIM, which combines GCN and sparse linear methods instead of combining GCN and MF to accommodate e-government service recommendation. In particular, GCNSLIM explicitly injects high-order collaborative signals obtained from multi-layer light graph convolutions into the item similarity matrix in the SLIM frame-work, effectively improving the recommendation accuracy. In addition, we propose two optimization measures, removing layer 0 embedding and adding nonlinear acti-vation, to further adapt to the characteristics of e-government service recommenda-tion scenarios. Furthermore, we propose a joint optimization mode to adapt to more diverse recommendation scenarios. We conduct extensive experiments on a real e-government service dataset and a common public dataset and demonstrate the ef-fectiveness of GCNSLIM in recommendation accuracy and operational performance.

### **5. Graph-guided Personalization for Federated Recommendation**

Chunxu Zhang, Guodong Long, Tianyi Zhou, Peng Yan, Zijjian Zhang, Bo Yang

https://arxiv.org/abs/2305.07866

Federated Recommendation is a new service architecture providing recommendations without sharing user data with the server. Existing methods deploy a recommendation model on each client and coordinate their training by synchronizing and aggregating item embeddings. However, while users usually hold diverse preferences toward certain items, these methods indiscriminately aggregate item embeddings from all clients, neutralizing underlying user-specific preferences. Such neglect will leave the aggregated embedding less discriminative and hinder personalized recommendations. This paper proposes a novel Graph-guided Personalization framework (GPFedRec) for the federated recommendation. The GPFedRec enhances cross-client collaboration by leveraging an adaptive graph structure to capture the correlation of user preferences. Besides, it guides training processes on clients by formulating them into a unified federated optimization framework, where models can simultaneously use shared and personalized user preferences. Experiments on five benchmark datasets demonstrate GPFedRec's superior performance in providing personalized recommendations.

### **6. Zero-shot Item-based Recommendation via Multi-task Product Knowledge  Graph Pre-Training**

Ziwei Fan, Zhiwei Liu, Shelby Heinecke, Jianguo Zhang, Huan Wang, Caiming Xiong, Philip S. Yu

https://arxiv.org/abs/2305.07633

Existing recommender systems face difficulties with zero-shot items, i.e. items that have no historical interactions with users during the training stage. Though recent works extract universal item representation via pre-trained language models (PLMs), they ignore the crucial item relationships. This paper presents a novel paradigm for the Zero-Shot Item-based Recommendation (ZSIR) task, which pre-trains a model on product knowledge graph (PKG) to refine the item features from PLMs. We identify three challenges for pre-training PKG, which are multi-type relations in PKG, semantic divergence between item generic information and relations and domain discrepancy from PKG to downstream ZSIR task. We address the challenges by proposing four pre-training tasks and novel task-oriented adaptation (ToA) layers. Moreover, this paper discusses how to fine-tune the model on new recommendation task such that the ToA layers are adapted to ZSIR task. Comprehensive experiments on 18 markets dataset are conducted to verify the effectiveness of the proposed model in both knowledge prediction and ZSIR task.

### **7. Is ChatGPT Fair for Recommendation? Evaluating Fairness in Large  Language Model Recommendation**

Jizhi Zhang, Keqin Bao, Yang Zhang, Wenjie Wang, Fuli Feng, Xiangnan He

https://arxiv.org/abs/2305.07609

The remarkable achievements of Large Language Models (LLMs) have led to the emergence of a novel recommendation paradigm -- Recommendation via LLM (RecLLM). Nevertheless, it is important to note that LLMs may contain social prejudices, and therefore, the fairness of recommendations made by RecLLM requires further investigation. To avoid the potential risks of RecLLM, it is imperative to evaluate the fairness of RecLLM with respect to various sensitive attributes on the user side. Due to the differences between the RecLLM paradigm and the traditional recommendation paradigm, it is problematic to directly use the fairness benchmark of traditional recommendation. To address the dilemma, we propose a novel benchmark called Fairness of Recommendation via LLM (FaiRLLM). This benchmark comprises carefully crafted metrics and a dataset that accounts for eight sensitive attributes1 in two recommendation scenarios: music and movies. By utilizing our FaiRLLM benchmark, we conducted an evaluation of ChatGPT and discovered that it still exhibits unfairness to some sensitive attributes when generating recommendations. Our code and dataset can be found at https://github.com/jizhi-zhang/FaiRLLM

### **8. Knowledge Soft Integration for Multimodal Recommendation**

Kai Ouyang, Chen Tang, Wenhao Zheng, Xiangjin Xie, Xuanji Xiao, Jian Dong, Hai-Tao Zheng, Zhi Wang

https://arxiv.org/abs/2305.07419

One of the main challenges in modern recommendation systems is how to effectively utilize multimodal content to achieve more personalized recommendations. Despite various proposed solutions, most of them overlook the mismatch between the knowledge gained from independent feature extraction processes and downstream recommendation tasks. Specifically, multimodal feature extraction processes do not incorporate prior knowledge relevant to recommendation tasks, while recommendation tasks often directly use these multimodal features as side information. This mismatch can lead to model fitting biases and performance degradation, which this paper refers to as the \textit{curse of knowledge} problem. To address this issue, we propose using knowledge soft integration to balance the utilization of multimodal features and the curse of knowledge problem it brings about. To achieve this, we put forward a Knowledge Soft Integration framework for the multimodal recommendation, abbreviated as KSI, which is composed of the Structure Efficiently Injection (SEI) module and the Semantic Soft Integration (SSI) module. In the SEI module, we model the modality correlation between items using Refined Graph Neural Network (RGNN), and introduce a regularization term to reduce the redundancy of user/item representations. In the SSI module, we design a self-supervised retrieval task to further indirectly integrate the semantic knowledge of multimodal features, and enhance the semantic discrimination of item representations. Extensive experiments on three benchmark datasets demonstrate the superiority of KSI and validate the effectiveness of its two modules.

### **9. Mem-Rec: Memory Efficient Recommendation System using Alternative  Representation**

Gopi Krishna Jha, Anthony Thomas, Nilesh Jain, Sameh Gobriel, Tajana Rosing, Ravi Iyer

https://arxiv.org/abs/2305.07205

Deep learning-based recommendation systems (e.g., DLRMs) are widely used AI models to provide high-quality personalized recommendations. Training data used for modern recommendation systems commonly includes categorical features taking on tens-of-millions of possible distinct values. These categorical tokens are typically assigned learned vector representations, that are stored in large embedding tables, on the order of 100s of GB. Storing and accessing these tables represent a substantial burden in commercial deployments. Our work proposes MEM-REC, a novel alternative representation approach for embedding tables. MEM-REC leverages bloom filters and hashing methods to encode categorical features using two cache-friendly embedding tables. The first table (token embedding) contains raw embeddings (i.e. learned vector representation), and the second table (weight embedding), which is much smaller, contains weights to scale these raw embeddings to provide better discriminative capability to each data point. We provide a detailed architecture, design and analysis of MEM-REC addressing trade-offs in accuracy and computation requirements, in comparison with state-of-the-art techniques. We show that MEM-REC can not only maintain the recommendation quality and significantly reduce the memory footprint for commercial scale recommendation models but can also improve the embedding latency. In particular, based on our results, MEM-REC compresses the MLPerf CriteoTB benchmark DLRM model size by 2900x and performs up to 3.4x faster embeddings while achieving the same AUC as that of the full uncompressed model.

### **10. Preference or Intent? Double Disentangled Collaborative Filtering**

Chao Wang, Hengshu Zhu, Dazhong Shen, Wei wu, Hui Xiong

https://arxiv.org/abs/2305.11084

People usually have different intents for choosing items, while their preferences under the same intent may also different. In traditional collaborative filtering approaches, both intent and preference factors are usually entangled in the modeling process, which significantly limits the robustness and interpretability of recommendation performances. For example, the low-rating items are always treated as negative feedback while they actually could provide positive information about user intent. To this end, in this paper, we propose a two-fold representation learning approach, namely Double Disentangled Collaborative Filtering (DDCF), for personalized recommendations. The first-level disentanglement is for separating the influence factors of intent and preference, while the second-level disentanglement is performed to build independent sparse preference representations under individual intent with limited computational complexity. Specifically, we employ two variational autoencoder networks, intent recognition network and preference decomposition network, to learn the intent and preference factors, respectively. In this way, the low-rating items will be treated as positive samples for modeling intents while the negative samples for modeling preferences. Finally, extensive experiments on three real-world datasets and four evaluation metrics clearly validate the effectiveness and the interpretability of DDCF.

### **11. Contrastive State Augmentations for Reinforcement Learning-Based  Recommender Systems, SIGIR2023**

Zhaochun Ren, Na Huang, Yidan Wang, Pengjie Ren, Jun Ma, Jiahuan Lei, Xinlei Shi, Hengliang Luo, Joemon M Jose, Xin Xin

https://arxiv.org/abs/2305.11081

Learning reinforcement learning (RL)-based recommenders from historical user-item interaction sequences is vital to generate high-reward recommendations and improve long-term cumulative benefits. However, existing RL recommendation methods encounter difficulties (i) to estimate the value functions for states which are not contained in the offline training data, and (ii) to learn effective state representations from user implicit feedback due to the lack of contrastive signals. In this work, we propose contrastive state augmentations (CSA) for the training of RL-based recommender systems. To tackle the first issue, we propose four state augmentation strategies to enlarge the state space of the offline data. The proposed method improves the generalization capability of the recommender by making the RL agent visit the local state regions and ensuring the learned value functions are similar between the original and augmented states. For the second issue, we propose introducing contrastive signals between augmented states and the state randomly sampled from other sessions to improve the state representation learning further. To verify the effectiveness of the proposed CSA, we conduct extensive experiments on two publicly accessible datasets and one dataset collected from a real-life e-commerce platform. We also conduct experiments on a simulated environment as the online evaluation setting. Experimental results demonstrate that CSA can effectively improve recommendation performance.

### **12. Integrating Item Relevance in Training Loss for Sequential Recommender Systems**

Andrea Bacciu, Federico Siciliano, Nicola Tonellotto, Fabrizio Silvestri

https://arxiv.org/abs/2305.10824

Sequential Recommender Systems (SRSs) are a popular type of recommender system that learns from a user's history to predict the next item they are likely to interact with. However, user interactions can be affected by noise stemming from account sharing, inconsistent preferences, or accidental clicks. To address this issue, we (i) propose a new evaluation protocol that takes multiple future items into account and (ii) introduce a novel relevance-aware loss function to train a SRS with multiple future items to make it more robust to noise. Our relevance-aware models obtain an improvement of ~1.2% of NDCG@10 and 0.88% in the traditional evaluation protocol, while in the new evaluation protocol, the improvement is ~1.63% of NDCG@10 and ~1.5% of HR w.r.t the best performing models.

### **13. Consumer-side Fairness in Recommender Systems: A Systematic Survey of  Methods and Evaluation**

Bjørnar Vassøy, Helge Langseth

https://arxiv.org/abs/2305.09330

In the current landscape of ever-increasing levels of digitalization, we are facing major challenges pertaining to scalability. Recommender systems have become irreplaceable both for helping users navigate the increasing amounts of data and, conversely, aiding providers in marketing products to interested users. The growing awareness of discrimination in machine learning methods has recently motivated both academia and industry to research how fairness can be ensured in recommender systems. For recommender systems, such issues are well exemplified by occupation recommendation, where biases in historical data may lead to recommender systems relating one gender to lower wages or to the propagation of stereotypes. In particular, consumer-side fairness, which focuses on mitigating discrimination experienced by users of recommender systems, has seen a vast number of diverse approaches for addressing different types of discrimination. The nature of said discrimination depends on the setting and the applied fairness interpretation, of which there are many variations. This survey serves as a systematic overview and discussion of the current research on consumer-side fairness in recommender systems. To that end, a novel taxonomy based on high-level fairness interpretation is proposed and used to categorize the research and their proposed fairness evaluation metrics. Finally, we highlight some suggestions for the future direction of the field.

### **14. Large Language Models are Zero-Shot Rankers for Recommender Systems**

Yupeng Hou, Junjie Zhang, Zihan Lin, Hongyu Lu, Ruobing Xie, Julian McAuley, Wayne Xin Zhao

https://arxiv.org/abs/2305.08845

Recently, large language models (LLMs) (e.g. GPT-4) have demonstrated impressive general-purpose task-solving abilities, including the potential to approach recommendation tasks. Along this line of research, this work aims to investigate the capacity of LLMs that act as the ranking model for recommender systems. To conduct our empirical study, we first formalize the recommendation problem as a conditional ranking task, considering sequential interaction histories as conditions and the items retrieved by the candidate generation model as candidates. We adopt a specific prompting approach to solving the ranking task by LLMs: we carefully design the prompting template by including the sequential interaction history, the candidate items, and the ranking instruction. We conduct extensive experiments on two widely-used datasets for recommender systems and derive several key findings for the use of LLMs in recommender systems. We show that LLMs have promising zero-shot ranking abilities, even competitive to or better than conventional recommendation models on candidates retrieved by multiple candidate generators. We also demonstrate that LLMs struggle to perceive the order of historical interactions and can be affected by biases like position bias, while these issues can be alleviated via specially designed prompting and bootstrapping strategies. The code to reproduce this work is available at https://github.com/RUCAIBox/LLMRank

### **15. Manipulating Visually-aware Federated Recommender Systems and Its  Countermeasures**

Wei Yuan, Shilong Yuan, Chaoqun Yang, Quoc Viet Hung Nguyen, Hongzhi Yin

https://arxiv.org/abs/2305.08183

Federated recommender systems (FedRecs) have been widely explored recently due to their ability to protect user data privacy. In FedRecs, a central server collaboratively learns recommendation models by sharing model public parameters with clients, thereby offering a privacy-preserving solution. Unfortunately, the exposure of model parameters leaves a backdoor for adversaries to manipulate FedRecs. Existing works about FedRec security already reveal that items can easily be promoted by malicious users via model poisoning attacks, but all of them mainly focus on FedRecs with only collaborative information (i.e., user-item interactions). We argue that these attacks are effective because of the data sparsity of collaborative signals. In practice, auxiliary information, such as products' visual descriptions, is used to alleviate collaborative filtering data's sparsity. Therefore, when incorporating visual information in FedRecs, all existing model poisoning attacks' effectiveness becomes questionable. In this paper, we conduct extensive experiments to verify that incorporating visual information can beat existing state-of-the-art attacks in reasonable settings. However, since visual information is usually provided by external sources, simply including it will create new security problems. Specifically, we propose a new kind of poisoning attack for visually-aware FedRecs, namely image poisoning attacks, where adversaries can gradually modify the uploaded image to manipulate item ranks during FedRecs' training process. Furthermore, we reveal that the potential collaboration between image poisoning attacks and model poisoning attacks will make visually-aware FedRecs more vulnerable to being manipulated. To safely use visual information, we employ a diffusion model in visually-aware FedRecs to purify each uploaded image and detect the adversarial images.

### **16. Leveraging Large Language Models in Conversational Recommender Systems, from Google**

Luke Friedman, Sameer Ahuja, David Allen, Zhenning Tan, Hakim Sidahmed, Changbo Long, Jun Xie, Gabriel Schubiner, Ajay Patel, Harsh Lara, Brian Chu, Zexi Chen, Manoj Tiwari

https://arxiv.org/abs/2305.07961

A Conversational Recommender System (CRS) offers increased transparency and control to users by enabling them to engage with the system through a real-time multi-turn dialogue. Recently, Large Language Models (LLMs) have exhibited an unprecedented ability to converse naturally and incorporate world knowledge and common-sense reasoning into language understanding, unlocking the potential of this paradigm. However, effectively leveraging LLMs within a CRS introduces new technical challenges, including properly understanding and controlling a complex conversation and retrieving from external sources of information. These issues are exacerbated by a large, evolving item corpus and a lack of conversational data for training. In this paper, we provide a roadmap for building an end-to-end large-scale CRS using LLMs. In particular, we propose new implementations for user preference understanding, flexible dialogue management and explainable recommendations as part of an integrated architecture powered by LLMs. For improved personalization, we describe how an LLM can consume interpretable natural language user profiles and use them to modulate session-level context. To overcome conversational data limitations in the absence of an existing production CRS, we propose techniques for building a controllable LLM-based user simulator to generate synthetic conversations. As a proof of concept we introduce RecLLM, a large-scale CRS for YouTube videos built on LaMDA, and demonstrate its fluency and diverse functionality through some illustrative example conversations.