Research Statement

Hady W. LAUW
School of Computing and Information Systems, Singapore Management University
Tel: (65) 6828-5277; Email: hadywlauw@smu.edu.sg
19 December 2024

Background

The Web has evolved from the broadcasting (few-to-many) paradigm of Web 1.0, where contents are published by established information providers (e.g., news agencies), to the participatory (many-to-many) paradigm of Web 2.0, where content production is now undertaken by many users, who contribute content on various social media platforms such as blogs, Twitter, review sites. This paradigm shift yields research issues on mining the wealth of user-generated contents on the Web, both in the macro, general sense (how to use the wisdom of the crowds to solve various problems), as well as in the micro, personalized sense (how to customize the experience of an individual).

Thus, mining user-generated data on the Web deals with the following ever-increasing challenges and opportunities:

- *Scale*. Algorithms should be cognizant of the large amount of data being generated, and how to deal with them efficiently. However, with the scale of data involved, comes the advantage of redundancies from many distinct sources, allowing us to tap into the wisdom of the crowds.

- *Network*. Every piece of information is connected to another in an inter-connected network. When a user writes a review on an object, that action affects the reputation of the user, the quality of the object being reviewed, other users' opinions of the object, etc. This network gives rise to the opportunity of exploiting the inter-connectivity of objects to propagate information from localities where there is abundant information to those where there is less.

- *User-centric.* In social media, users become the primary concern of analytics, because they are central to the questions of what the quality of information produced is, as well as who might be the potential audience. This focus on users allows us to better understand the behaviors of the different players, and how to derive analytical outputs that are more targeted for smaller groups or even individual users.

In this context, my central concern as a researcher can be summarized as:

> *"designing algorithms for mining user-generated data of various modalities for understanding the behaviors and preferences of users, individually and collectively, and applying the mined knowledge to develop user-centric applications"*

This research statement describes my interests over the years, with its naturally evolving sub-themes illustrated in Figure 1. In my current research undertaking at SMU, I lead the group **Preferred.AI** (https://preferred.ai) on a concerted effort towards preference learning and recommender systems.

Research Areas

A. Mining behaviors from user-generated data

My early research dealt with profiling user behaviors from large amounts of data. Specific questions included how social network among users could be inferred from indirect co-occurrence relations, and how user behaviors when evaluating objects could be discovered from the collective rating data. The guiding insight was the necessity to factor in the inter-connectivity among users, as well as between users and objects they interacted with, when inferring the behavior of users [TKDE08, KDD06,TWEB12, SDM07].

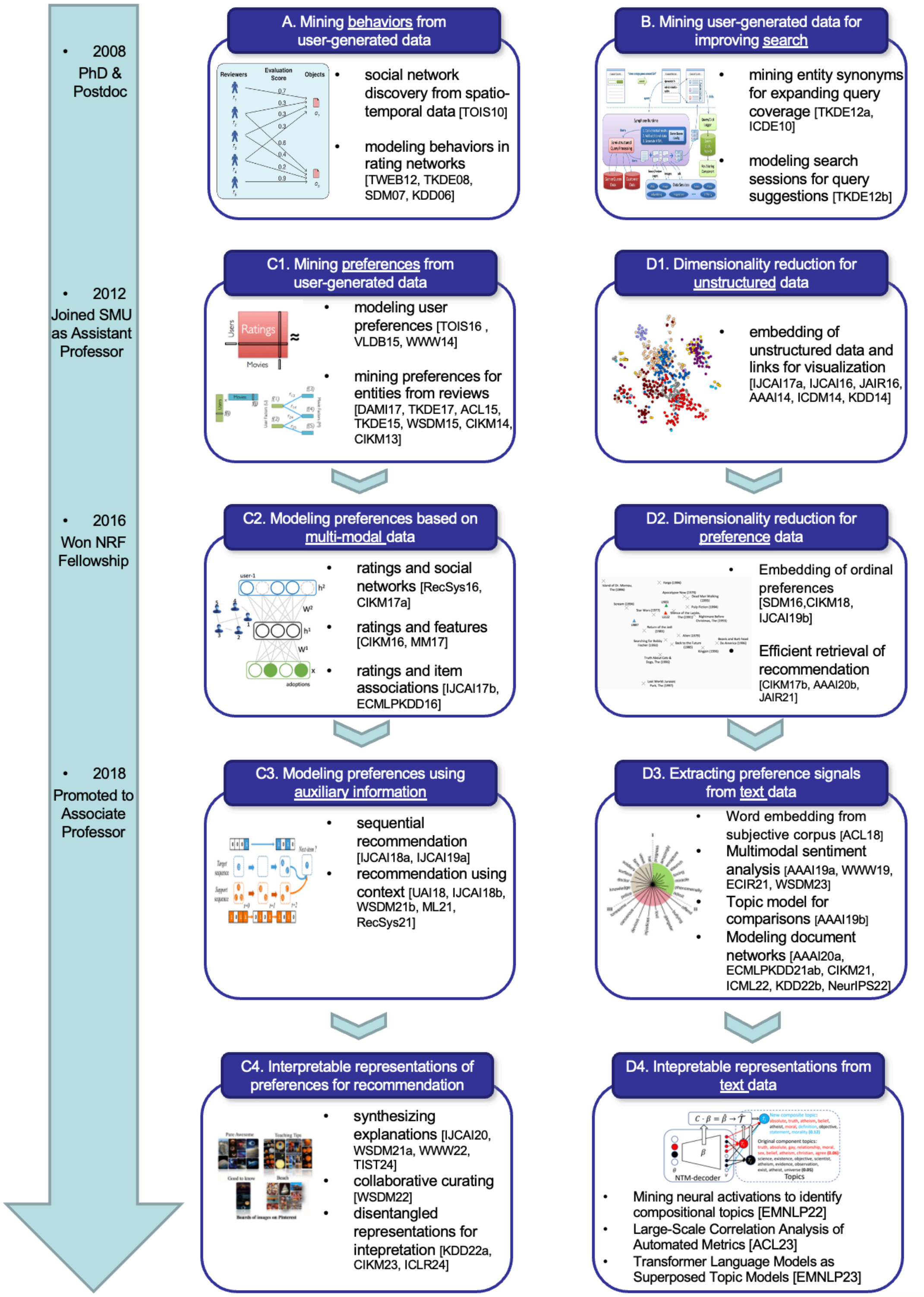

Figure 1 Evolution of My Research Directions

B. Mining user-generated data for improving search

As postdoc at Microsoft Research in Silicon Valley, my research evolved into how the mined knowledge on user behaviors could improve search. For one, we derive synonyms of each record in the database offline, to enable fuzzy matching of queries to records in real time [TKDE12a, ICDE10]. We mine these synonyms by bridging how users refer to specific entities (by examining the query logs capturing issued queries), and how users produce relevant information (by examining the click logs capturing clicked pages representing entities). For another, we track user sessions [TKDE12b] by mining query similarity signals (textual, temporal, and based on search logs such as reformulations and co-clicks) to capture when two queries belong to the same task.

C. Modeling user preferences for recommendation

Since joining SMU in 2012, I have been pursuing primarily two broad research directions. The first is learning not just the behaviors of users, but also their preferences, which could then inform personalization and recommendation in various applications.

C1. Mining preferences from user-generated data

Having some information on the preferences of users is crucial in many applications. The underlying principle is to predict a user's next action based on the historical actions of the user. To alleviate sparsity of information, one approach is to factor in "contexts" as additional information. In [TOIS16, WWW14], we propose a probabilistic model that estimates the probability that two users are likely to agree on a specific item. A user's preference for an item may also be conveyed, not explicitly in terms of ratings or consumption, but implicitly through what the user expresses in reviews. Often, this is done through comparison between pairs of items. In [ACL15], we explore how such comparisons may be identified from reviews, and in [TKDE17, CIKM14], we investigate how comparative sentences can be used to infer the direction of comparisons.

While textual reviews are useful, any individual reader is likely to be overwhelmed by the large quantity of reviews. One issue is how to identify a smaller section of high-quality reviews. In [TKDE15, CIKM13], we propose to use micro-reviews, or concise review-like statements in social media (e.g., Foursquare) frequently accompanied by "check-ins". In [DAMI17, WSDM15], we propose to "synthesize" new reviews by combining segments of existing reviews or micro-reviews.

In another scenario, we are interested in bundling, i.e., how to determine which items should be grouped together to extract the highest utility to sellers and/or consumers. In [VLDB15], we show that optimal bundling is computationally intractable for bundle sizes of three items or more, and develop efficient heuristics for bundle configuration.

C2. Mining preferences based on multi-modal data

Since winning the NRF Fellowship in 2016, I am increasingly focused on mining preferences from *multi-modal* data. One modality is the *social network*. In [RecSys16], we show that a deep model that incorporates social networks into the model structure for modeling rating observations performs better than a wing model that treats social networks as observations alongside ratings. In [CIKM17a], we further investigate the notion of user communities in encoding preferences that align well with topical preferences as learnt from the textual representations of items.

Another modality is the *features* of items or users. In [CIKM16], we model heterogeneous rankings arising from diverse preferences of different preference groups based on features. Examples of item features include its image representation. In [MM17, MM20], we propose augmented convolutional neural networks to detect the sentiments expressed by images, considering the user or item involved.

Associations among items is another modality. One association is which item is preferred by a user after consuming another. In [ECMLPKDD16], we develop sequential recommendation model that incorporates latent groups that affect emissions and context-dependent transition probabilities. Another association is co-occurrence in baskets. In [IJCAI17b], we develop factorization machines sensitive to the basket of items, for recommending the next item that a user may put into her basket.

C3. Modeling preferences using auxiliary information

It is important to pay attention not just to the target interaction, e.g., user's interaction with item, but also to auxiliary information or interactions that could enrich the learning for the target interaction. In [IJCAI18a], we investigate contemporaneous basket sequences, i.e., scenarios where we observe both a target sequence of interest (e.g., purchase) and a supporting sequence that provides context to the target (e.g., clicks). In [ML21], we look at how both organic feedback from an e-commerce site and logged bandit feedback from a target advertising site could be used to predict ads on the target site. In [RecSys21], we explore how auxiliary data from one domain could come in useful as prior to faclitate cross-domain recommendation.

Another form of auxiliary information is item context. What we refer to by 'context' could be coexisting items within the same shopping carts, being browsed within the same session, etc. Such item-item relationships constitute valuable information that would otherwise not easily be derivable from similarities in product attributes alone. In [IJCAI18b, UAI18], we develop probabilistic models based on Poisson Factorization. These models learn item representations both contextually based on their relatedness with other items, as well as collaboratively based on their interactions by users. Auxiliary information can also come from both users and items. While deep latent factor models are useful for collaborative filtering, previous works tend to be one-sided in encoding primarily user representations. We propose Bilateral VAE [WSDM21b] that is suitable for dyadic preference data and capable of accommodating auxiliary data from both sides.

Given the variety of various auxiliary information, we develop Cornac [JMLR20], a library of recommender systems algorithms that makes it easy to conduct comparative experiments involving algorithms and auxiliary data. The library contains 40 recommendation algorithms, and the number is still actively growing. The importance of multi-modality in recommender systems is bolstered by our findings [IC21] that recommendation algorithms designed for one modality often find practical use with other modalities as well, opening up the potential for cross-modality utilization of algorithms.

C4. Interpretable representations of preferences for recommendation

While the focus in most cases is accuracy, increasingly we also value interpretability and explainability. In [IJCAI20], we develop a way to synthesize an explanation for a recommendation, based on snippets gathered from existing reviews of a product. This increases the understandability of why a product is being recommended, lending greater credibility to the recommendation. In [WSDM21], we develop a comparative explanation by identifying aspects in which a product is better than another. In [WWW22], we consider explaining not individual products, but a set of products by summarizing their descriptions. In [TIST24], we model Question and Answers (QA) in an attention mechanism to identify useful reviews. Through joint modeling, we collectively form an explanation in terms of QA and review.

Visual imagery could also enhance interpretability. In [WSDM22], we study collaborative curating, whereby users organize data both in terms of similarity as well as in terms of preferences. Experiments on Pinterest data shows that the act of clustering represents users' conception of what are similar, whereas the act of adopting represents users' inherent preferences.

What makes recommendation models particularly difficult to interpret is the abstractness of the latent factors. In [KDD22a], we look into disentangling or separating the latent factors into multiple components such as by differentiating these components we get a better sense of the factors that drive users' preferences. Moreover, by correlating those disentangled factors with product meta data, these factors are more interpretable. In [CIKM23], we introduce iterative latent attention for personalized item grouping into VAE framework to infer multiple user's interests, incorporate implicit differentiation to ease the training process, and study the interactions between cluster prototypes via self-attention. In [ICLR24], we introduce FACETVAE to resolve shortcomings of VAE-based disentangled recommendation models, characterized by three main innovations 1) disentangling item space under multi-faceted manner, 2) binding compositional user interests from low-level ones discovered from item space and 3) effectively binding user interests via bi-directional binding block.

D. Dimensionality reduction for user-generated data

My second research direction since joining SMU is dimensionality reduction, turning complex high-dimensional data into more compact, yet useful lower-dimensional representations.

D1. Dimensionality reduction for unstructured data

A text document is often represented as a bag of words. This high dimensionality makes exploratory analysis challenging. Embedding is a technique to reduce the dimensionality into a few dimensions (as few as two or three for visualization). Current techniques are not geared towards semantic interpretability because it simply represents data points as coordinates. A semantic interpretation of text is more easily done if there is a topical representation. In [IJCAI17, JAIR16, AAAI14, KDD14], we propose to jointly model two tasks that are usually done separately but are highly complementary, namely: topic modeling for semantic interpretability and embedding for dimensionality reduction. We show that such joint modeling indeed results in more effective visualization that enables exploration of the topical probabilities in documents. Furthermore, since in the Web context, documents are often linked together as a network, we show that incorporating network information creates even more informative embedding of documents [ICDM14]. One application for such joint modeling of topics and embedding is to design word clouds that are more effective for visual comparison of documents by reflecting topical similarities across documents in visual similarities of their word clouds [IJCAI16].

D2. Dimensionality reduction for preference data

Appreciating the synergies between dimensionality reduction and preference mining, I begin exploring dimensionality reduction for preference data. By finding an appropriate embedding for ordinal data, we can place multiple types of objects (e.g., users and items) in the same space [SDM16]. This benefits recommendations via visualization and efficient retrieval [JAIR21]. Aside from preference over items, the user may also express different assessments of whether two items are similar. By incorporating multiple perspectives, we can model different viewpoints of object similarities as well [CIKM18, IJCAI19b]. To further improve the efficiency of top-k retrieval of recommended items, we develop ranking models compatible with multiple structures, including LSH, spatial tree, inverted index [CIKM17b, AAAI20b].

D3. Extracting preference signals from text data

We would like the representations that we extract from data to be able to encode the preferences expressed by the data. In [ACL18] we discover that word embeddings learnt from subjective corpora perform better at sentiment classification, resulting in a novel word embedding SentiVec which is infused with sentiment information from a lexical resource. Still on sentiment analysis, [AAAI19a] looks at the interplay between text and image of a review, discovering that photos in a review play a greater utility as an attention mechanism within a deep learning framework VistaNet, highlighting the salient aspects of an entity, rather than expressing sentiments independently of the text. [WSDM23] looks at the use of concept-oriented transformers. In some cases, these three-way association between words, images, and sentiments could be applied in text-to-image retrieval where different images could be retrieved based on the required sentiment [ECIR21]. We also discover that factoring in images could also help in generating natural language reviews that also help in predicting preferences [WWW19].

One form of dimensionality reduction for text is topic modeling. While previous supervised topic models consider document labels in an independent and pointwise manner, in [AAAI19b] we propose CompareLDA that learns topic distributions that comply with the pairwise comparison observations. In some document corpora, documents are connected to one another in a network. In [AAAI20a] we propose Adjacent-Encoder, neural encoder architectures that induce competition among documents for topic propagation and reconstruction among neighbors for semantic capture. Adjacent-Encoder incorporates the network structure implicitly, with similar number of parameters as Auto-Encoder family yet outperforms the latter. On top of the document network structure, we further incorporate additional information in the form of semi-supervision [ECMLPKDD21a], multi-layered network [ECMLPKDD21b], and listwise comparisons [CIKM21]. More recent approaches frame topic models as graph neural networks, which are further enhanced by temporality [ICML22], author information [KDD22b], supplementing short texts [NeurIPS22], and hyperbolid modeling [KDD23].

D3. Interpretable representations from text data

Another important direction is to examine the interpretability of topic models. Because neural networks involve interactions by multiple neurons, topics may not correspond to individual neurons, but rather compositions of neurons. We analyze the concurrent neural activations to mine compositional topics [EMNLP22]. We conduct a large-scale correlation analysis of coherence metrics for topic models [ACL23]. We also disentangle transformer language models into superposed topic models [EMNLP23].

Selected Publications

A. Mining behaviors from user-generated data

- [TOIS10] Hady W. Lauw, Ee-Peng Lim, HweeHwa Pang, and Teck-Tim Tan, "STEvent: Spatio-Temporal Event Model for Social Network Discovery," ACM Transactions on Information Systems, Vol 28, No. 3, Article 15 (2010).
- [TKDE08] Hady W. Lauw, Ee-Peng Lim, and Ke Wang, "Bias and Controversy in Evaluation Systems," IEEE Transactions on Knowledge and Data Engineering, Vol. 20, No. 11, 2008. (Earlier in KDD06.)
- [TWEB12] Hady W. Lauw, Ee-Peng Lim and Ke Wang, "Quality and Leniency in Online Collaborative Rating Systems," ACM Transactions on the Web, Vol 6, No. 1, 2012. (Earlier in SDM07.)

B. Mining user-generated data for improving search

- [TKDE12a] Tao Cheng, Hady W. Lauw, and Stelios Paparizos, "Entity Synonyms for Structured Web Search," IEEE Transactions on Knowledge and Data Engineering, Vol 24, No. 10, 2012. (Earlier in ICDE10.)
- [TKDE12b] Heasoo Hwang, Hady W. Lauw, Lise Getoor, and Alexandros Ntoulas, "Organizing User Search Histories," IEEE Transactions on Knowledge and Data Engineering, Vol 24, No. 5, 2012.

C. Mining preferences from user-generated data

- [TOIS16] Loc Do and Hady W. Lauw, "Probabilistic Models for Contextual Agreement in Preferences," ACM Transactions on Information Systems, Vol 34, No. 4, Article No. 21, Sep 2016. (Earlier in WWW14.)
- [VLDB15] Loc Do, Hady W. Lauw, and Ke Wang, "Mining Revenue-Maximizing Bundling Configuration," Proceedings of the VLDB Endowment, Vol 8, No. 5, 2015.
- [ACL15] Maksim Tkachenko and Hady W. Lauw, "A Convolution Kernel Approach to Identifying Comparisons in Text," Annual Meeting of the Association for Computational Linguistics, 2015.
- [TKDE17] Maksim Tkachenko and Hady W. Lauw, "Comparative Relation Generative Model," IEEE Transactions on Knowledge and Data Engineering, Vol 29, No. 4, Apr 2017. (Earlier in CIKM'14.)
- [TKDE15] Thanh-Son Nguyen, Hady W. Lauw, and Panayiotis Tsaparas, "Review Selection Using Micro-Reviews," IEEE Transactions on Knowledge and Data Engineering, Vol 27, No. 4, 2015. (Earlier in CIKM13.)
- [DAMI17] Thanh-Son Nguyen, Hady W. Lauw, Panayiotis Tsaparas, "Micro-Review Synthesis for Multi-Entity Summarization," Data Mining and Knowledge Discovery, Vol 31, No. 5, pages 1189--1217, 2017.
- [WSDM15] Thanh-Son Nguyen, Hady W. Lauw, and Panayiotis Tsaparas, "Review Synthesis for Micro-Review Summarization," ACM International Conference on Web Search and Data Mining, 2015.
- [RecSys16] Trong T. Nguyen and Hady W. Lauw, "Representation Learning for Homophilic Preferences," ACM Conference on Recommender Systems, 2016.
- [CIKM17a] Trong T. Nguyen and Hady W. Lauw, "Collaborative Topic Regression with Denoising AutoEncoder for Content and Community Co-Representation," ACM Conference on Information and Knowledge Management, Nov 2017.
- [CIKM16] Maksim Tkachenko and Hady W. Lauw, "Plackett-Luce Regression Mixture Model for Heterogeneous Rankings," ACM Conference on Information and Knowledge Management, 2016.
- [MM17] Quoc-Tuan Truong and Hady W. Lauw, "Visual Sentiment Analysis for Review Images with Item-Oriented and User-Oriented CNN," ACM Multimedia, 2017.
- [MM20] Quoc-Tuan Truong, Hady W. Lauw, Martin Aumüller, and Naoko Nitta, "Reproducibility Companion Paper: Visual Sentiment Analysis for Review Images with Item-Oriented and User-Oriented CNN," ACM Multimedia Conference (ACM MM'20), Reproducibility Track, Oct 2020.
- [ECMLPKDD16] Duc-Trong Le, Yuan Fang, and Hady W. Lauw, "Modeling Sequential Preferences with Dynamic User and Context Factors," European Conference on Machine Learning and Principles and Practice of Knowledge Discovery, 2016.
- [IJCAI17b] Duc-Trong Le, Hady W. Lauw, and Yuan Fang, "Basket-Sensitive Personalized Item Recommendation," International Joint Conference on Artificial Intelligence (IJCAI'17), Aug 2017.
- [IJCAI18a] Duc-Trong Le, Hady W. Lauw, and Yuan Fang, "Modeling Contemporaneous Basket Sequences with Twin Networks for Next-Item Recommendation," International Joint Conference on Artificial Intelligence (IJCAI'18), Jul 2018.
- [IJCAI18b] Aghiles Salah and Hady W. Lauw, "A Bayesian Latent Variable Model of User Preferences with Item Context," International Joint Conference on Artificial Intelligence (IJCAI'18), Jul 2018.
- [UAI18] Aghiles Salah and Hady W. Lauw, "Probabilistic Collaborative Representation Learning for Personalized Item Recommendation," Conference on Uncertainty in Artificial Intelligence (UAI'18), 2018.

- [WSDM21b] Quoc-Tuan Truong, Aghiles Salah, and Hady W. Lauw, "Bilateral Variational Autoencoder for Collaborative Filtering," ACM International Conference on Web Search and Data Mining (WSDM'21), Mar 2021.
- [ML21] Quoc-Tuan Truong and Hady W. Lauw, "Variational Learning from Implicit Bandit Feedback," Machine Learning, Vol. 110, No. 8, pp. 2085-2105, 2021.
- [RecSys21] Aghiles Salah, Thanh-Binh Tran, Hady W. Lauw, "Towards Source-Aligned Variational Models for Cross-Domain Recommendation," ACM Conference on Recommender Systems (RecSys'21), Sep 2021.
- [IJCAI19a] Duc-Trong Le, Hady W. Lauw, and Yuan Fang, "Correlation-Sensitive Next-Basket Recommendation," International Joint Conference on Artificial Intelligence (IJCAI'19), Jul 2019.
- [JMLR20] Aghiles Salah, Quoc-Tuan Truong, and Hady W. Lauw, "Cornac: A Comparative Framework for Multimodal Recommender Systems," Journal of Machine Learning Research (JMLR), Machine Learning Open Source Software, Vol. 21, No. 95, 2020.
- [IC21] Quoc-Tuan Truong, Aghiles Salah, Thanh-Binh Tran, Jingyao Guo, Hady W. Lauw, "Exploring Cross-Modality Utilization in Recommender Systems," IEEE Internet Computing, Vol. 25, No. 4, Jul/Aug 2021, pp. 50-57.
- [IJCAI20] Trung-Hoang Le and Hady W. Lauw, “Synthesizing Aspect-Driven Recommendation Explanations from Reviews,” International Joint Conference on Artificial Intelligence (IJCAI'20), Jul 2020. **Distinguished Paper Award**.
- [WSDM21a] Trung-Hoang Le and Hady W. Lauw, "Explainable Recommendation with Comparative Constraints on Product Aspects," ACM International Conference on Web Search and Data Mining (WSDM'21), Mar 2021.
- [WWW22] Quoc-Tuan Truong, Tong Zhao, Chenghe Yuan, Jin Li, Jim Chan, Soo-Min Pantel and Hady W. Lauw,, "AmpSum: Adaptive Multiple-Product Summarization towards Improving Recommendation Explainability," The Web Conference 2022 (TheWebConf'22), Apr 2022.
- [WSDM22] Dung D. Le and Hady W. Lauw, "Collaborative Curating for Discovery and Expansion of Visual Clusters," ACM International Conference on Web Search and Data Mining (WSDM'22), Feb 2022.
- [KDD22a] Nhu-Thuat Tran and Hady W. Lauw, "Aligning Dual Disentangled User Representations from Ratings and Textual Content," ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'22), Aug 2022.

D. <u>Dimensionality reduction for user-generated data</u>

- [IJCAI17a] Tuan M. V. Le and Hady W. Lauw, "Semantic Visualization for Short Texts with Word Embeddings," International Joint Conference on Artificial Intelligence, 2017.
- [JAIR16] Tuan M. V. Le and Hady W. Lauw, " Semantic Visualization with Neighborhood Graph Regularization," Journal of Artificial Intelligence Research, 2016. (Earlier in AAAI14.)
- [KDD14] Tuan M. V. Le and Hady W. Lauw, "Semantic Visualization for Spherical Representation," ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2014.
- [ICDM14] Tuan M. V. Le and Hady W. Lauw, "Probabilistic Latent Document Network Embedding," IEEE International Conference on Data Mining, 2014.
- [IJCAI16] Tuan M. V. Le and Hady W. Lauw, “Word Clouds with Latent Variable Analysis for Visual Comparison of Documents,” International Joint Conference on Artificial Intelligence, 2016.
- [SDM16] Dung D. Le and Hady W. Lauw, "Euclidean Co-Embedding of Ordinal Data for Multi-Type Visualization," SIAM International Conference on Data Mining, 2016.
- [CIKM18] Dung D. Le and Hady W. Lauw, "Multiperspective Graph-Theoretic Similarity Measure," ACM Conference on Information and Knowledge Management (CIKM'18), Oct 2018.
- [IJCAI19b] Dung D. Le and Hady W. Lauw, "Learning Multiple Maps from Conditional Ordinal Triplets," International Joint Conference on Artificial Intelligence (IJCAI'19), Jul 2019.
- [CIKM17b] Dung D. Le and Hady W. Lauw, "Indexable Bayesian Personalized Ranking for Efficient Top-k Recommendation," ACM Conference on Information and Knowledge Management, 2017.
- [AAAI20b] Dung D. Le and Hady W. Lauw, "Stochastically Robust Personalized Ranking for LSH Recommendation Retrieval," AAAI Conference on Artificial Intelligence (AAAI'20), Jan 2020.
- [JAIR21] Dung D. Le and Hady W. Lauw, "Efficient Retrieval of Matrix Factorization-Based Top-k Recommendations: A Survey of Recent Approaches," Journal of Artificial Intelligence Research (JAIR), Vol. 70, 2021. To be presented at the IJCAI 2021 Journal Track.
- [ACL18] Maksim Tkachenko, Chong Cher Chia, and Hady W. Lauw, "Searching for the X-Factor: Exploring Corpus Subjectivity for Word Embeddings," Annual Meeting of the Association for Computational Linguistics (ACL’18), Jul 2018.
- [WWW19] Quoc-Tuan Truong and Hady W. Lauw, "Multimodal Review Generation for Recommender Systems," The Web Conference (WWW'19), May 2019.
- [AAAI19a] Quoc-Tuan Truong and Hady W. Lauw, "VistaNet: Visual Aspect Attention Network for Multimodal Sentiment Analysis," AAAI Conference on Artificial Intelligence (AAAI'19), Jan 2019.
- [ECIR21] Quoc-Tuan Truong and Hady W. Lauw, "Sentiment-Oriented Metric Learning for Text-to-Image Retrieval", European Conference on Information Retrieval (ECIR'21), Mar 2021.

- [AAAI19b] Maksim Tkachenko and Hady W. Lauw, "CompareLDA: A Topic Model for Document Comparison," AAAI Conference on Artificial Intelligence (AAAI'19), Jan 2019.
- [AAAI20a] Ce Zhang and Hady W. Lauw, "Topic Modeling on Document Networks with Adjacent-Encoder," AAAI Conference on Artificial Intelligence (AAAI'20), Jan 2020.
- [ECMLPKDD21a] Delvin Ce Zhang and Hady W. Lauw, "Semi-Supervised Semantic Visualization for Networked Documents," European Conference on Machine Learning and Principles and Practice of Knowledge Discovery (ECMLPKDD'21), Sep 2021.
- [ECMLPKDD21b] Delvin Ce Zhang and Hady W. Lauw, "Representation Learning on Multi-Layered Heterogeneous Network," European Conference on Machine Learning and Principles and Practice of Knowledge Discovery (ECMLPKDD'21), Sep 2021.
- [CIKM21] Delvin Ce Zhang and Hady W. Lauw, "Topic Modeling for Multi-Aspect Listwise Comparison," ACM International Conference on Information and Knowledge Management (CIKM'21), Nov 2021.
- [ICML22] Delvin Ce Zhang and Hady W. Lauw, "Dynamic Topic Models for Temporal Document Networks," International Conference on Machine Learning (ICML'22), July 2022.
- [KDD22b] Delvin Ce Zhang and Hady W. Lauw, "Variational Graph Author Topic Modeling," ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'22), Aug 2022.
- [NeurIPS22] Delvin Ce Zhang and Hady W. Lauw, "Meta-Complementing the Semantics of Short Texts in Neural Topic Models," Conference on Neural Information Processing Systems (NeurIPS'22), Nov 2022.
- [EMNLP22] Jia Peng Lim and Hady W. Lauw, "Towards Reinterpreting Neural Topic Models via Composite Activations," Conference on Empirical Methods in Natural Language Processing (EMNLP'22), Dec 2022.
- [WSDM23] Quoc-Tuan Truong and Hady W. Lauw, Concept-Oriented Transformers for Visual Sentiment Analysis, ACM International Conference on Web Search and Data Mining (WSDM'23), Feb 2023.
- [ACL23] Jia Peng Lim and Hady W. Lauw, Large-Scale Correlation Analysis of Automated Metrics for Topic Models, Annual Meeting of the Association for Computational Linguistics (ACL'23), Jul 2023.
- [EMNLP23] Jia Peng Lim and Hady W. Lauw, Disentangling Transformer Language Models as Superposed Topic Models, Conference on Empirical Methods in Natural Language Processing (EMNLP'23), Dec 2023.
- [KDD23] Delvin Ce Zhang, Rex Ying, Hady W. Lauw, Hyperbolic Graph Topic Modeling Network with Continuously Updated Topic Tree, ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'23), Aug 2023.
- [ICLR24] Nhu-Thuat Tran and Hady W. Lauw, Learning Multi-Faceted Prototypical User Interests, International Conference on Learning Representations (ICLR'24), Jul 2024.
- [TIST24] Trung-Hoang Le and Hady W. Lauw, Question-Attentive Review-Level Explanation for Neural Rating Regression, ACM Transactions on Intelligent Systems and Technology, Dec 2024.

These papers, as well as my other publications, can be downloaded from: https://www.hadylauw.com/publications.