CURRICULUM VITAE

MD ZIA ULLAH

Institut de Recherche en Informatique de Toulouse (IRIT)
Centre national de la recherche scientifique (CNRS)
118 Route de Narbonne, F-31062 CEDEX 9
Toulouse, France

✉: mdzia.ullah@irit.fr, 🌐: zuacubd.github.io, �
: github/zuacubd
🐦: @zuacsea 📞: +33 (06) 35 92 86 62

RESEARCH INTERESTS

I am interested in theoretical and practical information retrieval (IR) problems, as well as applied machine learning for large-scale predictive problems. My research has focused on adaptive information retrieval, query performance prediction, search intent mining and result diversification, check-worthy claim prediction, bipartite graph-based ranking, and deep learning for IR. I am always interested in expanding my research topics in other related areas. I do research also on natural language processing and social media analysis.

TEACHING INTERESTS

I am motivated to teach the following graduate- and undergraduate-level courses of Computer Science discipline:

- **Graduate-level courses:** Information retrieval (IR), Data Analytics, Network analysis, Natural language processing (NLP), Applied machine learning, Deep learning, Visualization, Web Application, etc.
- **Undergraduate-level courses:** Algorithms, Data structures, Object oriented programming (Java), Python programming, C/C++ programming, Graph theory, Competitive programming, Database theory, etc.

EDUCATION

2013/10 – 2016/12	**Ph.D.** Computer Science and Engineering, Toyohashi University of Technology, Toyohashi, Aichi, Japan Advisor: Prof. Masaki Aono *Thesis: "Bipartite Graph-based Ranking Methods for Subtopic Mining and Genetic Disease Prediction"*
2011/10 – 2013/09	**M.Engg.** Computer Science and Engineering, Toyohashi University of Technology, Toyohashi, Aichi, Japan Advisor: Prof. Masaki Aono *Thesis: "Estimating a Ranked List of Diseases by Associating Two Sets of Bipartite Graphs"*
2005/05 – 2010/03	**B.Sc. (Honours)** Computer Science and Engineering, University of Chittagong (CU) Chittagong, Bangladesh

APPOINTMENTS

2019/09 – 2021/12	**Postdoctoral Researcher** Institut de Recherche en Informatique de Toulouse (IRIT) Centre national de la recherche scientifique (CNRS) Toulouse, France Mentor: Professor Josiane Mothe

2017/04 – 2019/08	**Postdoctoral Researcher**
	Institut de Recherche en Informatique de Toulouse (IRIT)
	Universite Toulouse 3 Paul Sabatier (UT3)
	Toulouse, France
	Mentor: Professor Josiane Mothe
2017/01 – 2017/03	**Assistant Professor**
	Department of Computer Science and Engineering
	Port City International University (PCIU)
	Chittagong, Bangladesh
2016/10 – 2016/12	**Research Assistant**
	Department of Computer Science and Engineering
	Toyohashi University of Technology (TUT)
	Chittagong, Bangladesh
2011/06 – 2011/09	**Lecturer**
	Department of Computer Science and Engineering
	Bangladesh University of Business and Technology (BUBT)
	Dhaka, Bangladesh
2011/04 – 2011/05	**Software Engineer**
	Proggasoft Limited
	Dhaka, Bangladesh

RESEARCH AREAS

General Areas	Information retrieval, Natural language processing, Applied machine learning, Deep learning
Specific Topics	Adaptive information retrieval, Query performance prediction, Query expansion, Social media analysis, Search intent mining and diversification, Learning to rank, Bipartite graph-based ranking, Fairness in ranking, Conversational IR.

TECHNICAL SKILLS

Programming Languages	Python, Java, R, Ruby, C/C++, PHP, AWK
Deep Learning Frameworks	TensorFlow, PyTorch, Caffe
Machine Learning Packages	RankLib, SVM-rank, Scikit-learn
Search Engine Tools	Indri IR, Terrier IR, Lucene, Elastic-Search
Web Apps (back and front-end)	Python Flask, RESTful, RabbitMQ, Java Script, CSS
DevOps	Docker, Singularity

TEACHING EXPERIENCE

MASTERS LEVEL	**Data Analytics**	**2017/09 – Present**
	Master 2IS - Innovative Information Systems	
	Université Toulouse I – Capitole	
	Toulouse, France	

- *Information retrieval*
- *Data analysis*
- *Python programming*
- *R programming*

BACHELOR LEVEL	**Principles of Operating systems**	**2018/09 – 2019/02**
	IUT Blagnac	
	Université Toulouse II – Jean Jaures	
	Toulouse, France	

- *Operating systems (OS)*
- *Python programming for OS*

Computer Science Programs   **2017/01 – 2017/03**
Port City International University
Chittagong, Bangladesh

- *Object-oriented programming (JAVA)*
- *C programming and computer fundamentals*
- *Competitive programming (Algorithms and Data structures)*

Computer Science Programs   **2011/06 – 2011/09**
Bangladesh University of Business and Technology
Dhaka, Bangladesh

- *Computer graphics*
- *Data communications*
- *Computer fundamentals*

PEDAGOGICAL TRAINING

Enhancing Learning	Université Toulouse I – Capitole Toulouse, France	**2020/06/29 – 07/01**
The Challenge of Internationalization	Université Toulouse I – Capitole Toulouse, France	**2019/06/26 – 06/28**
Why Teach?	Université Toulouse I – Capitole Toulouse, France	**2019/12/17 – 12/19**

MENTORING

Ph.D. Student	Abu Nowshed Chy, TUT, Japan	**2016/04 – 2019/03**
M.Engg. Student	Ismat Ara Reshma, TUT, Japan	**2012/10 – 2014/09**
	Abu Nowshed Chy, TUT, Japan	**2014/04 – 2016/03**
	Sun Yiwen, TUT, Japan	**2016/10 – 2016/12**

Guiding students to conduct original research (i.e., theory and coding), writing paper for conferences/journals and preparing interactive presentations/posters.

SCHOLARSHIPS

- **Japanese Student Services Organization (JASSO) Scholarship** awarded from April 2016 – December 2016 for demonstrating excellent academic and personal records at University.
- **Amano Institute of Technology Scholarship** awarded from April 2015 – March 2016 for achieving excellent publications as a 2nd year Ph.D. student.
- **Japanese Student Services Organization (JASSO) Scholarship** awarded from April 2014 – March 2015 for demonstrating excellent academic and personal records at University.
- **Japanese Ministry of Education, Culture, Sports, Science, and Technology (MEXT) Scholarship** awarded from October 2011 – September 2013 for achieving outstanding CGPA in undergraduate study.

AWARDS

- **Best Presentation Award** at the 12th Asia Information Retrieval Societies Conference (AIRS 2016).
- **SIGIR Student Travel Grant** for the 12th Asia Information Retrieval Societies Conference (AIRS 2016).
- **Best Paper Award** at International Conference on Advanced Informatics: Concepts, Theory and Applications, ICAICTA 2015, Chomburi, Thailand, August 19-22th, 2015.
- **Best Paper Award** at International Conference on Advanced Informatics: Concepts, Theory and Applications, ICAICTA 2014, Bandung, Indonesia, August 20-21th, 2014.
- **Ranked 1st in the ImageCLEF 2014** Scalable Concept Image Annotation Competition, (Large-scale noisy web image dataset), September 15-18th, 2014.

- **Awarded in the ACM-ICPC Dhaka Site** Regional Programming Contest held at East-West University, December 8, Dhaka, Bangladesh, 2007. **Ranked 12th (over 83 teams)**

RESEARCH GRANTS

- Hori Foundation Research Grant (**$ 5,000**) for my project **Search Intent Mining** from April 2014 – September 2016.

PROFESSIONAL ACTIVITIES

Program Committee Members:

SIGIR	2021, 2020, 2019
CIKM	2021, 2020
WSDM	2022, 2021
KDD	2021
ECIR	2021, 2020, 2019
CLEF	2021, 2020, 2019, 2018, 2017
Others	KSE2019, KES2020, ICRIIS2017

Journal Reviewers:

- Information Retrieval Journal.
- Information Processing and Management.
- Open Computer Science Journal.
- GigaScience Journal.
- Jordanian Journal of Computers and Information Technology.

Organizing Committee Chair:

- Joint Conference of the Information Retrieval Communities in Europe (CIRCLE 2020), Toulouse, France.
- Workshop on Machine Learning for Trend and Weak Signal Detection in Social Networks and Social Media (MTWD2020), Toulouse, France.
- 2nd International Conference on Advanced Information and Communication Technology (ICAICT2020), Dhaka, Bangladesh.

PROFESSIONAL MEMBERSHIPS

- ACM Special Interest Group on Information Retrieval (SIGIR)   **2018 – Present**
- ACM Special Interest Group on Applied Computing (SIGAPP)   **2017 – 2018**
- IEEE Engineering in Medicine and Biology Society (EMBS)   **2012 – 2014**

EU PROJECT ACTIVITY

H2020 PREVISION (Prediction and Visual Intelligence for Security Information)   **09/2019 – 12/2021**

- Leader of Task 3.2 in WP3: Smart fusion and incomplete data handling.
- Involvement in Work packages from WP3 – WP8.
- Working on Data fusion, Machine learning, Linguistic analysis, Predictive analytics, digital security, Knowledge-base, Tools Deployments, Testing, Deliverable writing, and Dissemination.

CURRENT PROJECTS

Query Performance Prediction

In information retrieval (IR), query performance prediction (QPP) aims at predicting the effectiveness of a system for a given search query without resorting to relevance judgments. QPP can be used to inform an IR system whether a given search query would be effective or not, allowing the system to process it differently. Concerning an ineffective search query, the system could either apply a query reformulation or engage in an interactive session with the user (i.e., conversational IR) to understand the search intent and provide a better search result. Developing an effective QPP predictor is a challenging topic. Existing QPPs use pre-retrieval features or post-retrieval features (i.e., using the retrieved documents) or a machine-learned combination of those features. In this project, we propose to develop new QPP features and learn an explainable machine-learning model to select and combine the pre-and post-retrieval QPP features to predict the effectiveness. We have introduced summarized LETOR features as QPP features and combined them using a transparent machine learning (ML) model to predict the system effectiveness. We have also developed several effective Post-retrieval QPP predictors. The experiments on standard TREC benchmarks show that our proposed approach outperforms the known related QPP methods.

- **Md Zia Ullah et al.**, Forward and backward feature selection for query performance prediction, *The 35th ACM/SIGAPP Symposium On Applied Computing (SAC)*, 2020.
- **Md Zia Ullah et al.**, Query Performance Prediction Focused on Summarized Letor Features, *The 41st International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 2018
- **Md Zia Ullah et al.**, Query Performance Prediction and Effectiveness Evaluation Without Relevance Judgments: Two Sides of the Same Coin, *The 41st International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 2018.

Adaptive Information Retrieval

Modern information retrieval (IR) systems have become more and more complex, involving a large number of components and hyper-parameters. An IR system may choose from a set of possible retrieval models (e.g., BM25, Language model, etc.) or various query expansion parameters (e.g., expansion model, number of expansion terms, etc.), whose values greatly influence the overall retrieval effectiveness. Traditionally, these parameters are set globally at a system level based on training queries using Grid search or random search. However, a global configuration of parameters for an IR system may not be effective for all search queries having diverse characteristics (e.g., board or ambiguous query). In this project, we propose to develop an adaptive IR system for treating individual search queries on a per-query basis and improved retrieval effectiveness. The proposed adaptive IR system predicts the best system configuration according to the query representation. We cast this problem as a ranking of different system configurations and choose the best one using a learning-to-rank technique. The experiments on standard ad hoc TREC benchmarks show that this approach can significantly outperform the traditional method (i.e., Grid search) of global tuning of the system and known related methods. In a further study, we found that having all system configurations in the pool to select the best one could not fit some queries. We propose choosing a reduced set of system configurations using Risk-sensitive criteria. There is still some room for improvement in this project. An improved query representation using query performance predictors or contextualized language model (e.g., Transformer) could increase retrieval effectiveness. A clustering of queries and affinity propagation could replace the ranking of system configurations to improve the system effectiveness.

- Josiane Mothe and **Md Zia Ullah**, Defining an Optimal Configuration Set for Selective Search Strategy – A Risk-Sensitive Approach, *The 30th ACM International Conference on Information and Knowledge Management (CIKM)*, 2021.
- Josiane Mothe and **Md Zia Ullah**, Information Retrieval Device and Method using a Set of Search Configurations Pre-selected using Efficiency and Risk Functions, *European patent*, 19305984.7, 2019.
- **Md Zia Ullah et al.**, Learning to Adaptively Rank Document Retrieval System Configurations, *ACM Transactions of Information Systems (ACM TOIS)*, 2019.

Statistical Analysis of Information Retrieval System Components

Search engines involve various components and hyper-parameters whose values greatly influence the overall retrieval effectiveness while treating individual queries. When handling a new collection, finding the most appropriate system component for tuning as the most priority is a challenge since many factors influence the system effectiveness, such as the system components, internal parameters, and document collection. In this project, the objective is to analyze which modules and parameters influence system effectiveness the most. We have performed this statistical analysis at various levels based on several data analysis methods; each method is appropriate to reveal different aspects of the problem. We use an analysis of variance (ANOVA) to explore the components that statistically significantly influence the effectiveness, classification and regression tree (CART) to model the impact of the different component modalities and data visualization. We are currently focusing on analyzing these IR components on a large scale for various collections.

- **Md Zia Ullah**, Studying the Variability of System Setting Effectiveness by Data Analytics and Visualization, *International Conference of the Cross-Language Evaluation Forum for European Languages (CLEF)*, 2019.

Check-worthy Claim Prediction

Social media eases information spreading, makes information diffusion quicker, and reaches potentially more people than traditional media, in many cases regardless of the information quality. Automatic fact-checking could be a solution to warn social media users or used to stop spreading fake news. In this project, our objective is to predict the check-worthy claims for prioritizing fack-checking. We represent the claims using heterogeneous features based on Information nutritional-based labels, word embedding, and linguistics. To predict the check-worthy claims, we learn a supervised classification model using the proposed representation. The experimental results on standard benchmark datasets show that our approach improves the performance compared to the known related methods. Currently, we work on extracting context-aware features using graph-based multi-modal information and performing a feature selection thorough a deeper analysis.

- **Md Zia Ullah**, Information Nutritional Label and Word Embedding to Estimate Information Check-Worthiness, *The 42nd ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 2019.

Histopathological Image-based Cancer Detection

Germinal center (GC) and non-germinal center (NGC) are two different types of lymphomas (cancer), and classification of these types are important in cancer diagnosis. Several algorithms were developed based on gene expression and some biological assumptions that are too costly to obtain. In this project, the objective is to classify GC vs NGC types from a whole-slide histopathology image. In this regard, we propose to classify the whole-slide image (WSI) by leveraging the region of interest (ROI) based on whole-slide annotated images. First, we consider the image patch from the bounding box of the annotated slide at a specific resolution level. Then, we transform the extracted annotated image patch from the annotation slide to the original slide at a higher resolution level and extract the regular patches from the original whole-slide image. For GC vs NGC classification, we consider multiple Convolutional neural networks (CNN), including GoogleNet, AlexNet, VGGNet, and ResNet, to be enabled diverse learning from the patches. We also ensemble the prediction probabilities of multiple deep models to estimate an accurate prediction. We evaluate our approach using a recent WSI dataset from Oncopole Toulouse. Currently, we are focusing on combining the prediction of classifiers learned from higher and lower resolution levels of images.

Jargon Detection and its Application to Cyber-Security

Jargon detection aims at extracting meaningful key-phrases from text documents where the key-phrases provide a compact representation of document's content. The extracted keyphrases can serve as seeds in various potential applications, such as building a contextual lexicon, filtering documents, clustering text documents, identifying possibly related documents or scoring the text documents using the lexicon. A key-phrase is a sequence of one or more words (e.g., n-grams). The extraction pipeline is comprised of several steps such as pre-processing of the text (e.g., Parsing, POS tagging, Stop-word removal, Stemming or Lemmatization), candidate key-phrase extraction, and scoring and ranking the candidates. In this project, our objective is to develop tools to help the extract of key-phrases and the creation/updating of specific and topic-oriented lexicon by selecting key-phrases. The contextualized lexicon is then used to score and rank risk-sensitive documents in Cyberspace.

PAST PROJECTS

Search Intent Mining and Result Diversification

With an information need in mind, the user usually formulates a query and issues it to a search engine. The search engine responds with an ad hoc ranked list of documents/answers to fulfill the information need. Web search queries are usually vague, ambiguous, or tend to have multiple facets. Users may have different search intents while issuing the same search query (e.g., the word -address- may mean the place someone lives or a formal speech delivered to an audience). If the issued search query conveys a variety of interpretations, the ad hoc search result may be far from "what the user wants to have." In this project, our objectives are to understand the search query by exploring possible search intents (i.e., query understanding) and rerank the search results accordingly (e.g., diversification). To understand the search query, we have extracted the subtopics covering the dynamic search intents of the users. We have also developed a method for ranking the subtopics of the search query by exploiting the locally trained word-embedding based features, a bipartite graph-based ranking, and estimating the novelty of subtopics by combining contextual and categorical similarities. The experiments on two benchmark datasets on a large web corpus show that this approach significantly outperforms the known related works.

- **Md Zia Ullah** and Masaki Aono, A Bipartite Graph-based Ranking Approach to Query Subtopics Diversification Focused on Word Embedding Features, *IEICE Transactions on Information and Systems (IEICE TOIS)*, 2016.
- **Md Zia Ullah et al.**, Query Subtopic Mining Exploiting Word Embedding for Search Result Diversification, *Asia Information Retrieval Societies Conference (AIRS)*, 2016.
- **Md Zia Ullah** and Masaki Aono, Query Subtopic Mining for Search Result Diversification, *1st International Conference of Advanced Informatics: Concept, Theory and Application (ICAICTA)*, 2014.

Information Retrieval in the Microblog Sphere

Social network platforms are not only places for maintaining social relationships but also act as valuable information sources. Users often turn to social network platforms for sharing their personal views, experiences, and important news while also getting some information. Among several social network sites, Twitter is now the most popular, where users post short-text tweets or engage in interaction with a reply/mention whenever a notable event occurs. That is why Information retrieval in the Twitter-sphere has made a hit with many complacencies. In this project, our objective is to retrieve and re-rank the tweets incorporating the temporal dimension for a search query. Vocabulary mismatch and temporal burst are two challenging issues of short-text in microblog retrieval. To re-rank the short-text tweets, we have proposed efficient and effective learning to rank approach while addressing vocabulary mismatch and temporal burst issues based on content and context-aware features. We have experimented with, evaluated our proposed approach on two benchmark datasets, and shown better performance than the known related methods.

- **Md Zia Ullah et al.**, Query Expansion for Microblog Retrieval Focusing on an Ensemble of Features, *Journal of Information Processing (JIP)*, 2018.
- **Md Zia Ullah et al.**, Microblog Retrieval Using Ensemble of Feature Sets Through Supervised Feature Selection, *IEICE Transactions on Information and Systems (IEICE TOIS)*, 2017.

Multi-modal Information Fusion for Image Annotation

With the spread of various social network services, including Facebook, Twitter, Instagram, and Flickr, there has been an explosive growth of images on the Internet. Such collections of images could be leveraged in various potential applications such as recommending restaurant menus, tourist guides and entertainment facilities, traffic congestions, local weather information, crime investigation, and so on. There should have semantic links or annotations among these images to develop such an application. However, there is no such annotation for most of the image data on the Internet. Some description of the image is available on the Web page where the image appears; however, the relationship between the surrounding text and images is redundant and irrelevant, varies greatly. Despite the vast applicability of such multi-modal information in machine learning, it makes weak supervision, which is a challenging problem. In this project, our objective is to annotate images with semantically relevant keywords or concepts. We have proposed a new method based on ontology, graph structure immerged from multi-modal information, and machine learning to annotate images. First, we have developed an ontology-based approach to harvest training examples from the noisy-labelled images collected from the Web. Second, we have introduced a graph structure to model the context around the image and define a new kernel function to propagate ontology-based text and image features across multi-domain concepts. Third, we have an ensemble of voting strategies and probability estimates from multiple binary classifiers for tackling multi-class multi-label problems. We have experimented with, evaluated the proposed method using two benchmark datasets (the ImageCLEF 2013 and 2014), and shown that our approach outperforms the known related methods.

- **Md Zia Ullah et al.**, Ontology-based Classification for Multi-label Image Annotation, *1st International Conference of Advanced Informatics: Concept, Theory and Application (ICAICTA)*, 2014.
- **Md Zia Ullah et al.**, KDEVIR at ImageCLEF 2014 Scalable Concept Image Annotation Task: Ontology-based Automatic Image Annotation, *International Conference of the Cross-Language Evaluation Forum for European Languages (CLEF)*, 2014.

Bipartite graph and its application to Health informatics

With a vast amount of medical knowledge data available on the Internet, it is becoming increasingly practical to help doctors in clinical diagnostics by suggesting plausible diseases predicted by applying data and text mining technologies. Since genetic diseases are difficult to diagnose because of their diverse symptoms, patients are often misdiagnosed or experienced long diagnostic delays. In this project, our objective is to retrieve and rank possible genetic diseases linked through causative genes, given a set of clinical phenotypes. First, we have analyzed the human disease network (HDN) and protein-protein interaction (PPI) network to predict causal genes and explored the pathways from phenotype to genetic diseases through their causal genes. Second, we have associated two sets of bipartite graphs and introduced a weighting scheme to approximate the weight of the edge. We have experimented and evaluated our proposed method on publicly available datasets, and the result shows that our proposed approach outperforms the known related methods.

- **Md Zia Ullah et al.**, Estimating a Ranked List of Human Genetic Diseases by Associating Phenotype-Gene with Gene-Disease Bipartite Graphs, *ACM Transactions on Intelligent Systems and Technology (ACM TIST)*, 2015.
- **Md Zia Ullah et al.**, Estimating a Ranked List of Human Hereditary Diseases for Clinical Phenotypes by Using Weighted Bipartite Networks, *IEEE Engineering in Medicine and Biology Society (EMBS)*, 2013.

PATENTS

1 Josiane Mothe and **Md Zia Ullah**, Information Retrieval Device and Method using a Set of Search Configurations Pre-selected using Efficiency and Risk Functions, *European Patent*, 19305984.7, 2019.

BOOK CHAPTERS

1 **Md Zia Ullah et al.**, Comparison of machine learning models for -early- depression detection from users' posts, Springer, 2021 [In press]

PUBLICATIONS

Journal Articles (Peer-Reviewed and Published)

1 Deveaud R, Mothe J, **Ullah MZ**, Nie JY (2019), Learning to Adaptively Rank Document Retrieval System Configurations, ▶ *ACM Transactions of Information Systems (ACM TOIS)*, 41 pages, pp.3:1-3:41, Volume 37, Issue 1, doi: https://doi.org/10.1145/3231937. [Google Scholar citations: 12].

2 Chy AN, **Ullah MZ**, and Aono M (2018), Query Expansion for Microblog Retrieval Focusing on an Ensemble of Features, ▶ *Journal of Information Processing*, 16 pages, pp. 61-76, Volume 27, Issue 4, 2, doi: https://doi.org/10.2197/ipsjjip.27.61. [Google Scholar citations: 7].

3 Chy AN, **Ullah MZ**, and Aono M (2017), Microblog Retrieval Using Ensemble of Feature Sets through Supervised Feature Selection, ▶ *IEICE Transactions on Information and Systems (IEICE TOIS)*, 13 pages, pp. 793-806, Volume E100-D, Issue 4, 2, doi: https://doi.org/0.1587/transinf.2016DAP0032. [Google Scholar citations: 7].

4 **Ullah MZ**, and Aono M (2016), A Bipartite Graph-based Ranking Approach to Query Subtopics Diversification Focused on Word Embedding Features, ▶ *IEICE Transactions on Information and Systems (IEICE TOIS)*, 11 pages, pp. 3090-3100, Volume E99-D, Issue 12, doi: https://doi.org/10.1587/transinf.2016EDP7190. [Google Scholar citations: 2].

5 **Ullah MZ**, and Aono M, and Seddiqui MH (2015), Estimating a Ranked List of Human Genetic Diseases by Associating Phenotype-Gene with Gene-Disease Bipartite Graphs, ▶ *ACM Transactions on Intelligent System and Technology (ACM TIST)*, 21 pages, pp.56:1 56:21, Volume 6, Issue 4, doi: https://doi.org/10.1145/2700487. [Google Scholar citations: 5].

International Conference Papers (Peer-Reviewed and Published)

1 Mothe J, **Ullah MZ** (2021, November), Defining an Optimal Configuration Set for Selective Search Strategy – A Risk-Sensitive Approach, ▶ *Proceedings of the 30th ACM International Conference on Information and Knowledge Management (CIKM 2021)*, Virtual Event, QLD, Australia, doi: https://doi.org/10.1145/3459637.3482422. **Rank: A**

2 Chifu AG, Mothe J, **Ullah MZ** (2020, March), Fair Exposure of Documents in Information Retrieval: a Community Detection Approach, ▶ *Joint Conference of the Information Retrieval Communities in Europe (CIRCLE 2020)*, Toulouse, France.

3 Demestichas K, Hoang TBN, Mothe J, Teste O, and **Ullah MZ** (2020, March), Prediction and Visual Intelligence for Security Information: The Prevision H2020 Project, ▶ *Joint Conference of the Information Retrieval Communities in Europe (CIRCLE 2020)*, Toulouse, France.

4 Déjean S, Ionescu RT, Mothe J, **Ullah MZ** (2020, March), Forward and Backward Feature Selection for Query Performance Prediction, ▶ *Proceedings of the 35th Annual ACM Symposium on Applied Computing (ACM SAC)*, pp. 690-697, Brno, Czech Republic. doi: https://doi.org/10.1145/3341105.3373904 [Google Scholar citations: 4] **Rank: B**

5 Déjean S, Mothe J, **Ullah MZ** (2019, September), Studying the Variability of System Setting Effectiveness by Data Analytics and Visualization, ▶ *International Conference of the Cross-Language Evaluation Forum for European Languages (CLEF2019)*, pp. 62-74, Lugano, Switzerland. doi: `https://doi.org/10.1007/978-3-030-28577-7_3`. [Google Scholar citations: 1]

6 Lespagnol C, Mothe J, **Ullah MZ** (2019, July), Information Nutritional Label and Word Embedding to Estimate Information Check-Worthiness, ▶ *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2019)*, pp. 941–944, Paris, France. doi: `https://doi.org/10.1007/978-3-030-28577-7_3`. [Google Scholar citations: 9] **Rank: A***

7 Mizzaro S, Mothe J, Roitero K, and **Ullah MZ** (2018, July), Query Performance Prediction and Effectiveness Evaluation Without Relevance Judgments: Two Sides of the Same Coin, ▶ *The 41st International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2018)*, pp. 1233-1236, Ann Arbor, MI, USA. doi: `https://doi.org/10.1145/3209978.3210146`. [Google Scholar citations: 9] **Rank: A***

8 Chifu AG, Laporte L, Mothe J, and **Ullah MZ** (2018, July), Query Performance Prediction Focused on Summarized Letor Features, ▶ *The 41st International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2018)*, pp. 1177-1180, Ann Arbor, MI, USA. doi: `https://doi.org/10.1145/3209978.3210121`. [Google Scholar citations: 10] **Rank: A***

9 Molina S, Mothe J, Roques D, Tanguy L, and **Ullah MZ** (2017, September), IRIT QFR: IRIT Query Feature Resource, ▶ *International Conference of the Cross-Language Evaluation Forum for European Languages (CLEF2017)*, pp. 69-81, Dublin, Ireland. doi: `https://doi.org/10.1007/978-3-319-65813-1_6`. [Google Scholar citations: 4]

10 **Ullah MZ**, Chy AN, and Aono M (2016, November), Query Subtopic Mining Exploiting Word Embedding for Search Result Diversification, ▶ *Twelfth Asia Information Retrieval Societies Conference(AIRS 2016)*, pp. 308-314, Beijing, China. doi: `https://doi.org/10.1007/978-3-319-48051-0_24`. [Google Scholar citations: 11] **Best Presentation Award Rank: C**

11 Chy AN, **Ullah MZ**, and Aono M (2015, August), Combining Temporal and Content-Aware Features for Microblog Retrieval, ▶ *IEEE International Conference on Advanced Informatics: Concepts, Theory and Applications (ICAICTA 2015)*, pp. 1-6, Chonburi, Thailand. doi: `https://doi.org/10.1109/ICAICTA.2015.7335353`. [Google Scholar citations: 3] **Best Paper Award**

12 **Ullah MZ** and Aono M (2014, August), Query Subtopic Mining for Search Result Diversification, ▶ *IEEE International Conference on Advanced Informatics: Concepts, Theory and Applications (ICAICTA 2014)*, pp. 309-314, Bandung, Indonesia. doi: `10.1109/ICAICTA.2014.7005960`. [Google Scholar citations: 6]

13 Reshma IA, **Ullah MZ**, and Aono M (2014, August), Ontology-based Classification for Multi-label Image Annotation, ▶ *IEEE International Conference on Advanced Informatics: Concepts, Theory and Applications (ICAICTA 2014)*, pp. 226-231, Bandung, Indonesia, doi: `10.1109/ICAICTA.2014.7005945`. [Google Scholar citations: 5] **Best Paper Award**

14 **Ullah MZ**, Aono M, and Seddiqui MH (2013, July), Estimating a Ranked List of Human Hereditary Diseases for Clinical Phenotypes by Using Weighted Bipartite Networks, ▶ *35th IEEE International Conference on Engineering in Medicine and Biology Society (EMBS)*, pp. 3475-3478, Osaka, Japan. doi: `10.1109/EMBC.2013.6610290` **Rank: C**

International Workshops

1 **Ullah MZ** and Aono M (2016, September), KDEIR at CLEF eHealth 2016: Health Documents Re-ranking Based on Query Variations, ▶ *CLEF 2016 Evaluation Labs and Workshop, Online Working Notes (CEUR)*, pp. 167-170, vol. 1609, Èvora, Portugal. [Google Scholar citations: 1]

2 **Ullah MZ** and Aono M (2016, June), KDEIM at NTCIR-12 IMine-2 Search Intent Mining Task: Query Understanding through Diversified Ranking of Subtopics, ▶ *Proceedings of the 12th NTCIR Conference on Evaluation of Information Access Technologies*, NTCIR, pp. 60–63, Tokyo, Japan. [Google Scholar citations: 1] **Ranked 1st position**

3 Abu Nowshed Chy, **Ullah MZ**, and Aono M (2016, June), KDETM at NTCIR-12 Temporalia Task: Combining a Rule-based Classifier with Weakly Supervised Learning for Temporal Intent Disambiguation, ▶ *Proceedings of the 12th NTCIR Conference on Evaluation of Information Access Technologies*, NTCIR, pp. 281-284, Tokyo, Japan. [Google Scholar citations: 2]

4 **Ullah MZ** and Masaki Aono (2015, September), KDEVIR at ImageCLEF 2015 Scalable Image Annotation, Localization, and Sentence Generation task: Ontology based Multi-label Image Annotation, ▶ *CLEF 2015 Evaluation Labs and Workshop, Online Working Notes (CEUR)*, 6 pages, Toulouse, France. [Google Scholar citations: 1]

5   Abu Nowshed Chy, **Ullah MZ**, and Masaki Aono (2015, May), A Time and Context-Aware Re-ranker for Microblog Retrieval, ▶ *The 29th Annual Conference of the Japanese Society for Artificial Intelligence*, Hokkaido, Japan. [Google Scholar citations: 1]

6   **Ullah MZ** and Masaki Aono (2015, December), SEM13 at the NTCIR-11 IMINE Task: Subtopic Mining and Document Ranking Subtasks, ▶ *Proceedings of the 11th NTCIR Conference on Evaluation of Information Access Technologies*, NTCIR, pp. 64-68, Japan. [Google Scholar citations: 0]

7   Ismat Ara Reshma, **Ullah MZ**, and Masaki Aono (2014, September), KDEVIR at ImageCLEF 2014 Scalable Concept Image Annotation Task: Ontology-based Automatic Image Annotation, ▶ *CLEF 2014 Evaluation Labs and Workshop, Online Working Notes (CEUR)*, CLEF 2014, 9 pages, Sheffield, UK. [Google Scholar citations: 7] **Ranked 1st position**

8   Ismat Ara Reshma, **Ullah MZ**, Masaki Aono (2014, August), Ontology based Supervised Learning for Image Annotation, ▶ *The Institute of Electronics, Information and Communication Engineers (IEICE), Image Engineering (IE)*, IE2014-22-IE2014-29, Technical Report, Vol. 114, No.172, PP. 41-56, Chiba, Japan. [Google Scholar citations: 0]

9   Ismat Ara Reshma, **Ullah MZ**, and Masaki Aono (2013, September), KDEVIR at ImageCLEF 2013 Image Annotation Task, ▶ *CLEF 2013 Evaluation Labs and Workshop, Online Working Notes (CEUR)*, CLEF 2013, Valencia, Spain. [Google Scholar citations: 4]

10   **Ullah MZ**, Masaki Aono, and Md Hanif Seddiqui (2012, June), SEM12 at the NTCIR-10 INTENT-2 English Subtopic Mining Subtask, ▶ *Proceedings of the 10th NTCIR Conference on Evaluation of Information Access Technologies*, NTCIR, Tokyo, Japan. [Google Scholar citations: 7]

TALKS

○ Presentation on **Introduction to Convolutional Neural Network & Different Architectures**, *H2020 InnEO Space PhD Summer School*, **2021**. (Oral)

○ Presentation on **Learning to Adaptively Rank Document Retrieval System Configurations**, *ACM SIGIR*, **2020**. (Oral)

○ Presentation on **Fair Exposure of Documents in Information Retrieval: a Community Detection Approach**, *CIRCLE*, **2020**. (Oral)

○ Presentation on **Forward and Backward Feature Selection for Query Performance Prediction**, *ACM SAC*, **2020**. (Oral)

○ Presentation on **Information Nutritional Label to Predict Information Check-worthiness**, *MTWD*, **2020**. (Oral)

○ Presentation on **Information Nutritional Label and Word Embedding to Estimate Information Check-Worthiness**, *ACM SIGIR*, **2019**. (Poster)

○ Presentation on **Studying the Variability of System Setting Effectiveness by Data Analytics and Visualization**, *CLEF*, **2019**. (Oral)

○ Presentation on **Query Performance Prediction Focused on Summarized LETOR Features**, *SIGIR*, **2018**. (Poster)

○ Presentation on **Query Performance Prediction and Effectiveness Evaluation Without Relevance Judgments: Two Sides of the Same Coin**, *SIGIR*, **2018**. (Poster)

○ Presentation on **IRIT-QFR: Query Feature Resource**, *CLEF*, **2017**. (Oral)

○ Presentation on **Query Subtopic Mining Exploiting Word Embedding for Search Result Diversification**, *AIRS*, **2016**. (Poster)

○ Presentation on **KDEIM at NTCIR-12 IMINE-2: Intent Mining Through Diversified Ranking of Subtopics**, *NTCIR*, **2016**. (Poster)

○ Presentation on **Query Subtopic Mining for Search Result Diversification**, *ICAICTA*, **2014**. (Oral)

○ Presentation on **Estimating a Ranked List of Human Hereditary Diseases for Clinical Phenotypes by Using Weighted Bipartite Network**, *EMBS*, **2013**. (Poster)

○ Presentation on **SEM12 at the NTCIR-10 INTENT-2 English Subtopic Mining**, *NTCIR*, **2013**. (Poster)

LANGUAGE SKILLS

BENGALI	Mother's tongue
ENGLISH	Fluent
FRENCH	Speaking, listening, and writing (A2-1)
JAPANESE	Speaking, listening, and writing

REFEREES

- Josiane Mothe, Ph.D.
  Professor
  Institut de Recherche en Informatique de Toulouse (IRIT),
  ESPE, Uni. de. Toulouse,
  Toulouse, France
  Phone: +33(0)5 61 55 64 44
  Email: josiane.mothe@irit.fr
  Homepage: https://www.irit.fr/~Josiane.Mothe

- Masaki Aono, Ph.D.
  Professor
  Department of Computer Science & Engineering,
  Toyohashi University of Technology,
  Toyohashi, Aichi, Japan.
  TEL: 81+532-44-6764
  Email:aono@tut.jp
  Homepage: https://www.kde.cs.tut.ac.jp/~aono