Curriculum Vitae
Marinka Zitnik

Department of Biomedical Informatics

Harvard University

10 Shattuck Street

Boston, MA 02115

marinka@hms.harvard.edu

+1 (650)-308-6763

http://zitniklab.hms.harvard.edu

https://scholar.harvard.edu/marinka

Personal Statement

Artificial intelligence (AI) and machine learning (ML) hold tremendous promise in enabling scientific breakthroughs and discoveries in diverse areas. I investigate applied machine learning with a current focus on networked systems that require infusing structure and domain knowledge. This research is theoretically grounded and spans several areas of machine learning. My research strategy is to create foundational models (e.g., pre-trained models, self-supervised models, general-purpose models, multi-purpose models, and multi-modal models) that are trained on broad data at scale and can be adapted to a range of downstream tasks. This research opens up new avenues for understanding network and disease biology, developing safe and effective medicines, and can impact the way predictive modeling is performed today at the fundamental level.

My research realizes a scientific approach that consists of: (i) inventing ways to infuse domain knowledge and structure into complex and multi-scale datasets to reduce uncertainty and enable generalization to entirely new scenarios not seen during training; (ii) developing methods that produce actionable and trustworthy representations and can reason over massive datasets; and (iii) translating machine learning research into innovative applications.

My research contributions are exemplified by publications in top-tier machine learning and data mining venues; top-tier bioinformatics and computational biology venues; Best Paper and Research Awards from the International Society for Computational Biology (ISCB); International Conference on Machine Learning (ICML); Bayer Early Excellence in Science Award; Amazon Research Faculty Award; and two Rising Star Awards from MIT Electrical Engineering and Computer Science (EECS) and The Broad Institute of MIT and Harvard, being the only young scientist who received such recognition in both EECS and Biomedicine.

My algorithms and methods have had a tangible impact, which has garnered interests of government, academic, and industry researchers and has put new tools in the hands of practitioners. Along with the scientific advances, as a scientist and engineer, I also release open-source tools and AI-ready datasets that the research community can immediately use and build upon.

Academic Position

Harvard University, USA   December 2019-
Assistant Professor, Department of Biomedical Informatics
Associate Member, Broad Institute of MIT and Harvard
Faculty Affiliate, Harvard Data Science Initiative

Education

Stanford University, USA   2016-2019
Postdoctoral Research Scholar
Department of Computer Science
Advisor: Prof. Jure Leskovec
Biohub Postdoctoral Research Fellow

University of Ljubljana, Slovenia   2012-2015
Ph.D. in Computer Science
Department of Computer and Information Science
Advisor: Prof. Blaz Zupan
Committee: Peter Semrl, Igor Kononenko, Saso Dzeroski, Florian Markowetz
summa cum laude (GPA 10.00/10)
Jozef Stefan Golden Emblem

Stanford University, Department of Computer Science, USA   2014
     Research Fellow with Prof. Jure Leskovec

Baylor College of Medicine, Department of Molecular and Human Genetics, USA   2013-2014
     Predoctoral Fellow with Prof. Gad Shaulsky and Prof. Adam Kuspa

Imperial College London, Department of Computing, UK   2012
     Research Student with Prof. Natasa Przulj

University of Toronto, Donnelly Centre for Cellular and Biomolecular Research, Canada   2012
     Research Student with Prof. Charles Boone

University of Ljubljana, Slovenia   2008-2012
     B.Sc. in Computer Science and Mathematics
     Department of Computer and Information Science
     Department of Mathematics and Physics
     summa cum laude (GPA 10.00/10)
     University of Ljubljana Prešeren Award

Selected Honors

◇ **ELLIS Scholar**, elected by ELLIS Society, European Laboratory for Learning and Intelligent Systems   2021
◇ **Roche Alliance with Distinguished Scientists (ROADS) Award**   2021
◇ **AstraZeneca Research Award**   2021
◇ **Best Paper Award, Interpretable ML in Healthcare, Int'l Conf. on Machine Learning (ICML)**   2021
◇ **Best Poster Award, Computational Biology, Int'l Conf. on Machine Learning (ICML)**   2021
◇ **Amazon Research Faculty Award**   2021
◇ **Bayer Early Excellence in Science Award**   2020
◇ **Next Generation in Biomedicine**, 20 awardees worldwide, Broad Institute of MIT and Harvard   2018
◇ **Rising Star in Electrical Engineering and Computer Science**, 60 awardees in North America, MIT   2018
◇ **National Guest Scholar** at Stanford Clinical Excellence Research Center (CERC), invited to teach the CERC fellows about data science   2018
◇ **Jozef Stefan Golden Emblem**, Highest national award for Ph.D. dissertation in natural sciences, medicine, and biotechnology, a maximum of 3 prizes are awarded for theses in the last 3 years   2017
◇ **Nominated for the Best European Doctoral Dissertation in Artificial Intelligence** by the European Association for Artificial Intelligence   2016
◇ **National Award for Early Graduation from PhD**, 2 years and 10 months after my B.Sc. degree   2015
◇ **DREAM Olfaction Prediction Challenge**, Competition on predicting how a molecule smells from its physical and chemical features; rank: 3/19 and invited as a co-author on a *Science* paper   2015
◇ *University of Ljubljana Outstanding Research Award*, Highest research award given to 10 out of 40,000 researchers   2015
◇ **Best Poster Award** at the 11th Basel Computational Biology Conference (BC^2'15), awarded to 2 posters out of several hundred at the conference   2015
◇ *Google Anita Borg & Grace Hopper Fellowship*, 1 out of 20 winners worldwide   2014
◇ **Best Paper Award for Outstanding Research** in Critical Assessment of Massive Data Analysis at the 23rd International Conference on Intelligent Systems for Molecular Biology (ISMB)   2014
◇ **Heidelberg Laureate Forum**, Young researcher at the forum of Turing, Fields, Nevanlinna, and Abel Prize laureates; 1 out of 100 computer science graduate students worldwide   2014
◇ **Best Poster Award** at the 18th Annual International Conference on Research in Computational Molecular Biology (RECOMB'14), awarded to 2 posters out of several hundred at the conference   2014
◇ **Winner of the BioNLP Gene Regulation Network Challenge**, Competition on gene regulatory network inference; rank: 1/30 and invited to give a talk at the 51st Annual Meeting of the Association for Computational Linguistics (ACL'13)   2013
◇ **Best Paper Award for Outstanding Research** in Critical Assessment of Massive Data Analysis at the 22nd International Conference on Intelligent Systems for Molecular Biology (ISMB)   2013
◇ **Winner of the Biomedical Literature Search Competition**, Competition on classifying biomedical research literature; rank: 1/126 and invited to give a talk at the 7th Joint Rough Set Symposium (JRS'12)   2012

<table>
<tr><td>◇ EMC Israel Data Science Challenge, Competition on classifying programming code from open-source projects; rank: 2/91 and awarded a prize of $10,000</td><td>2012</td></tr>
<tr><td>◇ University of Ljubljana Prešeren Award, Highest university award for students given to 10 out of 40,000 students</td><td>2012</td></tr>
<tr><td>◇ University of Ljubljana Best Student Award, Highest research award given to 10 students nationwide</td><td>2011</td></tr>
<tr><td>◇ Google Anita Borg Scholarship, 1 out of 20 winners in Europe, Google</td><td>2011</td></tr>
<tr><td>◇ Outstanding Student Award, 4 times awarded as the best student in the class of 300 (based on average grade) at the University of Ljubljana</td><td>2009-2012</td></tr>
<tr><td>◇ Swiss Talent Forum, Young researcher at the Meeting of Nobel Laureates; 1 out of 60 students in Europe</td><td>2009</td></tr>
<tr><td>◇ 19th European Union Contest for Young Scientists, Prize for a chemical Strunz classification of the largest national collection of minerals and rocks</td><td>2007</td></tr>
<tr><td>◇ First-prize Winner of the Year for research on the impact of EU accession on technology, Hewlett-Packard</td><td>2007</td></tr>
</table>

Publications

Publications: https://scholar.google.com/citations?user=YtUDgPIAAAAJ&hl=en

◇ **Articles in Peer-Reviewed Journals**

[80] M. M. Li, K. Huang, **M. Zitnik**. Graph Representationa Learning in Biomedicine. *In Review*, 2021 [arXiv:2104.04883]

[79] D. Stupp, E. Sharon, I. Bloch, **M. Zitnik**, O. Zuk, Y. Tabach. Co-Evolution Based Machine-Learning for Predicting Functional Interactions Between Human Genes. *Nature Communications*, 12, 6454, 2021.

[78] X. Zhang, M. Sumathipala, **M. Zitnik**. Population-Scale Identification of Differential Adverse Events Before and During a Pandemic. *Nature Computational Science*, 1, 666–677, 2021.

[77] D. M. Gysi*, I. Do Valle*, **M. Zitnik***, A. Ameli*, X. Gan*, O. Varol, H. Sanchez, R.M. Baron, D. Ghiassian, J. Loscalzo, A.-L. Barabasi. Network Medicine Framework for Identifying Drug Repurposing Opportunities for COVID-19. *Proceedings of the National Academy of Sciences (PNAS)*, 118, 19:e2025581118, 2021.

[76] C. Ruiz, **M. Zitnik**, J. Leskovec. Identification of Disease Treatment Mechanisms through the Multiscale Interactome. *Nature Communications*, 12, 1796, 2021.

[75] S. Wang, A.O. Pisco, A. McGeever, M. Brbic, **M. Zitnik**, S. Darmanis, J. Leskovec, J. Karkanias, R.B. Altman. Leveraging the Cell Ontology to Classify Unseen Cell Types. *Nature Communications*, 12, 5556, 2021.

[74] J.J. Patten, P.T. Keiser, D. Gysi, G. Menichetti, H. Mori, C.J. Donahue, X. Gan, I. Do Valle, K. Geoghegan-Barek, M. Anantpadma, J. L. Berrigan, S. Jalloh, T. Ayazika, F. Wagner, **M. Zitnik**, S. Ayehunie, D. Anderson, J. Loscalzo, S. Gummuluru, M.N. Namchuk, A. L. Barabasi, R. A. Davey. Multidose Evaluation of 6,710 Drug Repurposing Library Identifies Potent SARS-CoV-2 Infection Inhibitors In Vitro and In Vivo, *In Review*, 2021 [bioRxiv: 2021.04.20.440626]

[73] T. Wanyan, H. Honarvar, S.K. Jaladanki, C. Zang, N. Naik, S. Somani, J.K. De Freitas, I. Paranjpe, A. Vaid, J. Zhang, R. Miotto, Z. Wang, G.N. Nadkarni, **M. Zitnik**, A. Azad, F. Wang, Y. Ding, B.S. Glicksberg. Contrastive Learning Improves Critical Event Prediction in COVID-19 Patients. *Patterns*, 2:12, 100389, 2021.

[72] S. N. Kunz, D. Helkey, **M. Zitnik**, C. S. Phibbs, J. Rigdon, J. A. F. Zupancic, J. Profit. Quantifying the Variation in Neonatal Transport Referral Patterns Using Network Analysis, *Journal of Perinatology*, s41372-021-01091, 2021.

[71] M. Brbic, **M. Zitnik**, S. Wang, A.O. Pisco, R.B. Altman, S. Darmanis, J. Leskovec. MARS: Discovering Novel Cell Types Across Heterogeneous Single-Cell Experiments. *Nature Methods*, 17, 1200–1206, 2020.

[70] K. Huang, T. Fu, L. Glass, **M. Zitnik**, C. Xiao, J. Sun. DeepPurpose: A Deep Learning Library for Drug-Target Interaction Prediction. *Bioinformatics*, 36, 22-23:5545–5547, 2020.

[69] K. Huang, C. Xiao, L. Glass, **M. Zitnik**, J. Sun. SkipGNN: Predicting Molecular Interactions with Skip-Graph Networks. *Scientific Reports*, 10, 21092, 2020.

[68] G. Stiglic, P. Kocbek, N. Fijacko, **M. Zitnik**, K. Verbert, L. Cilar. Interpretability of Machine Learning based Prediction Models in Healthcare. *WIREs Data Mining and Knowledge Discovery*, 10:e1379, 2020.

[67] **M. Zitnik**, R. Sosic, M.W. Feldman, J. Leskovec. Evolution of Resilience in Protein Interactomes Across the Tree of Life. *Proceedings of the National Academy of Sciences (PNAS)*, 116, 10:4426-4433, 2019.

[66] W. Nelson, **M. Zitnik**, B. Wang, J. Leskovec, A. Goldenberg, R. Sharan. To Embed or Not: Network Embedding as a Paradigm in Computational Biology. *Frontiers in Genetics*, 10, 381, 2019.

[65] **M. Zitnik**, F. Nguyen, B. Wang, J. Leskovec, A. Goldenberg, M.M. Hoffman. Machine Learning for Integrating Data in Biology and Medicine: Principles, Practice, and Opportunities. *Information Fusion*, 50, 71-91, 2019.

[64] A. Copar, B. Zupan, **M. Zitnik**. Fast Optimization of Non-negative Matrix Tri-factorization. *PLoS One*, 14, 6, e0217994, 2019.

[63] **M. Zitnik**, R. Sosic, J. Leskovec. Prioritizing Network Communities. *Nature Communications*, 9, 2544, 2018.

[62] B. Wang*, A. Pourshafeie*, **M. Zitnik***, J. Zhu, C.D. Bustamante, S. Batzoglou, J. Leskovec. Network Enhancement as a General Method to Denoise Weighted Biological Networks. *Nature Communications*, 9, 3108, 2018. *(* first co-authors)*

[61] **M. Zitnik**, M. Agrawal, J. Leskovec. Modeling Polypharmacy Side Effects with Graph Convolutional Networks. *Bioinformatics*, 34, 13:457-466, 2018.
Oral presentation at ISMB 2018.
Covered by 30 news outlets in the US and Europe.
6th most-read paper in *Bioinformatics* (source: Altmetric by Oxford University Press).

[60] X. Wang, Y. Zhang, X. Ren, Y. Zhang, **M. Zitnik**, J. Shang, C. Langlotz, J. Han. Cross-type Biomedical Named Entity Recognition with Deep Multi-Task Learning. *Bioinformatics*, 35, 10:1745-1752, 2018.

[59] A. Copar, **M. Zitnik**, B. Zupan. Scalable Non-Negative Matrix Tri-Factorization. *BioData Mining*, 10, 41, 2017.

[58] J.C. Puigvert, **M. Zitnik**, A.-S. Jemth, M. Carter, J. Unterlass, B. Hallstrom, O. Loseva, Z. Karem, J.M. Calderon-Montano, C. Lindskog, P.H. Edqvist, D. Matuszewski, H.A. Blal, R. Berntsson, M. Haggblad, U. Martens, M. Studham, M. Uhlen, B. Lundgren, C. Wahlby, E. Sonnhammer, E. Lundberg, P. Stenmark, B. Zupan, T. Helleday. A Comprehensive Structural, Biochemical and Biological Profiling of the Human NUDIX Hydrolase Family. *Nature Communications*, 8, 1541, 2017.
I led all computational analyses in this project.

[57] **M. Zitnik**, J. Leskovec. Predicting Multicellular Function Through Multi-Layer Tissue Networks. *Bioinformatics*, 33, 14:190-198, 2017.
Oral presentation at ISMB 2017.

[56] A. Keller, R.C. Gerkin, Y. Guan, A. Dhurandhar, G. Turu, B. Szalai, J.D. Mainland, Y. Ihara, C.W. Yu, R. Wolfinger, C. Vens, L. Schietgat, K. De Grave, R. Norel, **DREAM Olfaction Consortium**, G. Stolovitzky, G. Cecchi, L.B. Vosshall, P. Meyer. Predicting Human Olfactory Perception from Chemical Features of Odor Molecules. *Science*, 355, 6327:820-826, 2017.

[55] **M. Zitnik**, B. Zupan. Jumping Across Biomedical Contexts Using Compressive Data Fusion. *Bioinformatics*, 32, 12:90-100, 2016.
Oral presentation at ISMB 2016.

[54] M. Strazar, **M. Zitnik**, B. Zupan, J. Ule, T. Curk. Orthogonal Matrix Factorization Enables Integrative Analysis of Multiple RNA Binding Proteins. *Bioinformatics*, 32, 10:1527-1535, 2016.

[53] S. Zitnik, **M. Zitnik**, B. Zupan, M. Bajec. Sieve-Based Relation Extraction of Gene Regulatory Networks from Biological Literature. *BMC Bioinformatics*, 16, Suppl 16:S1, 2015.

[52] **M. Zitnik**, E. A. Nam, C. Dinh, A. Kuspa, G. Shaulsky, B. Zupan. Gene Prioritization by Compressive Data Fusion and Chaining. *PLoS Computational Biology*, 11, 10:e1004552, 2015.
The computer science department award for outstanding research.

[51] **M. Zitnik**, B. Zupan. Gene Network Inference by Fusing Data from Diverse Distributions. *Bioinformatics*, 31, 12:230-239, 2015.
Oral presentation at ISMB 2015.

[50] M. M. Usaj*, M. Brloznik*, P. Kaferle*, **M. Zitnik**, H. Wolinski, F. Leitner, S. D. Kohlwein, B. Zupan, U. Petrovic. Genome-Wide Localization Study of Yeast Pex11 Identifies Peroxisome-Mitochondria Interactions

through the ERMES Complex. *Journal of Molecular Biology*, 427, 11:2072–2087, 2015. *(* first co-authors)*

[49] **M. Zitnik**, B. Zupan. Survival Regression by Data Fusion. *Systems Biomedicine*, 2, 3:47-53, 2015.

[48] **M. Zitnik**, B. Zupan. Data Imputation in Epistatic MAPs by Network-Guided Matrix Completion. *Journal of Computational Biology*, 22, 6:595-608, 2015.

[47] **M. Zitnik**, B. Zupan. Data Fusion by Matrix Factorization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37, 1:41-53, 2015.

[46] **M. Zitnik**, B. Zupan. Gene Network Inference by Probabilistic Scoring of Relationships from a Factorized Model of Interactions. *Bioinformatics*, 30, 12:246-254, 2014.
Oral presentation at ISMB 2014.
The computer science department award for outstanding research.

[45] **M. Zitnik**, B. Zupan. Matrix Factorization-Based Data Fusion for Drug-Induced Liver Injury Prediction. *Systems Biomedicine*, 2, 1:16-22, 2014.

[44] **M. Zitnik**, V. Janjic, C. Larminie, B. Zupan, N. Przulj. Discovering Disease-Disease Associations by Fusing Systems-Level Molecular Data. *Scientific Reports*, 3, e3202, 2013.
The computer science department award for outstanding research.

[43] J. Demsar, ... **M. Zitnik**, *et al.* Orange: Data Mining Toolbox in Python. *Journal of Machine Learning Research*, 14:2349-2353, 2013.

[42] **M. Zitnik**, B. Zupan. Nimfa: A Python Library for Nonnegative Matrix Factorization. *Journal of Machine Learning Research*, 13:849-853, 2012.

◇ **Articles in Peer-Reviewed Proceedings**

[41] X. Zhang, M. Zeman, T. Tsiligkaridis, **M. Zitnik**. Graph-Guided Network for Irregularly Sampled Multivariate Time Series. *In Review*. [arXiv:2110.05357]

[40] C. Agarwal, M. Zitnik, H. Lakkaraju. Towards a Rigorous Theoretical Analysis and Evaluation of GNN Explanations. *In Review*. [arXiv:2106.09078]

[39] Q. Wang, K. Huang, P. Chandak, N. Gehlenborg, **M. Zitnik**. Interactive Visual Explanations for Deep Drug Repurposing. Proceedings of the International Conference on Machine Learning, ICML, Workshop on Interpretable ML in Healthcare, 2021.
Best Paper Award

[38] M.M. Li, **M. Zitnik**. Deep Contextual Learners for Protein Networks. Proceedings of the International Conference on Machine Learning, ICML, Workshop on Computational Biology, 2021.
Best Poster Award

[37] K. Huang, T. Fu, W. Gao, Y. Zhao, Y. Roohani, J. Leskovec, C.W. Coley, C. Xiao, J. Sun, **M. Zitnik**. Therapeutics Data Commons: Machine Learning Datasets and Tasks for Drug Discovery and Development. Proceedings of Neural Information Processing Systems, NeurIPS Datasets and Benchmarks, 35, 2021.

[36] M.B.A. McDermott, B. Yap, P. Szolovits, **M. Zitnik**. Structure Inducing Pre-Training. *In Review*. [arXiv:2103.10334]

[35] C. Agarwal, H. Lakkaraju, **M. Zitnik**. Towards a Unified Framework for Fair and Stable Graph Representation Learning. Proceedings of International Conference on Uncertainty in Artificial Intelligence, UAI, 37, 2021.

[34] E. Alsentzer, S.G. Finlayson, M.M. Li, **M. Zitnik**. Subgraph Neural Networks. Proceedings of Neural Information Processing Systems, NeurIPS, 34, 2020.

[33] K. Huang, **M. Zitnik**. Graph Meta Learning via Local Subgraphs. In Proceedings of Neural Information Processing Systems, NeurIPS, 34, 2020.

[32] X. Zhang, **M. Zitnik**. GNNGuard: Defending Graph Neural Networks against Adversarial Attacks. In Proceedings of Neural Information Processing Systems, NeurIPS, 34, 2020.

[31] W. Hu, M. Fey, **M. Zitnik**, Y. Dong, H. Ren, B. Liu, M. Catasta, J. Leskovec. Open Graph Benchmark: Datasets for Machine Learning on Graphs. In Proceedings of Neural Information Processing Systems, NeurIPS, 34, 2020.

[30] K. Huang, T. Fu, D. Khan, A. Abid, A. Abdalla, A.Abid, L.M. Glass, **M. Zitnik**, C. Xiao, J. Sun. MolDe-

signer: Interactive Design of Efficacious Drugs with Deep Learning. In Proceedings of Neural Information Processing Systems, Demo, NeurIPS, 34, 2020.

[29] W. Hu, B. Liu, J. Gomes, **M. Zitnik**, P. Liang, V. Pande, J. Leskovec. Pre-training Graph Neural Networks. In Proceedings of the International Conference on Learning Representations, ICLR, 8, 2020.

[28] R. Ying, D. Bourgeois, J. You, **M. Zitnik**, J. Leskovec. GNN Explainer: Generating Explanations for Graph Neural Networks. In Proceedings of Neural Information Processing Systems, NeurIPS, 33, 2019.

[27] W. Hamilton, P. Bajaj, **M. Zitnik**, D. Jurafsky, J. Leskovec. Embedding Logical Queries on Knowledge Graphs. In Proceedings of Neural Information Processing Systems, NeurIPS, 32, 2018.

[26] C. Donnat, **M. Zitnik**, D. Hallac, J. Leskovec. Learning Structural Node Embeddings via Diffusion Wavelets. In ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD, 2018.

[25] M. Agrawal*, **M. Zitnik***, J. Leskovec. Large-Scale Analysis of Disease Pathways in the Human Interactome. In Pacific Symposium on Biocomputing, 23:111-122, 2018. *(* first co-authors)*

[24] **M. Zitnik**, B. Zupan. Collective Pairwise Classification for Multi-Way Data Analysis of Diseases and Drugs. In Pacific Symposium on Biocomputing, 21:81-92, 2016.

[23] **M. Zitnik**, B. Zupan. Large-Scale Data Fusion. In Minisymposia at 37th Annual International Conference of the IEEE Engineering in Medicine and Biology, EMBC, 2015.

[22] **M. Zitnik**, B. Zupan. Survival Regression by Data Fusion. In International Conference on Intelligent Systems for Molecular Biology, ISMB, CAMDA, 2014.
Best paper award for outstanding research at ISMB/CAMDA 2014.

[21] **M. Zitnik**, B. Zupan. Imputation of Quantitative Genetic Interactions in Epistatic MAPs by Interaction Propagation Matrix Completion. In Proceedings of 18th Annual International Conference on Research in Computational Molecular Biology, RECOMB, 8394:448-462, 2014.

[20] **M. Zitnik**, B. Zupan. Matrix Factorization-Based Data Fusion For Gene Function Prediction in Baker's Yeast and Slime Mold. In Pacific Symposium on Biocomputing, 19:400-411, 2014.

[19] **M. Zitnik**, B. Zupan. Matrix Factorization-Based Data Fusion for Drug-Induced Liver Injury Prediction. In International Conference on Intelligent Systems for Molecular Biology, ISMB, CAMDA, 2013.
best paper award for outstanding research at ISMB/CAMDA 2013.

[18] S. Zitnik, **M. Zitnik**, B. Zupan, M. Bajec. Extracting Gene Regulation Networks Using Linear-Chain Conditional Random Fields and Rules. In Association for Computational Linguistics, ACL, BioNLP, 2013.
First-prize winner in gene regulatory network extraction at ACL BioNLP 2013.

[17] J. Zbontar, **M. Zitnik**, M. Zidar, G. Majcen, M. Potocnik, B. Zupan. Team ULjubljana's Solution to the JRS 2012 Data Mining Competition. In Rough Sets and Current Trends in Computing, 7413/2012, 471-478, 2012.

◇ **Abstracts in Peer-Reviewed Proceedings**

[16] S. Kunz, **M. Zitnik**, J. Zupancic, J. Leskovec, J. Rigdon, J. Profit. Variation in Neonatal Referral Networks across Patient Groups. Pediatric Academic Societies Meeting, PAS, 2019.

[15] **M. Zitnik**, M. Agrawal, J. Leskovec. Modeling Polypharmacy Side Effects with Graph Convolutional Networks. 26th International Conference on Intelligent Systems for Molecular Biology, ISMB/ECCB, 2018.

[14] M. Agrawal*, **M. Zitnik***, J. Leskovec. Large-Scale Analysis of Disease Pathways in the Human Interactome. Pacific Symposium on Biocomputing, PSB, 2018. *(* first co-authors)*

[13] **M. Zitnik**, J. Leskovec. Predicting Multicellular Function Through Multi-Layer Tissue Networks. 25th International Conference on Intelligent Systems for Molecular Biology, ISMB/ECCB, 2017.

[12] **M. Zitnik**, B. Zupan. Jumping Across Gene-Disease Contexts Using Compressive Data Fusion. 24th International Conference on Intelligent Systems for Molecular Biology, ISMB, 2016.

[11] **M. Zitnik**, B. Zupan. Jumping Across Biomedical Contexts Using Compressive Data Fusion. Pacific Symposium on Biocomputing, 2016.

[10] **M. Zitnik**, B. Zupan. Integrate Everything but the Kitchen Sink: Data Set Selection and Sensitivity Estimation in Collective Factor Models. 23rd International Conference on Intelligent Systems for Molecular

Biology, ISMB, 2015.

[9] **M. Zitnik**, E. A. Nam, C. Dinh, A. Kuspa, G. Shaulsky, B. Zupan. Gene Prioritization by Compressive Data Fusion and Chaining. Basel Computational Biology Conference, BC^2, 2015.
Best poster award.

[8] M. M. Usaj, M. Usaj, **M. Zitnik**, D. Kablawi, B. Zupan, B. J. Andrews, C. Boone. Exploring the Yeast Endocytic Pathway by Combining High-Throughput Genetics and High-Content Microscopy. International Specialised Symposium on Yeast, 2014.

[7] **M. Zitnik**. Learning by Fusing Heterogeneous Data. Heidelberg Laureate Forum, 2014.

[6] **M. Zitnik**, E. A. Nam, C. Dinh, A. Kuspa, G. Shaulsky, B. Zupan. Data Fusion for Gene Prioritization. 22nd International Conference on Intelligent Systems for Molecular Biology, ISMB, 2014.

[5] **M. Zitnik**, B. Zupan. Biomedical Data Fusion by Simultaneous Matrix Tri-factorization. 18th Annual International Conference on Research in Computational Molecular Biology, RECOMB, 2014.
Best poster award.

[4] **M. Zitnik**, B. Zupan. Matrix Factorization-Based Data Fusion For Gene Function Prediction in Baker's Yeast and Slime Mold. Pacific Symposium on Biocomputing, 2014.

[3] M. M. Usaj, M. Usaj, **M. Zitnik**, B. Zupan, D. G. Drubin, B. J. Andrews, C. Boone. Studying Yeast Endocytosis with High-throughput Multi-channel Fluorescence Microscopy: a Tale of Cortical Actin Patches, Endosomes and Vacuoles. 13th International Conference on Systems Biology, ICSB, 2012.

◇ **Theses**

[2] **M. Zitnik**. A Matrix Factorization Approach for Inference from Heterogeneous Data. B.Sc. Thesis, University of Ljubljana, 2012.
Summa cum laude.
University of Ljubljana Prešeren Award.

[1] **M. Zitnik**. Learning by Fusing Heterogeneous Data, Ph.D. Thesis, University of Ljubljana, 2015.
Summa cum laude.
Jozef Stefan Golden Emblem.
Nominated for the Best European Doctoral Dissertation in Artificial Intelligence by the European Association for Artificial Intelligence.

Workshop & Symposium Organizer

◇ Trustworthy AI for Healthcare, Workshop at the 36th AAAI Conference on Artificial Intelligence, AAAI, 2022.

◇ AI for Science: Mind the Gaps, Workshop at the 35th International Conference on Neural Information Processing Systems, NeurIPS, 2021.

◇ International Workshop on AI in Health: Transferring and Integrating Knowledge for Better Health, Workshop at the Web Conference, WWW, 2021.

◇ Trustworthy AI for Healthcare, Workshop at the 35th AAAI Conference on Artificial Intelligence, AAAI, 2021.

◇ National Symposium on Drug Repurposing for Future Pandemics, 2020.

◇ PhD Forum, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, ECML-PKDD, 2020.

◇ Graph Representation Learning and Beyond. Workshop at the 37th International Conference on Machine Learning, ICML, 2020.

◇ Representation Learning on Graphs and Manifolds. Workshop at the 7th International Conference on Learning Representations, ICLR, 2019.

◇ Machine Learning for Drug Development. Tutorial at the 29th International Joint Conference on Artificial Intelligence, IJCAI, 2020.

◇ Deep Learning for Network Biology. Tutorial at the 26th International Conference on Intelligent Systems for Molecular Biology, ISMB, 2018.

◇ Large-Scale Data Fusion by Collective Matrix Factorization. Tutorial at the 11th Basel Computational Biology Conference, BC^2, 2015.

◇ Data Fusion of Everything. Tutorial at the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC, 2015.

Datasets	◇ **M. Zitnik**, R. Sosic, S. Maheshwari, J. Leskovec. BioSNAP Datasets: Stanford Biomedical Network Dataset Collection. http://snap.stanford.edu/biodata, 2018.
	◇ K. Huang, T. Fu, W. Gao, Y. Zhao, Y. Roohani, J. Leskovec, C.W. Coley, C. Xiao, J. Sun, **M. Zitnik**. Therapeutics Data Commons: Machine Learning Datasets and Tasks for Drug Discovery and Development. https://tdcommons.ai, 2021.

Scientific Community Activities

◇ **Selected Invited Talks and Seminars** (see https://zitniklab.hms.harvard.edu/talks)
- IEEE Big Data, Graph Techniques for Adversarial Activity Analytics (12/2021) -- *Keynote*
- SIMBig 2021, International Conference on Information Management and Big Data (12/2021) -- *Keynote*
- University of Cambridge, Making Connections - Brains & Other Complex Systems (11/2021)
- University of Cambridge, Isaac Newton Institute for Mathematical Sciences (11/2021)
- Yale University, Yale Network Science Institute Seminar (11/2021)
- AMIA 2021, Graph Based Machine Learning for Healthcare (11/2021)
- Google Research, Scalable Algorithms for Semi-supervised and Unsupervised Learning (10/2021)
- International Conference on Automated Knowledge Base Construction (AKBC), Explainable Graph-based Machine Learning (10/2021)
- KDD 2021, KDD Workshop on Data Mining in Bioinformatics (08/2021) -- *Keynote*
- KDD 2021, KDD Workshop on Knowledge Graphs (08/2021) -- *Keynote*
- KDD 2021, KDD Cup Workshop (08/2021)
- 2021 AIMI Symposium, Stanford Symposium on Artificial Intelligence in Medicine and Imaging (08/2021)
- IBM Research, IBM Research Seminar (08/2021)
- ISMB/ECCB 2021, Bio-Ontologies (07/2021) -- *Keynote*
- ISMB/ECCB 2021, Network Biology (07/2021) -- *Keynote*
- Johnson & Johnson, Johnson & Johnson Seminar (07/2021)
- Trustworthy Machine Learning Initiative, Trustworthy ML Seminar (07/2021)
- London Geometry and Machine Learning, Geometry and Deep Learning (07/2021)
- Cleveland Clinic, NSF Workshop on The Present and the Future of AI in Biomedical Research (06/2021)
- AIME 2021, Keynote at eXplainable AI in Healthcare (06/2021) -- *Keynote*
- Bayer Foundation, Bayer Research Seminar (05/2021)
- 23andMe, 23andMe Research Seminar (05/2021)
- Genentech, Genentech Research Seminar (05/2021)
- MIT Lincoln Laboratory, Graph Exploitation Symposium (GraphEx) (05/2021)
- Broad Institute of MIT and Harvard, Models, Inference & Algorithms Seminar (05/2021)
- ICLR, Workshop Geometrical and Topological Representation Learning (05/2021)
- MLSys, Workshop on Graph Neural Networks and Systems (04/2021) -- *Keynote*
- Cold Spring Harbor Laboratory, Network Biology Meeting (03/2021)
- University of Massachusetts Amherst, Machine Learning Seminar (12/2020)
- University of Massachusetts Lowell, CS Colloquium (12/2020)
- Purdue University, Department of Computer Science (12/2020)
- MLCB, Machine Learning in Computational Biology (11/2020) -- *Keynote*
- University of Cambridge, Department of Computer Science and Technology (11/2020)
- Harvard Center of Mathematical Sciences and Applications, New Technologies in Mathematics (10/2020)
- KDD, Workshop on Knowledge Graphs: Mining Knowledge Graphs for Deep Insights (07/2020)
- ISMB BioVis (07/2020) -- *Keynote*
- NSF Workshop on Future Directions in Network Biology (06/2020)
- Graph Exploitation Symposium, MIT Lincoln Laboratory (04/2020)
- Dagstuhl Seminar, Visualization of Biological Data -- From Analysis to Communication (04/2020)
- B3D Seminar, Harvard Biomedical Informatics and Biostatistics (04/2020)
- MIT Bioinformatics Seminar, MIT Math & MIT CSAIL (03/2020)
- Swiss Applied Machine Learning Days (01/2020)
- Beth Israel Deaconess Medical Center, Harvard Newborn Health Services and Epidemiology (01/2020)
- NeurIPS Workshop on Graph Representation Learning (GRL) (12/2019)
- DS, 22nd International Conference on Discovery Science (10/2019) -- *Keynote*

- US Food and Drug Administration (FDA) Science Forum (09/2019)
- ISMB/ECCB Machine Learning in Computational & Systems Biology (07/2019) -- *Keynote*
- ICGEB-TRAIN High Content Imaging and Data Science for Virtual Screening and Drug Discovery (05/2019)
- NetSci Meeting on Network Medicine (05/2019)
- Harvard University (04/2019)
- University of Pennsylvania (04/2019)
- Georgia Institute of Technology (04/2019)
- University of California, San Diego (04/2019)
- University of California, Santa Barbara (04/2019)
- University of Illinois at Urbana-Champaign (04/2019)
- University of California, Irvine (04/2019)
- Emory University (04/2019)
- EPFL Ecole Polytechnique Federale de Lausanne (03/2019)
- University of California, San Francisco (03/2019)
- Johns Hopkins University (03/2019)
- Northeastern University (03/2019)
- University of Washington (03/2019)
- Stanford Symposium of Computational, Evolutionary and Human Genomics (CEHG) (02/2019)
- Boston University (02/2019)
- Yale University (02/2019)
- Carnegie Mellon University (02/2019)
- Cornell University (02/2019)
- University of Notre Dame (02/2019)
- University of California, Berkeley (01/2019)
- University of California, Los Angeles (01/2019)
- NeurIPS Workshop on Relational Representation Learning (12/2018)
- The Broad Institute of Harvard and Medicine, Next Generation Symposium (11/2018)
- The US National Academies of Sciences, Engineering, and Medicine (10/2018)
- European Bioinformatics Institute (EMBL-EBI) (09/2018)
- Stanford Clinical Excellence Research Center (CERC) (09/2018)
- AI in Medicine: Inclusion & Equity (AiMIE) (08/2018)
- AMS-MAA Joint Mathematics Meetings, Clustering of Graphs: Theory, Practice, and Applications (01/2018)
- NetSci Meeting on Network Medicine (06/2017)
- International Conference for Big Data and AI in Medicine (04/2017)
- Jozef Stefan Institute (01/2017)
- DARPA Cancer Analytics Workshop, White House Cancer Moonshot (10/2016)
- AIME Workshop on Matrix Computations in Biomedical Informatics (06/2015)
- TEDx University of Ljubljana (10/2014)
- Baylor College of Medicine, Department of Molecular and Human Genetics (01/2014)
- University of Toronto, Donnelly Centre for Cellular and Biomolecular Research (09/2012)

◇ **Conference Chair**
- Demo Co-Chair for ECML/PKDD 2017
- PhD Forum Co-Chair for ECML/PKDD 2020
- General Chair for NSF National Symposium on Drugs for Future Pandemics, 2020
- Co-Chair for Network Biology Track at ISMB 2020
- Co-Chair for Tutorial Track at WWW 2021
- Co-Chair for Biomedical Informatics Area at ISMB 2021
- Co-Chair for Systems Biology & Networks Area at ISMB 2022

◇ **Editorial Board Member and Associate Editor**
- Editorial Board Member, Neurocomputing (2021-)
- IEEE Transactions on Big Data (Guest Editor for Special Issue on AI for COVID-19, 2020)
- IEEE/ACM Transactions on Computational Biology and Bioinformatics (Guest Editor for Special Issue on Deep Learning and Graph Embeddings for Network Biology, 2020)
- Editorial Board Member, Data Mining and Knowledge Discovery (DAMI) (2020-)

- · Editorial Board Member, Artificial Intelligence in Medicine (AIM) (2019-)
- · Frontiers in Genetics (Guest Editor, 2018)
- · ECML PKDD (Editorial Board Member, 2017-2018, 2020)
- · ACM XRDS Crossroads (Department Leader and Editor, 2012-2016)

⬦ **Program Committee Membership.** KDD 2016-Present, ECML PKDD 2017-2019, AIME 2017, ISMB 2018-Present, PSB 2016-2020, SDM 2019-2020, ICML 2019-Present, ICLR 2019-Present, WWW 2019-2020, IJCAI 2019-2020, UAI 2019, NeurIPS 2018-Present, MLCB 2019, AAAI 2020, RECOMB 2020.

⬦ **Journal Reviewing.** Nature Communications, Nature Neuroscience, Proceedings of the National Academy of Sciences of the United States of America (PNAS), Bioinformatics, Journal of Machine Learning Research, IEEE Transactions on Neural Networks and Learning Systems, IEEE Journal of Biomedical and Health Informatics, Computers in Biology and Medicine, IEEE Signal Processing Letters, Journal of Cheminformatics, IEEE Transactions on Computational Biology and Bioinformatics, Nucleic Acids Research, Journal of Biomedical Informatics, Knowledge and Information Systems, BMC Bioinformatics, Machine Learning Journal, BioData Mining, Data Mining and Knowledge Discovery Journal, IEEE Transactions on Knowledge and Data Engineering, PLoS One, PLoS Computational Biology, Nature Scientific Reports, Nature Machine Intelligence