

2020

## Explainable AI Using Knowledge Graphs

Manas Gaur

*University of South Carolina - Columbia*

Ankit Desai

*Embibe Inc.*

Keyur Faldu

*Embibe Inc.*

Amit Sheth

*University of South Carolina - Columbia*

Follow this and additional works at: [https://scholarcommons.sc.edu/aii\\_fac\\_pub](https://scholarcommons.sc.edu/aii_fac_pub)



Part of the [Computer Engineering Commons](#), [Educational Methods Commons](#), [Electrical and Computer Engineering Commons](#), and the [Public Health Commons](#)

---

### Publication Info

Preprint version *ACM CoDS-COMAD Conference*, 2020.

© Association for Computing Machinery, 2020.

This Article is brought to you by the Artificial Intelligence Institute at Scholar Commons. It has been accepted for inclusion in Publications by an authorized administrator of Scholar Commons. For more information, please contact [digres@mailbox.sc.edu](mailto:digres@mailbox.sc.edu).

# Explainable AI using Knowledge Graphs

Manas Gaur  
Artificial Intelligence Institute  
Columbia, South Carolina, USA  
mgaur@email.sc.edu

Ankit Desai  
Embibe Inc.  
Bengaluru, Karnataka, India  
ankit.desai@embibe.com

Keyur Faldu  
Embibe Inc.  
Bengaluru, Karnataka, India  
k@embibe.com

Amit Sheth  
Artificial Intelligence Institute  
Columbia, South Carolina, USA  
amit@sc.edu

## ABSTRACT

During the last decade, traditional data-driven deep learning (DL) has shown remarkable success in essential natural language processing tasks, such as relation extraction. Yet, challenges remain in developing artificial intelligence (AI) methods in real-world cases that require explainability through human interpretable and traceable outcomes. The scarcity of labeled data for downstream supervised tasks and entangled embeddings produced as an outcome of self-supervised pre-training objectives also hinders interpretability and explainability. Additionally, data labeling in multiple unstructured domains, particularly healthcare and education, is computationally expensive as it requires a pool of human expertise. Consider Education Technology, where AI systems fall along a “capability spectrum” depending on how extensively they exploit various resources, such as academic content, granularity in student engagement, academic domain experts, and knowledge bases to identify concepts that would help achieve knowledge mastery for student goals. Likewise, the task of assessing human health using online conversations raises challenges for current statistical DL methods through evolving cultural and context-specific discussions. Hence, developing strategies that merge AI with stratified knowledge to identify concepts that would delineate healthcare conversations patterns and help healthcare professionals decide. Such technological innovations are imperative as they provide consistency and explainability in outcomes. This tutorial discusses the notion of explainability and interpretability through the use of knowledge graphs in (1) Healthcare on the Web, (2) Education Technology. This tutorial will provide details of knowledge-infused learning algorithms and its contribution to explainability for the above two applications that can be applied to any other domain using knowledge graphs.

## KEYWORDS

Knowledge infusion, Knowledge graphs, Public Health, Education Technology, Explainability, and Traceability in AI, Responsible Data Science.

## 1 GOAL AND OBJECTIVE OF THE TUTORIAL

Recently, there is increasing attention to developing methods to enable the easy adoption of AI in practice. These methods comprise analyzability, interpretability, traceability, and explainability of AI models and its prediction using statistical natural language processing (NLP), information extraction (IE), and deep learning (DL). At the center of this upsurge are the knowledge graphs (KG),

a large network of entities, their semantic types, properties, and relationships between entities. For instance, Wikidata [17], DBpedia [12], UMLS [1], ConceptNet [15]. The utility of KG in DL is to provide relative importance scores (or semantic weighting) to learnable features for interpretation of the outcomes [16, 18]. A recent study [3] leverage the Columbia-Suicide Severity Rating Scale (C-SSRS) to clinically assess the varying suicidality on Reddit and identify behavioral cues to extract supportive users, which improved the recall of the overall process. Likewise, in the domain of education, the use of deep learning and knowledge infusion methods to extend Bayesian Knowledge Tracing (BKT) and Deep Knowledge Tracing (DKT) pair the capability to present concept-level masteries for easy intervention with the traditional prediction of student’s overall mastery. Further, the student’s goal contextualizes the knowledge graph with a relevant curriculum, whereas his concept-level mastery can be used to personalize the knowledge graph. Which enables the intervention layer to generate better learning outcomes. For instance, a recent study used Blooms’ Taxonomy to extend the BKT model of a student to interpretably evaluate knowledge mastery [2, 10, 11]. Such utilization of KG/taxonomies is of immediate concentration in the domain of explainable AI and data science, a paradigm of importance to the community in the Big Data Analytics Conference. Through demos, implementable details, and resources, the focus of the tutorial is to provide directions in developing AI methods involving the fusion of knowledge for contextualizing the content, generating attention weights, and evaluating predictions to reason over the healthcare state or learning activity of an individual for improving the human experience with AI system. Beginners in the area of Knowledge Graphs would learn an introduction to the knowledge graph, its construction, and utility through knowledge-infused learning. Experts in the domain of Knowledge Graphs (KGs) and its application in AI would appreciate the capability of knowledge-infused learning in generating explanations to classification problems in Healthcare on the Web and deriving concepts mastery in Education. Moreover, the attendees will gain in-depth understanding of the methods used for fusing knowledge through demonstrations, theory, and conceptual outlining of the tutorial. The tutorial is appropriate for the community engaging in the Conference on Big Data Analytics as it highlights the essence of contextualized knowledge representation and knowledge-based systems in facilitating human-guided interventions in Education and Healthcare on the Web for effective outcomes. It will provide the community with tools to overcome obstacles in social good domains that lack high-quality training

data and poor interpretability. We believe the resources provided during the tutorial will further the research in responsible data science in healthcare and education.

## 2 TUTORIAL DESCRIPTION AND OUTLINE

The critical issues in healthcare, particularly mental healthcare, such as estimating the mental health state of an individual and education, such as calibrating and improving knowledge mastery of an individual, have been highlighted by research in artificial intelligence. The methods focus on principles of IE (e.g., predicate invention [9]), DL (e.g., convolutional neural networks [14], deep knowledge tracing using the recurrent neural networks [19]), and behavioral NLP to predict either severity of mental illness or knowledge mastery in education. No doubt, deep learning-based approaches have achieved state-of-the-art performance in knowledge state and mental state prediction. Still, it neither provides explicit explanations of its outcome nor provides actionable suggestions on achieving clinical relevance in healthcare or desired mastery in education. The domain of healthcare and education technology is stocked with publicly available data in various forms (e.g., social media crawls, ebooks/PDFs). Leveraging such a large corpus of unstructured text for improving counseling services in healthcare and learning outcomes in education, a paradigm coalesced with external curated knowledge is desired. This tutorial presents a paradigm, Knowledge-infused Learning, which describes methodologies of incorporating domain-knowledge in deep learning approaches for bringing consistency and robustness in predictions. Specifically, we will provide implementable details on Shallow Infusion, Semi-Deep Infusion, and Deep Infusion of Knowledge, which are methods to augment auxiliary knowledge for helping the mental healthcare providers educationist to understand key features contributing to the current knowledge state or mental health state of an individual. Further, we will discuss methods for contextualization and abstraction of unstructured text, independently in healthcare and education using knowledge graphs/taxonomy [3, 5, 7, 8], to derive actionable knowledge for mental healthcare providers and educationists. We will describe strategies for evaluating methods of knowledge infusion with examples from mental healthcare and online education.

### 2.1 Tutorial Outline

We begin with the introduction to Knowledge Graph-based Learning comprising of (a) description on the role of KGs in model interpretability, traceability of predictions to a KG, in achieving explainability of the outcome with examples and (b) procedure to construct KGs and use them at scale. Following which, we would motivate the audience on the concept of Knowledge-infused AI which would include (a) the theoretical underpinnings of knowledge-infused learning, (b) the different alternatives for combining KGs and Learning methods, (c) KG-based AI frameworks in practice, and (d) different evaluation strategies for such frameworks. Thereafter, we would give a preface of using KGs towards improving learning outcomes (e.g. Amazon Alexa, Coursera, eDX) and in healthcare informatics. This introductory session would cover prior research covering shallow and semi-deep infusion of knowledge in deep learning or machine learning procedures [6]. Subsequently, we will take up

two key applications to provide a practical tour on the theme of the tutorial:

- (1) **Knowledge-infused Learning in Education:** would deliver insights into different KG-type resources in education, methods of construction, and ways to infuse KG in the current state-of-the-art knowledge tracing approaches in education: Deep Knowledge Tracing [13]. During this part of the tutorial, we will provide a demo of an industry use-case at Embibe and different strategies of evaluating this use-case.
- (2) **Knowledge-infused Learning in Healthcare:** will concentrate on two important use-cases of “Healthcare on the Web”: (a) Utilization of Diagnostic Statistical Manual for Mental Health Disorder (DSM-5) to understand Reddit communication and (b) Dynamic peer-support group formation on Reddit using this understanding. In the process of addressing research challenges in these expositions, we would discuss methods on associating Medical Knowledge Graphs (e.g. SNOMED-CT, UMLS [4]) with Healthcare on the Web.

## 3 TARGET AUDIENCE

This tutorial will bring researchers in academia, industry, humanitarian organizations, and healthcare practitioners at the confluence of knowledge representation, natural language understanding, and deep learning. Prior exposure to the basic concepts in NLP and DL is desirable, however, there are no prerequisites for attending the tutorial. We will cover basics and advanced techniques with sufficient use cases and demonstrations. Newcomers in the area will learn the basic principles of data science and the fundamentals of knowledge-infused learning. Expert attendees will appreciate promising, reliable, and practical approaches to overcoming familiar technical obstacles in social good domains.

## 4 PRESENTERS' BIOGRAPHIES

### (1) Amit P. Sheth<sup>1</sup>

Prof. Amit Sheth is an Educator, Researcher, and Entrepreneur. He is the founding director of the university-wide Artificial Intelligence Institute at the University of South Carolina (#AIISC). Previously, he was the LexisNexis Ohio Eminent Scholar and the executive director of Ohio Center of Excellence in Knowledge-enabled Computing. He is a Fellow of IEEE, AAAI, and AAAS. He has organized 100 international events (general/program chair, organization committee chair), 65+ keynotes, given many well-attended tutorials, and is among the well-cited computer scientists. He has founded three companies by licensing his university research outcomes, including the first Semantic Web company in 1999 that pioneered technology similar to what is found today in Google Semantic Search and Knowledge Graph. Several commercial products and deployed systems have resulted from his research.

### (2) Manas Gaur<sup>2</sup>

Manas Gaur is currently a Ph.D. student in the Artificial Intelligence Institute at the University of South Carolina. He has been a Data Science and AI for Social Good Fellow with

<sup>1</sup><https://www.linkedin.com/in/amitsheth/>

<sup>2</sup><https://manasgaur.github.io/>

the University of Chicago and Dataminr Inc. His interdisciplinary research funded by NIH and NSF operationalizes the use of Knowledge Graphs, Natural Language Understanding, and Machine Learning to solve social good problems in the domain of Mental Health, Cyber Social Harms, and Crisis Response. His work has appeared in premier AI and Data Science conferences (CIKM, WWW, AAAI, CSCW), journals in science (PLOS One, Springer-Nature, IEEE Internet Computing), and healthcare-specific meetings (NIMH MHSR, AMIA).

(3) **Keyur Faldu**<sup>3</sup>

Keyur Faldu is currently working as a chief data scientist at Embibe. He has founded and built the data science lab at Embibe, which is primarily responsible for building AI platforms for Learning Outcomes using intelligent content authoring, and intelligent intervention leveraging educational knowledge base and a big data lake. He has more than 14 years of industry experience in data science, machine learning, deep learning, and natural language processing. He worked at startups Veveo Inc, and Runa, and also was part of the founding team for data science at McKinsey Digital Labs at McKinsey. He holds post-graduation in Computer Science from the Indian Institute of Science, Bangalore.

(4) **Ankit Desai**<sup>4</sup>

Ankit Desai (Ph.D.) is currently working as an Associate Principal Data Scientist at Data Science Lab at Embibe. He is involved in a process of architecting the Knowledge Graph. Overall 11+ years of experience. Apart from working at Embibe, he has published multiple research papers in international conferences and journals. His current research interests include Educational Data Mining, Graph Mining, Mining Massive Data sets, and Cost-sensitive Data Mining. Architecting the data science projects from scratch to production level interests him the most. Moreover, he has served as a review committee member for conferences of ACM and IEEE.

## 5 PAST EXPERIENCES

- (1) Knowledge-infused Learning for Healthcare, PyData Virtual Conference at Salamanca, 2020 (Slides<sup>5</sup>, Video<sup>6</sup>)
- (2) Knowledge-infused Statistical Learning for Social Good Applications, PyData Virtual Conference at Berlin 2020 (Slides<sup>7</sup>, Video<sup>8</sup>)
- (3) Knowledge-infused Deep Learning, Tutorial at 31st ACM Hypertext and Social Media (Virtual) 2020. (Slides<sup>9</sup>, Video<sup>10</sup>)

- (4) Knowledge will propel Machine Understanding of Big Data, China Conference on Knowledge Graph and Semantic Computing (CCKS) 2017. (Slides<sup>11</sup>, Video<sup>12</sup>)
- (5) Knowledge Graphs and their Central Role in Big Data Processing: Past, Present, and Future, Keynote at 7th ACM CoDS and 25th COMAD 2020. (Slides<sup>13</sup>)
- (6) AI in Education: Transforming education using Personalized Adaptive Learning, Open Data Science Conference 2019. (Video<sup>14</sup>)
- (7) Explainability of Medical AI through Domain Knowledge, Ontology Summit 2019. (Slides<sup>15</sup>, Video<sup>16</sup>)
- (8) Workshop on Knowledge-infused Mining and Learning, Co-located with 26th ACM SIGKDD conference. (Slides<sup>17</sup>, Proceedings<sup>18</sup>).

## 6 ACKNOWLEDGEMENT

We acknowledge partial support from the National Institutes of Health (NIH) award: MH105384-01A1: "Modeling Social Behavior for Healthcare Utilization in Depression", National Science Foundation (NSF) award 1761880: "Spokes: MEDIUM: MID-WEST: Collaborative: Community-Driven Data Engineering for Substance Abuse Prevention in the Rural Midwest". Any opinions, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

## REFERENCES

- [1] Olivier Bodenreider. 2004. The unified medical language system (UMLS): integrating biomedical terminology. *Nucleic acids research* 32, suppl\_1 (2004), D267–D270.
- [2] Soma Dhavala, Chirag Bhatia, Joy Bose, Keyur Faldu, and Aditi Avasthi. 2020. Auto Generation of Diagnostic Assessments and Their Quality Evaluation. *International Educational Data Mining Society* (2020).
- [3] Manas Gaur, Amanuel Alambo, Joy Prakash Sain, Ugur Kursuncu, Krishnaprasad Thirunarayan, Ramakanth Kavuluru, Amit Sheth, Randy Welton, and Jyotishman Pathak. 2019. Knowledge-aware assessment of severity of suicide risk for early intervention. In *The World Wide Web Conference*. 514–525.
- [4] Manas Gaur, Keyur Faldu, and Amit Sheth. 2020. Semantics of the Black-Box: Can knowledge graphs help make deep learning systems more interpretable and explainable? *arXiv preprint arXiv:2010.08660* (2020).
- [5] Manas Gaur, Ugur Kursuncu, Amanuel Alambo, Amit Sheth, Raminta Daniulaityte, Krishnaprasad Thirunarayan, and Jyotishman Pathak. 2018. "Let Me Tell You About Your Mental Health!" Contextualized Classification of Reddit Posts to DSM-5 for Web-based Intervention. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*. 753–762.
- [6] Manas Gaur, Ugur Kursuncu, Amit Sheth, Ruwan Wickramarachchi, and Shweta Yadav. 2020. Knowledge-infused Deep Learning. In *Proceedings of the 31st ACM Conference on Hypertext and Social Media*. 309–310.
- [7] Christian Glahn and Marion R Gruber. 2020. Designing for context-aware and contextualized learning. In *Emerging Technologies and Pedagogies in the Curriculum*. Springer, 21–40.
- [8] Amelia Gyrard, Manas Gaur, Saeedeh Shekarpour, Krishnaprasad Thirunarayan, and Amit Sheth. 2018. Personalized health knowledge graph. (2018).

<sup>3</sup><https://in.linkedin.com/in/keyur-faldu-681b6633>

<sup>4</sup><https://www.linkedin.com/in/desaiankitb/>

<sup>5</sup><https://docs.google.com/presentation/d/1F69f6s0ft4zjjp5o02qlaAt5pyP0H3FLr2joRBDZUU/edit?usp=sharing>

<sup>6</sup><https://www.youtube.com/watch?v=AVEXOtEL0-A&t=24s>

<sup>7</sup>[https://docs.google.com/presentation/d/1bOSgTRl0zJR2i4tB\\_FEN9jhb7XY3mml4gUOjzYwOQ/edit?usp=sharing](https://docs.google.com/presentation/d/1bOSgTRl0zJR2i4tB_FEN9jhb7XY3mml4gUOjzYwOQ/edit?usp=sharing)

<sup>8</sup><https://www.youtube.com/watch?v=NHTSm5TW7E0&t=2s>

<sup>9</sup><https://www.slideshare.net/aiisc/acm-ht20-tutorialknowledgeinfuseddeeplearningaiinstitute-236984854>

<sup>10</sup><https://www.dropbox.com/s/5wit0i9qrx2hkww/Tutorial.mp4?dl=0>

<sup>11</sup><https://www.slideshare.net/apsheth/knowledge-will-propel-machine-understanding-of-big-data>

<sup>12</sup><https://www.youtube.com/watch?v=4e0dV7CTWM&list=PLqJzTtkUiq57UUmay2DEf09vDGI8zvEqR&index=2>

<sup>13</sup><http://bit.ly/iSEMANTICS2020>

<sup>14</sup><https://www.youtube.com/watch?v=4ppenYfVHNE>

<sup>15</sup>[https://docs.google.com/presentation/d/1Osmve6J7dXrP\\_O9Z-F8W9JfY34dZt2CF-JKp6sWDS4/edit?usp=sharing](https://docs.google.com/presentation/d/1Osmve6J7dXrP_O9Z-F8W9JfY34dZt2CF-JKp6sWDS4/edit?usp=sharing)

<sup>16</sup>[https://s3.amazonaws.com/ontologforum/OntologySummit2019/Medical/MedicalExplanation--UgurKursuncu-ManasGaur\\_20190327.mp4](https://s3.amazonaws.com/ontologforum/OntologySummit2019/Medical/MedicalExplanation--UgurKursuncu-ManasGaur_20190327.mp4)

<sup>17</sup>[https://www.youtube.com/watch?v=Ev6OVHdiQd4&list=PLqJzTtkUiq54bk4w92GzoUguE7jHP\\_LUA](https://www.youtube.com/watch?v=Ev6OVHdiQd4&list=PLqJzTtkUiq54bk4w92GzoUguE7jHP_LUA)

<sup>18</sup><http://ceur-ws.org/Vol-2657/>

- [9] Stanley Kok and Pedro Domingos. 2007. Statistical predicate invention. In *Proceedings of the 24th international conference on Machine learning*. 433–440.
- [10] Amar Lalwani and Sweetey Agrawal. 2017. Few hundred parameters outperform few hundred thousand. In *Proceedings of the 10th International Conference on Educational Data Mining, EDM*, Vol. 17. ERIC, 448–453.
- [11] Amar Lalwani and Sweetey Agrawal. 2018. Validating revised bloom's taxonomy using deep knowledge tracing. In *International Conference on Artificial Intelligence in Education*. Springer, 225–238.
- [12] Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N Mendes, Sebastian Hellmann, Mohamed Morsey, Patrick Van Kleef, Sören Auer, et al. 2015. DBpedia—a large-scale, multilingual knowledge base extracted from Wikipedia. *Semantic web* 6, 2 (2015), 167–195.
- [13] Chris Piech, Jonathan Bassen, Jonathan Huang, Surya Ganguli, Mehran Sahami, Leonidas J Guibas, and Jascha Sohl-Dickstein. 2015. Deep knowledge tracing. In *Advances in neural information processing systems*. 505–513.
- [14] Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne Van Den Berg, Ivan Titov, and Max Welling. 2018. Modeling relational data with graph convolutional networks. In *European Semantic Web Conference*. Springer, 593–607.
- [15] Robyn Speer, Joshua Chin, and Catherine Havasi. 2016. Conceptnet 5.5: An open multilingual graph of general knowledge. *arXiv preprint arXiv:1612.03975* (2016).
- [16] Laura von Rueden, Sebastian Mayer, Katharina Beckh, Bogdan Georgiev, Sven Giesselbach, Raoul Heese, Birgit Kirsch, Julius Pfrommer, Annika Pick, Rajkumar Ramamurthy, et al. 2019. Informed Machine Learning—A Taxonomy and Survey of Integrating Knowledge into Learning Systems. *arXiv preprint arXiv:1903.12394* (2019).
- [17] Denny Vrandečić and Markus Krötzsch. 2014. Wikidata: a free collaborative knowledgebase. *Commun. ACM* 57, 10 (2014), 78–85.
- [18] Xiaozhi Wang, Tianyu Gao, Zhaocheng Zhu, Zhiyuan Liu, Juanzi Li, and Jian Tang. 2019. KEPLER: A unified model for knowledge embedding and pre-trained language representation. *arXiv preprint arXiv:1911.06136* (2019).
- [19] Jinjin Zhao, Shreyansh Bhatt, Candace Thille, Dawn Zimmario, and Neelesh Gattani. 2020. Interpretable Personalized Knowledge Tracing and Next Learning Activity Recommendation. In *Proceedings of the Seventh ACM Conference on Learning@Scale*. 325–328.