**ECML-PKDD 2020: Workshop Proposal**

**GEM@ECML-PKDD 2020**

**https://gem-ecmlpkdd.github.io/**

**Workshop Title:** GEM: Graph Embedding and Mining (2nd edition)

**Contact information and affiliations**

**Remy Cazabet**

Lyon University (LIRIS/UCBL)

[http://cazabetremy.fr](http://cazabetremy.fr/)

[remy.cazabet@univ-lyon1.fr](mailto:remy.cazabet@univ-lyon1.fr)

**Bo Kang**

Ghent University

[http://bokang.io](http://bokang.io/)

[bo.kang@ugent.be](mailto:bo.kang@ugent.be)

**Charlotte Laclau**

Hubert Curien Laboratory, Lyon University (LabHC)

https://laclauc.github.io/

charlotte.laclau@univ-st-etienne.fr

**Christine Largeron**

Hubert Curien Laboratory, Lyon University (LabHC)

<https://perso.univ-st-etienne.fr/largeron/>

[largeron@univ-st-etienne.fr](mailto:largeron@univ-st-etienne.fr)

33 (0)478487406

**Tiphaine Viard**

Univ. Paris XIII

**Pascal Welke**

University of Bonn

<https://mlai.cs.uni-bonn.de/people/pascal-welke>

welke@cs.uni-bonn.de

**Proposed duration of the workshop:** full day

**Preferred workshop date:** September 14, 2020 (2nd choice: September 18)

**Topic**

The goal of this workshop is to further the understanding of how to discover useful and interesting knowledge from network data of various types, forms, and scales. The scope includes theory, algorithms, systems, and applications related to this topic. More specific research areas are listed below.

**General motivation of the area**

Networks of various types, such as property graphs, linked data / semantic web / RDF data, attributed graphs, and more, are increasingly used in practice as versatile and efficient models for data as encountered in today’s data-intensive research and industries. They allow capturing not only the information about entities (using attributes or properties), but also the relational structure between them. This adds significant additional flexibility as compared to data that is formalized as a set of unrelated data points -- a data format that was dominant in the machine learning and data mining literature until quite recently.

The ability to discover knowledge from such network data is therefore gaining in importance quickly. Research is needed both for the formalization of new problem types that match well with practical use cases, as well as into the algorithmic, statistical and information theoretic aspects of such problems.

The aim of this workshop, to be held in conjunction with ECML-PKDD Conference, is to be a discussion forum for presenting the most recent advances on these topics. We will encourage both theoretical and practical contributions in order to stimulate interactions between participants.

**Topics and themes:**

Representation learning

* Network embedding
* Graph compression
* Knowledge graph embedding
* Graph neural networks
* Entity resolution/deduplication
* Clustering

Pattern mining

* Local patterns: community detection, subgraph mining
* Graph summarization
* Subgroup discovery on graphs

Complex Network analysis

* Attributed networks
* Evolving graphs
* Probabilistic and uncertain Networks

Supervised learning

* Classification
* Link prediction
* Knowledge graph completion

Exploration

* Visualization
* Visual analytics

Modeling

* Random graph models
* Opinion formation
* Information propagation

Data management for graph analytics

* Data models and structures
* Index structures

Applications

* Social network analysis
* Biological networks and life science data
* Communication networks
* Urban data, traffic networks

**Why this Workshop should be added to ECML-PKDD 2020?**

Over the recent decades, networks have rapidly grown in importance. Effective methods for analysis and utilization of network data are not only relevant to industry but also pose interesting challenges for academia. Networks of various types are thus increasingly relevant to the ECML-PKDD community, and indeed they are receiving increasing amounts of attention. Yet, they are arguably not perceived as a key focus of the conference, and papers on the topic are often primarily submitted to conferences more explicitly dedicated to networks such as The Web Conference, ACM WSDM, and IEEE/ACM ASONAM, Complex networks as well as to large workshops such as Machine Learning with Graphs (MLG) (which has been co-organized with KDD in recent years). We think that this is particularly relevant at ECML-PKDD since, to the best of our knowledge, there is no well organized European community on this topic yet.

**Previous edition**

# It should be noticed that the first edition of this workshop has been jointly organized with ECML-PKDD2019. This first edition was a real success and has attracted more than 50 participants with 15 accepted papers and a keynote given by Stephan Günnemann (Technical University of Munich).

Detailed information about the program and presented works can be found on the workshop web site: https://gem-ecmlpkdd.github.io/

This new edition is partially organized by the same team.

**Workshop Attendees and Format:**

The proposed workshop is mainly geared towards students, researchers and practitioners actively working on topics related to representation learning for relational data and graph embedding. Given the large popularity of the research theme under consideration, and thanks to the privileged channel offered by the ECML-PKDD conference, we expect a positive response in terms of participants, around 40 people and around 25 submissions.

A dedicated website will be hosted at the address https://gem-ecmlpkdd.github.io/ in order to diffuse the information concerning the workshop, and we will diffuse the call on mailing lists and social media, with the help of the program committee members.

We target a **full day workshop** which will consist of up to one or two invited keynote speaker (45 minutes), followed by up to 4 to 6 selected paper presentations (long and short). The afternoon will be spent with a structured community discussion, focusing on various topics and a poster session presenting the accepted submissions. This shall foster networking among researchers and practitioners and ideally results in new collaborations or project proposals.

We will solicit short position papers (1-2 pages) to present ongoing and preliminary works, long research papers (4-8 pages) and demos. Depending on the number of papers selected for oral presentation, long papers will be 20-30 minutes and short papers 15-20 minutes including Q&A. All accepted papers shall be presented in the afternoon poster session.

**Keynote Speakers:**

Tijl De Bie from Ghent University has already accepted to give an invited talk and we plan to invite another speaker in case of acceptance of this workshop for a full day.

We propose to attract an expert in the field by offering him the waived registration fee for attendance to the conference and the accommodation.

**Paper Submission and selection process:**

Workshop submissions will be evaluated based on their relevance to the workshop topics and goals as well as their novel contribution. Paper submission and publication will preferably follow instructions in accord with the choices made for the main conference. We will use EasyChair to manage the submission process. Every submitted paper will be subject to peer-review by 2-3 reviewers selected from the PC. After the review, we will select papers for full or short papers, and for demo papers, acceptance will imply that at least one author will attend the workshop to present its work.

**Program Committee (to be confirmed)**

The program committee of GEM@ECML-PKDD2019 was composed of the following researchers. We expect them to accept to be part of the PC for this new edition:

* Aleksandar Bojchevski (Technical University of Munich, Germany)
* Amedeo Napoli (LORIA Nancy, France)
* Anes Bendimerad (INSA Lyon, France)
* Anton Tsitsulin (University of Bonn, Germany)
* Baptiste Jeudy (Université de Saint-Étienne, France)
* Chedy Raïssi (INRIA Nancy, France)
* Cigdem Aslay (Aalto University, Finland)
* Darío García García (Facebook AI, USA)
* Di Jin (University of Michigan, USA)
* Florian Adriaens (Ghent University, Belgium)
* Jan Ramon (INRIA Lille, France)
* Jilles Vreeken (Helmholtz Center for Information Security, Germany)
* Kristian Kersting (Technische Universität Darmstadt, Germany)
* Leonardo Gutierrez (Université Catholique de Louvain, Belgium)
* Marc Plantevit (CNRS Lyon, France)
* Mark Heimann (University of Michigan, USA)
* Mehwish Alam (CNR Rome, Italy)
* Mykola Pechenizkiy (Eindhoven University of Technology, Netherlands)
* Nikolaj Tatti (University of Helsinki, Finland)
* Polina Rozenshtein (Aalto University, Finland)
* Raúl Santos-Rodríguez (University of Bristol, UK)
* Renaud Lambiotte (University of Oxford, UK)
* Robert Pienta (Symantec Atlanta, USA)
* Sandra Mitrović (Katholieke Universiteit Leuven, Belgium)
* Stephan Günnemann (Technical University of Munich, Germany)
* Tamás Horváth (University of Bonn, Germany)
* Thomas Gärtner (Technical University of Vienna, Austria)
* Yujun Yan (University of Michigan, USA)

**Workshop chairs**

**Remy Cazabet** is associate professor at Lyon 1 University, Lyon, France. His research focus on graph mining, in particular graph clustering, dynamic network analysis and graph embedding in large graph contexts.

He has organized 4 international workshops on graph related topics in the last 3 years, and is currently leading a 4 years, french National Research Agency (ANR) project (BITUNAM, <http://cazabetremy.fr/BITUNAM.html>) on using machine learning on graphs to analyze the transaction network of crypto-currencies.

Contact: remy.cazabet@univ-lyon1.fr

Website: [http://cazabetremy.fr](http://cazabetremy.fr/)

------------

**Bo Kang** is a postdoctoral researcher at IDLAB, Ghent University, Belgium. His primary interests are data mining and machine learning, with a specific emphasis on dimensionality reduction, network embedding, and data visualization. He has co-authored papers in conferences including KDD, ICLR, ICDE, IEEE VIS, and ECML-PKDD and journals including Data Mining and Knowledge Discovery (DMKD) and IEEE Transactions on Knowledge and Data Engineering (TKDE). He was a research assistant in Fraunhofer IAIS and a research intern in Facebook AI. He frequently reviews for conferences and workshops such as ECML-PKDD, IJCAI, Discovery Science, ECAI, SIGKDD IDEA and journals such as Machine Learning Journal (MLJ) and IEEE Transactions on Knowledge and Data Engineering (TKDE).

Contact info:

Dept. of Electronics and Information Systems, IDLab, Ghent University

Email: [bo.kang@ugent.be](mailto:bo.kang@ugent.be)   
Tel: +32 485 613 164

Website: [http://bokang.io](http://bokang.io/)

------------

**Charlotte Laclau** is an associate professor at Hubert Curien Laboratory, Jean Monnet University. Her main line of research is statistical machine learning with a focus on unsupervised methods for high-dimensional and complex data. She received a grant from the IDEX program of University of Lyon to work on dynamic graph embedding and she is currently co-supervising a PhD student on fair node embedding. She has co-authored papers in international conferences including AAAI, ICML, SIGIR and journals including Data Mining and Knowledge Discovery (DMKD) and Neurocomputing. She is also a reviewer for several IA/machine learning conferences, such as IJCAI, ICML and NeurIPS.

Email: [charlotte.laclau@univ-st-etienne.fr](mailto:charlotte.laclau@univ-st-etienne.fr)

Website: https://laclauc.github.io/index.html

------------

**Christine Largeron** is a full professor in computer science. She is Professor at Jean Monnet University (France) since 2006 and, she is the head of the Data Mining and Information Retrieval group at Hubert Curien Laboratory. Her research interests focus on machine learning, data mining, information retrieval, text mining, social mining, network analysis. She regularly co-organizes on a number of international workshops, conferences on the topics of data science and network science, in particular on community detection and mining attributed networks such as SDAIN@ASONAM2013, MANeM@ASONAM2015and GEM@ECML-PKDD2019. She also gave a tutorial on Mining Attributed Networks during the World Wide Web conference 2018. She presented invited talks on related topics at School E’EGC 2018, Multidisciplinary International Social Networks Conference MISNC in (Bangkok-2017), International Workshop on High Dimensional Data Mining (Italia-2016), Workshop on Advances in Data Science (China, 2016).

She can be contacted at [Christine.Largeron@univ-st-etienne.fr](mailto:Christine.Largeron@univ-st-etienne.fr) and her web site is at <https://perso.univ-st-etienne.fr/largeron/>

[Contact info: Laboratoire Hubert Curien - UMR CNRS 5516 (LHC), Jean Monnet University, Saint-Etienne 18, rue Benoit Lauras 42000 Saint-Etienne, Tel: +33-(0)4 77 91 57 56 (57 80) Christine.Largeron@univ-st-etienne.fr](https://perso.univ-st-etienne.fr/largeron/)

------------

**Tiphaine Viard** is a postdoctoral researcher at University Paris 13. Her research interests focus on at the intersection of large complex network analysis (including their extensions : dynamic, multilayer, attributed networks) and machine learning. In particular, she is interested in what those complex networks and their models can contribute to transparency in machine learning. She has been regularly part of the organizing committee (ASONAM 2015, MARAMI 2015, AISR 2017, InWoRS 2018) and the program committee (Conference on Complex Systems 2016, DyNo 2018) of national and international workshops and conferences. She also has served as elected member at the Council of the Complex Systems society (2014-2017).

Contact info: Laboratoire d’Informatique de Paris Nord (LIPN), Université Sorbonne Paris Nord, 99 avenue Jean-Baptiste Clément, 93430 Villetaneuse; [viard@lipn.univ-paris13.fr](mailto:viard@lipn.univ-paris13.fr)

Website: tiphaineviard.com

------------

**Pascal Welke** is a postdoctoral researcher at the University of Bonn, Germany. He is interested in efficient graph mining and similarity measures, such as kernels, for graphs, but also broader topics in data mining, machine learning, and data analysis.

He has served as PC member in workshops (e.g. GEM@ECML-PKDD2019, DMLE@ECML-PKDD2019), reviewed for conferences (e.g. KDD, AISTATS) and journals (e.g. AMAI, Machine Learning), and is program chair for LWDA2020.

Contact info:

MLAI lab, CS Department, University of Bonn

Endenicher Allee 19a, 53115 Bonn, Germany

website: <https://mlai.cs.uni-bonn.de/people/pascal-welke>

email: [welke@cs.uni-bonn.de](mailto:welke@cs.uni-bonn.de)

phone: +49 228 734514

**Requirements regarding logistics:**

*-* video projector

- infrastructure to hang ~15 posters (walls, poster stands, or similar)

**Call For Papers GEM Workshop, Graph Embedding and Mining Co-located with ECML-PKDD, September 2020**

[Please accept our apologies if you receive multiple copies of this Call for Papers (CFP)]

We invite contributions in the area of exploration and analytics on graphs (algorithms, models, tools, evaluations, etc.) to be presented at the GEM workshop which is to be held at the European Conference on Machine Learning and Principles and Practices of Knowledge Discovery from Data (ECML-PKDD), in Ghent, Belgium - 14 – 18 September 2020.

**--- Topics of interests include ---**

Unsupervised and representation learning

- Network embedding

- Graph compression

- Entity resolution/deduplication

- Clustering

Pattern mining

- Local patterns: community detection, subgraph mining

- Graph summarization

- Subgroup discovery on graphs

Supervised learning

- Classification

- Link prediction

Exploration

- Visualization

- Visual analytics

Modeling

- Random graph models

- Opinion formation

- Information propagation

Data management for graph analytics

- Data models and structures

- Index structures

Applications

- Social network analysis

- Biological networks and life science data

- Communication networks

- Urban data, traffic networks

All types of approaches are welcome, e.g., graph neural networks, traditional ML, random walk methods, Bayesian inference, information theoretical approaches, and we encourage authors to consider the breadth of graph data types (attributed, dynamic graphs, etc.).

--- The workshop ---

The workshop will feature

- keynote speakers,

- a few contributed talks,

- a poster session in the afternoon, and

- small-group discussions on hot topics that are aimed to bring together academia and industry or end-users in application areas.

Keynote speakers and the program committee will be announced soon!

--- Submission Information ---

All papers will be peer-reviewed, single-blinded. We welcome many kinds of papers, such as (and not limited to):

- Novel research papers

- Demo papers

- Work-in-progress papers

- Visionary papers (white papers)

- Appraisal papers of existing methods and tools (e.g., lessons learned)

- Relevant work that has been previously published

Authors should indicate in their abstracts the kind of submissions that the paper belongs to, to help reviewers better understand their contributions. Submissions must be in PDF, written in English, not longer than 16 pages (including references and supplemental material), and formatted according to the standard Springer LNCS style. Shorter papers (including extended abstracts) are also welcome.

Accepted papers will be presented as a poster in the poster session and a few will be selected to also give an oral presentation. The papers will be posted on the workshop website (unless the authors opt out), but will not be published in a formal proceeding.

For accepted papers, at least one author must register for the conference and attend the workshop to present the work.

Submit via EasyChair: https://easychair.org/conferences/?conf=gem2019

--- Important Dates ---

Submission: June 20, 2019

Notification: July 12, 2019

Camera-ready: July 26, 2019

Workshop date: September 16, 2019

--- Further information and Contact ---

Organizers: Remy Cazabet, Bo Kang, Christine Largeron, Charlotte Laclau, Tiphaine Viard, Pascal Welke

Website URL: https://gem-ecmlpkdd.github.io

E-mail: [gem.ecmlpkdd@gmail.com](mailto:gem.ecmlpkdd@gmail.com)