

BOUNDARY-SPANNING SEMINAR SERIES

# Bridging the gap between literature data extraction and materials informatics

Seminar by Elsa Olivetti  
Massachusetts Institute of Technology



**Abstract:** Data has become a fundamental ingredient for accelerating and optimizing materials design and synthesis. Advances in applying natural language processing (NLP) to material science text has greatly increased the size and acquisition speed of materials science data from the published literature. This presentation will describe work to extract information from peer reviewed academic literature across a range of materials. Applying NLP pipelines to these types of materials science systems can be challenging due to the general schema and the noisiness of automatically extraction data. I will present data engineering techniques and discuss an optimal balance between automatic and manual data extraction.

**Bio:** Elsa Olivetti is the Esther and Harold E. Edgerton Career Development Professor in the Department of Materials Science and Engineering (DMSE) at the Massachusetts Institute of Technology. Her research focuses on improving the environmental and economic sustainability of materials in the context of rapid-expanding global demand. Dr. Olivetti received her B.S. degree in Engineering Science from the University of Virginia and her Ph.D. in Materials Science Engineering from MIT.

**9:00 a.m.**  
**October 27, 2020**

**Join us by Zoom:**

<https://kansas.zoom.us/j/94193463789>  
Passcode: 2020

[cebc@ku.edu](mailto:cebc@ku.edu) | [www.cebc.ku.edu](http://www.cebc.ku.edu)  
[www.nrt.ku.edu](http://www.nrt.ku.edu)

**Seminar overview:** this seminar series nurtures learning across academic boundaries. Expect rich conversations, spanning machine learning to catalysts, from lab bench to smokestack, with an eye toward protecting nature. Sponsoring units include: KU's Center for Environmentally Beneficial Catalysis (CEBC), the Internet of Catalysis National Science Foundation Research Traineeship (NRT), Information and Telecommunication Technology Center (ITTC), and the Departments of Chemistry, Chemical and Petroleum Engineering, and Electrical Engineering and Computer Science

**You belong here!** We invite everyone from the KU community to join us for this inclusive seminar.