Contribution ID: **18**   Type: **not specified**

Industrial Text Data Analysis: From Word Counting to GenAI Applications

Friday, 30 May 2025 15:45 (20 minutes)

In the Chemical Manufacturing Industry, a diverse array of sensor technologies and data collection methods provide valuable insights into monitoring physical and chemical phenomena, equipment status, process conditions, raw material attributes, product quality, emissions, and logistics. Despite the extensive use of sensors, critical process information such as leaks, corrosion, and insulation degradation often goes undetected, often hidden by control actions.

To address this gap, alternative information sources are essential for gaining insights into the state and health of industrial processes. One promising alternative is industrial text data found in reports, alarm tags, quality notes, and process memorandums. Although numerous text processing methods exist, their application in the Chemical Manufacturing Industry remains underexplored. This talk cover the efficacy of text processing techniques for information retrieval from industrial text data.

Industrial text data presents unique challenges for Natural Language Processing (NLP), including specific vocabulary, incomplete information, sparse domain coverage, and complex underlying phenomena. To overcome these challenges, fine-tuning NLP and Machine Learning (ML) models for specific production contexts is a potential solution. This approach requires a high-quality dataset representative of the targeted manufacturing process but poses limitations in generalizing to other processes and sites.

Generative AI (GenAI) models offer a promising solution for industrial text analysis and targeted information extraction. By serving as foundational models, GenAI can support the analysis of industrial text data. This talk will present the journey of industrial text analysis, covering from most basic concepts such as word counting to advanced ones like GenAI applications, highlighting the potential of the industrial text data analysis and including at least one case study.

Type of presentation

Contributed Talk

Primary author: Dr STRELET, Eugeniu (The Dow Chemical Company)

Co-authors: Mrs ZINK, Anna (The Dow Chemical Company); Dr CASTILLO, Ivan (The Dow Chemical Company); Prof. REIS, Marco S. (Univ Coimbra, CERES, Department of Chemical Engineering, Rua Sílvio Lima, Pólo II –Pinhal de Marrocos, 3030-790 Coimbra, Portugal); Dr RENDALL, Ricardo (The Dow Chemical Company); Dr CHIN, Swee-Teng (The Dow Chemical Company); Dr PENG, You (The Dow Chemical Company)

Presenter: Dr STRELET, Eugeniu (The Dow Chemical Company)

Session Classification: Session

Track Classification: Spring Meeting