



**Event:** WorldFAIR Chemistry Webinar: “What is a Chemical? Applying Chemical Data to Industry Challenges”. October 13, 2022. 12-1:30 pm EDT.

- 
- ***This space is to share notes/comments and discuss the topic of the event above. This document is available to be downloaded and shared.***
  - **Agenda:**
    - Flash Talks (Moderated by: Ian Bruno)**
    - 1. Artificial Intelligence and Machine Learning: **Lutz Weber**
      - BigQuery Open data set at [https://console.cloud.google.com/bigquery?project=sciwalker-open-data&organizationId=359740966731&pli=1&ws=!1m5!1m4!4m3!1ssciwalker-open-data!2schemistry\\_compounds!3soc\\_registry](https://console.cloud.google.com/bigquery?project=sciwalker-open-data&organizationId=359740966731&pli=1&ws=!1m5!1m4!4m3!1ssciwalker-open-data!2schemistry_compounds!3soc_registry)
    - 2. Artificial Intelligence and Machine Learning: **Teo Laino**
      - Text extraction of chemical reaction substances and interpretation of the workflow can then be applied to machine learning
      - Can also be used to predict routes of retrosynthesis
    - 3. Agri-Chemistry: **Yannick Djombou Feunang**
      - We need a new expressive line notation - try Wiswesser?
      - We need a IUPAC name resolver as open source software <https://github.com/Kohulan/Smiles-TO-iUpac-Translator>
    - 4. Textiles and Dyes: **Nelson Vinueza Benitez**
      - Dyes and pigments are used interchangeably but they are different types of chemicals
      - Colorant structures can have limited availability due to protection in patents
      - [Weaver's historic accessible collection of synthetic dyes: a cheminformatics analysis - Chemical Science \(RSC Publishing\)](#)
    - 5. Pharmaceutical Research Cheminformatics: **Nick Lynch**
      - In any data, it's a long tail (we never know everything at any data point)
      - Data should be born FAIR, meaning the addition of digital identifiers as metadata
      - Challenge: Research never happens in just one place
      - Also important to maintain stability of identifiers and we change data system
      - There are levels of FAIR that apply at different levels of locality

*\*WorldFAIR “Global cooperation on FAIR data policy and practice” is funded by the EC HORIZON-WIDERA-2021-ERA-01-41 Coordination and Support Action under Grant Agreement No. 101058393.*

Contact us : [FAIRChemistry@iupac.org](mailto:FAIRChemistry@iupac.org)



@FAIRChemistry



FAIRChemistry



**Event:** WorldFAIR Chemistry Webinar: “What is a Chemical? Applying Chemical Data to Industry Challenges”. October 13, 2022. 12-1:30 pm EDT.

---

- At the heart of high quality of ML and AI we need high quality data
  - 6. Pharmaceutical Informatics-Regulatory: **Gunther Schadow**
    - ISO 11283 (ICH, IDMP) defines chemical substance
    - Standardized the Drug Application Package (CTD)
      - Active ingredients
      - Inactive ingredients
      - Module 3 is where the chemical substance are discussed
    - Differences between raw chemicals and dosage form (using InChI)
    - The form of a substance is very important to automated methods
    - Names are normally from the INN
- <https://www.who.int/teams/health-product-and-policy-standards/inn>

### Panel Discussion

Lutz: Need a more expressive molecular representation -> across all nominal identifiers

Teo: Supporter of technology to process natural language in very disordered information -> explore and adopt technologies to aid chemists as they generate data...

Yannick: How can we organize and share physical property data with structures? This could lead to models that are more interoperable (represent molecular interactions). Expressive and easy to parse...

Nelson: Defining rules for structures? Make molecules available for everyone...

Nick: Best practice. Make the most of the standards we already have. We will never get a perfect standard.

Gunther: Low technical complexity, add tag, don't constrain them... Build InChI into larger structures!

*\*WorldFAIR “Global cooperation on FAIR data policy and practice” is funded by the EC HORIZON-WIDERA-2021-ERA-01-41 Coordination and Support Action under Grant Agreement No. 101058393.*

Contact us : [FAIRChemistry@iupac.org](mailto:FAIRChemistry@iupac.org)



@FAIRChemistry



FAIRChemistry