Supporting Prudent Practices in Research Laboratories

by Leah McEwen

In response to a sequence of dramatic accidents in research laboratories, [1-3] over the last five years the academic chemistry sector in the United States has focused on improving its safety culture and adopting more prudent laboratory practices. [4] Crucial to this effort is improved accessibility and management of chemical health and safety data and information tools that support laboratory process risk assessment. To this end, chemical safety, information, and education professionals are collaborating to facilitate and improve the risk assessment process for chemistry researchers and research students, as reported previously in *Chemistry International*. [5] The goal is to connect existing chemical safety and education communities and information technologies to improve *F*indability, *A*ccessibility, *I*nteroperability and *R*e-use of chemical safety data for both researchers and information systems. These are referred to as the FAIR principles for sharing research data. [6] There are a number of collaborative efforts applying this concept in order to facilitate the use of existing data, as well as to identify data gaps in the support of hazard analysis and risk assessment for academic research scenarios.

These efforts have identified a number of critical challenges to supporting safety data management at the scale and breadth needed for institutions researching novel compounds. For example:

- Pertinent data for assessing and managing risks are scattered across many agency and industry resources, nationally and internationally, and in many different formats.
- Reporting requirements vary by sector and region, and cause difficulties for exchanging and evaluating data.
- Diverse schemes are employed to identify chemicals and mixed substances are often only indexed by the primary component.
- No formal practice exists for reporting incidents to provide valuable "lessons learned" data.
- Few sources are designed for machine handling of the data to support local applications.

Addressing these challenges engages technologies and expertise from several national and international efforts in chemical data management. These include the IUPAC International Chemical Identifier (InChI); [7] the National Center for Biotechnology Information (NCBI) in the US; [8] the Pistoia Alliance; [9] the Research Data Alliance (RDA); [10] and the technical Divisions of Chemical Health and Safety (CHAS), [11] Information (CINF), [12] and Education (CHED) [13] of the American Chemical Society (ACS). [14] These tools are relevant to supporting the **R**ecognize, **A**ssess, **M**inimize, **P**repare (RAMP) paradigm recently popularized by the American Chemical Society. [15]

Preliminary analysis of the results of a recent survey of the academic chemistry community in the US points to a specific example of these problems. Of those surveyed, Safety Data Sheets (SDS) from chemical vendors were overwhelmingly the go-to source for chemical safety information. [16] There is concern, however, in the professional chemical safety community about the efficacy of using the SDS out of context to support laboratory risk assessment. While the SDS provides a general safety profile for an individual known chemical, chemistry research involves reactions among multiple compounds and substances combined in various proportions under diverse conditions. The chemical safety context is further influenced by processes within the experimental procedure, the management of substances from inventory through disposal, and setup and maintenance of the laboratory environment. Very little information relevant to this type of analysis has been incorporated systematically into the laboratory safety process.

There is a body of research dedicated to analyzing the operations and conditions of large scale chemical processes in industrial settings, where these processes are well-defined and carefully specified as part of the planning process. However, such analysis is rarely conducted for chemical procedures developed iteratively at the laboratory level as defined by OSHA regulations in the United States. [17] Very often, the most important source of relevant information is from reports of incidents where safe control was exceeded and the influence of reactivity and process factors can be considered in retrospect. However, these reports are not the focus of regular research reporting and tend to be brief mentions found sporadically in letters to editors of journals, [18] or as news items, [19] or occasionally rephrased as caution statements in vetted procedures. [20]

The value of compiling this data across many compounds is realized in the Bretherick's Handbook of

Reactive Chemical Hazards. [21] Not only does this handbook serve as a reference for individual chemicals, but similar reactivity hazards are grouped by chemical classes or atom groupings, suggesting other possible hazardous combinations. This tool is a step toward filling the significant need to project potential hazards beyond well-characterized chemicals when designing novel experiments. However, these reports are still locked in text and not easily parsed for incorporation into digital information workflows.

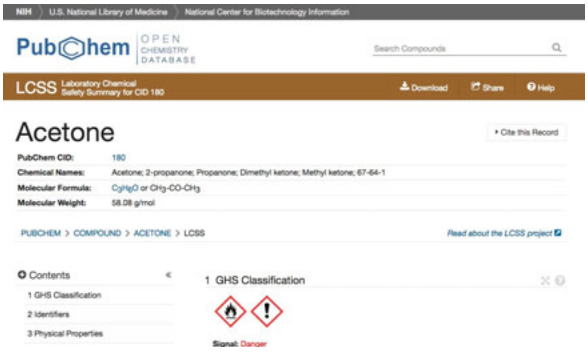
The current focus on chemical safety in research has inspired a number of additional efforts to fill this gap in experiential information. The Pistoia Alliance recently launched an open tool prototype for collecting pre-competitive reaction hazard data that can be downloaded for use in local inventory and planning systems. [22] The ACS is implementing a requirement that authors publishing in their journals include descriptions of potential hazards during their research processes that require safety measures beyond normal precautions. [23] This policy was prompted by a study of the ACS Committee on Chemical Safety last year that found very little mention of chemical safety reporting in the author guidelines of hundreds of scientific journals. [24] If other scientific societies and publishers embrace similar initiatives, reporting on chemical procedure hazards may begin to permeate the literature. The ability to make this information discoverable at the time of need will depend to some extent on the systematic description of these hazard scenarios, which will further depend on chemical safety terminology. Benefits can be realized by both traditional indexing and semantic methods.

Reframing these challenges and opportunities from a data management perspective, where machine processing is a critical requirement, raises the following questions:

1. How can chemical safety information be collated for further application?
2. How can chemical specific data be mapped to real-world chemical work?
3. How can incident data be described systematically in a laboratory context?

Collating data across sources adds value for analysis and enables researchers to more readily assess risk, especially that of novel compounds and evolving procedures. In the United States, the NCBI at the National Library of Medicine hosts the publically available PubChem database, [25] which collates public data on over 90 million compounds, semantically represented and programmatically accessible for reuse in local applications. Chemical safety information for over 100 thousand chemicals, from authoritative agency sources worldwide, is dynamically presented as Laboratory Chemical Safety Summaries (LCSS, see Figure 1 [26]), [27] based on the format described by the US National Research Council (NRC) in *Prudent Practices in the Laboratory*. [4] This data augments the information generally provided by Safety Data Sheets, and includes many chemical and physical properties, GHS Hazard Classifications, and specific examples of incompatible reactions, as sourced from the NLM Hazardous Substances DataBank, Sigma Aldrich Safety Center notes, the US National Fire Protection Agency (NFPA) documents, Sax's Dangerous Properties of Industrial Materials, and Bretherick's Handbook of Reactive Chemical Hazards, among others. Data are reported directly from multiple sources with full provenance, allowing users to make comparisons and determine their applicability to local contexts. The data are also programmatically accessible and can be provided as a stream to support integration into local systems and applications.

Figure 1. PubChem's Laboratory Chemical Safety Summary for Acetone [26]



In the safety context, hazards can be presented by any component in a chemical system, such as solvents. In practice, no compound is pure and composition impacts chemical reactivity, unintentionally or by design. Characterization of components provides better experimental results, in addition to safer laboratory functions. Managing information about mixed substances becomes a significant challenge in a world of 100+ million characterized compounds. In the real world, many articles of trade exist as defined or partially defined mixtures, yet often only the components of interest are noted. The identification of chemical entities involved in multi-component systems could enable the connection of safety data, predominately ordered

by individual chemicals, to commercial products and real-world chemical work. For these reasons, there is a project underway to leverage the IUPAC InChI algorithm to articulate the defined chemical aspects in a mixture system (see Figure 2). [28-29]

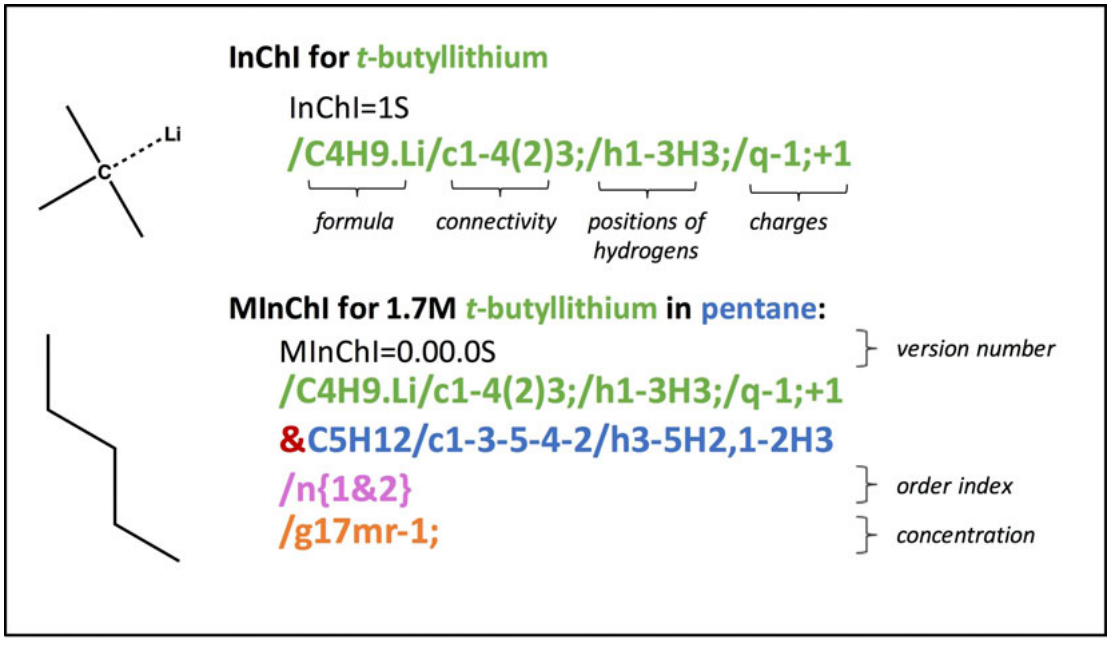
There are many factors at play in conducting a laboratory procedure that may contribute to the potential risk of a given situation. The safety community is looking to practices in industry to formalize useful approaches to risk assessment in the research laboratory context. [30-32] Analyzing procedures and coupling these with incident data can potentially bring to light incompatible combinations and problematic operations, as well as aid in planning for adjustments to experimental parameters. This opportunity can be greatly informed by tapping into collective information across many documented processes. Domain terminology that describes key factors can enable the systematic analysis of relationships, such as combinations of chemicals, or substances under different conditions. This approach has been used for single analysis of M/SDS documents, [33] and chemical procedures. [34] Melding these vocabularies to consider data from incidents more systematically and across compiled reports is the forefront of another community effort. [35]

The safety data ecosystem is exceedingly complex and a classic management challenge, involving many different uses, applications, stakeholders, expectations, and requirements. The scope is dynamic and exceeds the information space of well-documented chemicals. Several hundreds of thousands of different substances, involving hundreds of sources and thousands of users, may be present at a large research campus. Turning this data into improved safety in practice is a community challenge, engaging the expertise and dedication of a broad range of professionals in the research, education, and service sectors. The good news is that this challenge is increasingly recognized by the chemistry community and emerging information sources and tools show important potential in helping to address this opportunity for improvement.

References

1. http://cenblog.org/the-safety-zone/2010/11/ucla-calosha-settlement
2. http://cenblog.org/the-safety-zone/2017/01/civil-suit-filed-against-university-of-hawaii-for-lab-explosion
3. http://cenblog.org/the-safety-zone/2014/11/webinar-on-u-minnesota-azide-explosion-on-dec-3/
4. National Research Council. *Prudent Practices in the Laboratory: Handling and Management of Chemical Hazards, Updated Version*; The National Academies Press: Washington, DC, 2011. https://doi.org/10.17226/12654
5. McEwen, L. R. and Stuart, R. B. Meeting the Google Expectation for Chemical Safety Information. *Chemistry International* **37**(5-6):12-16, 2015. http://dx.doi.org/10.1515/ci-2015-0505
6. www.force11.org/group/fairgroup/fairprinciples

Figure 2. InChI Extension for Mixture Composition (Draft Specification)



7. www.inchi-trust.org
8. https://pubchem.ncbi.nlm.nih.gov/lcss
9. www.pistoiaalliance.org/projects/chemical-safety-library
10. www.rd-alliance.org/groups/chemistry-research-data-interest-group.html
11. http://dchas.org
12. http://acscinf.org
13. www.divched.org
14. www.acs.org/safety
15. Stuart, R. B. and McEwen, L. R. The Safety "Use Case": Co-Developing Chemical Information Management and Laboratory Safety Skills. *J. Chem. Ed.* **93**(3):516–526, 2016. https://doi.org/10.1021/acs.jchemed.5b00511
16. http://dchas.org/2017/04/05/safety-training-content
17. Occupational exposure to hazardous chemicals in the laboratory. 29 CFR 1910.1450. www.osha.gov/pls/oshaweb/owadisp.show_document?p_id=10106&p_table=STANDARDS
18. http://pubs.acs.org/cen/safety
19. https://dchas.org/the-dchas-l-list
20. http://cenblog.org/the-safety-zone/2016/02/oprds-safety-notables-from-the-literature/
21. Urben, P. (ed.) *Bretherick's Handbook of Reactive Chemical Hazards (8th Edition)*; Elsevier: Amsterdam, 2017. www.elsevier.com/books/brethericks-handbook-of-reactive-chemical-hazards/urben/978-0-08-100971-0
22. www.pistoiaalliance.org/projects/chemical-safety-library
23. Bertozzi, C. R. Ingredients for a Positive Safety Culture. *ACS Cent. Sci.* **2**(11):764–766, 2016. https://doi.org/10.1021/acscentsci.6b00341
24. Grabowski, L. E. and Goode, S. R. Review and analysis of safety policies of chemical journals. *J. Chem. Health. Safety* **23**(3):30-35, 2016. https://doi.org/10.1016/j.jchas.2015.10.001
25. https://pubchem.ncbi.nlm.nih.gov
26. National Center for Biotechnology Information. PubChem Compound Database; CID=180, datasheet=lcss,https://pubchem.ncbi.nlm.nih.gov/compound/180#datasheet=lcss (accessed May 3, 2017).
27. https://pubchem.ncbi.nlm.nih.gov/lcss
28. InChI Extension for Mixture Composition, https://iupac.org/project/2015-025-4-800
29. Kemsley, J. N. The United Nations of Chemistry. *Chem. Eng. News* **95**(10):25-26, 2017. http://cen.acs.org/articles/95/i10/United-Nations-chemistry.html
30. American Chemical Society, Committee on Chemical Safety, *Hazard Identification and Evaluation Task Force. Identifying and Evaluating Hazards in Research Laboratories*; American Chemical Society: Washington, DC, 2015. www.acs.org/content/dam/acsorg/about/governance/committees/chemicalsafety/publications/identifying-and-evaluating-hazards-in-research-laboratories.pdf
31. www.acs.org/hazardassessment
32. Mulcahy, M. B., *et al.* Using bowtie methodology to support laboratory hazard identification, risk management, and incident analysis. *J. Chem. Health. Safety* **24**(3):14-20, 2017. https://doi.org/10.1016/j.jchas.2016.10.003
33. www.ilpi.com/msds/ref/demystify.html
34. http://chemicaltagger.ch.cam.ac.uk
35. https://dchas.org/2017/04/05/information-flow-in-environmental-health-safety

Leah McEwen <lrm1@cornell.edu> is chemistry librarian at Cornell University, USA. She is a member of the IUPAC Committee on Publications and Cheminformatics Data Standards (CPCDS), and co-chair of the CPCDS Subcommittee on Cheminformatics Data Standards. ORCID.org/0000-0003-2968-1674