The Safety Use Case for Chemical Safety Information

Ralph Stuart
Chemical Hygiene Officer, Keene State College
PS: Alas, in these modern times, these young folks just don't know their descriptive chemistry like us old guys do. I predict disaster and catastrophe - as we old guys die off, the world will be left with chemists who don't know descriptive chemistry. Alas and Alak! (or whatever)

- Jay Young

My version of this statement is that 21st Century Chemistry Education focuses on discovery processes, using specific chemical techniques and information management. This does not necessarily translate into transferable bench-top skills; of particular concern is safety skills.
The Continuing Evolution of Science

Lab science in the 21st Century is an emerging complex system which highly values converging knowledge.

Changing Science Learning Styles

Science Paradigms

- Thousand years ago: science was empirical describing natural phenomena
- Last few hundred years: theoretical branch using models, generalizations
- Last few decades: a computational branch simulating complex phenomena
- Today: data exploration (eScience) unify theory, experiment, and simulation
  - Data captured by instruments or generated by simulator
  - Processed by software
  - Information/knowledge stored in computer
  - Scientist analyzes database/files using data management and statistics

FIGURE 4-2. The web of faculty interactions created by Bio-X. The network of faculty interactions across Stanford has expanded since the establishment of the Bio-X program. The resulting network reportedly appeals to technology com-
The Current Context of Chemical Safety Information

- **Caveat emptor**: Chemistry textbooks and laboratory manuals provide a overview of generic rules, followed by "see the MSDS".
- Wikipedia provides links to random MSDS sources with no evidence of why that source is selected; some sources are kaput, many are dated.
- Reports by the National Research Council, the ACS, NFPA after specific laboratory safety incidents found this approach to chemical safety education and information inadequate.
Use Cases to Consider

• **Teaching laboratory setting:** short term use of specific chemical concentrations in procedures with expected outcomes; close oversight of inexperienced lab workers by experienced personnel can be assumed.

• **Research laboratory setting:** evolving use of chemicals with uncertain process outcomes for lengths of time determined by results of work; diverse group of lab workers with loose supervision by experienced personnel.

• **Service laboratory setting:** long term use of specific chemicals in similar processes with reproducible outcomes on an long term basis.
Non Lab Use Cases Identified by *Prudent Practices*

- **Non Lab Use Cases**
  - Household use of commercial chemical products
  - Large scale manufacturing use of chemicals
  - Medical aspects of long term exposures (drug use, etc.)
  - Environmental impacts
  - HAZMAT response
  - Transportation considerations
The Lab Use Case Questions

• Does the use of this chemical require the use of a fume hood or other local ventilation system?
• What PPE is appropriate for the use of this chemical?
• What waste disposal protocols are required to legally dispose of this chemical?
• Are unusual emergency response protocols necessary for work involving this chemical?
• Are the specific chemical reactivity hazards associated with the use of this chemical that all users should be aware of?
The Structure of Chemical Safety Information

Increasingly well-defined; also increasing voluminous

Chemical Safety Information Source Hierarchy

- 1990's: Basis for Data Provided
- 2000's: NRC Examples
- 1980's: GHS Definitions
- 1970's: Suppliers' Legal Requirements
- Raw Data in Disparate Sources
- MSDS's
- GHS SDS and labels
- LCSSs
Less Structured Chemical Safety Information

Scientific Audiences (Principal Investigators)
- Flashpoint
- LD50's
- Odor threshold

Bench Chemists (lab staff)
- Flinn's List of the “40 Devils”
- Not Voodoo “Rookie Mistakes”
- Word of Mouth

Lab Lore and Lessons Learned

PubChem data collection and data views

Wikipedia ChemBoxes

Process Oriented

External Audiences (Funders and Regulators)
- Emergency Response Plans
- Waste Disposal Programs
- Ventilation Support
- General training Support

Chemical Based

The General Public (Students)
- GHS statements
- NFPA diamonds
- Miscellaneous notes in the wiki article
Looking for Structure in the Electronic Data

- How large is the PubChem chem safety information universe?
- How high quality is it (including consistency & provenance information)?

Key safety information fields:
- GHS class designation(s) and signal word
- NFPA diamond information

PubChem data collection and data views

- Millions of chemicals; 3500 with GHS info

How much overlap is there between the two?

Wikipedia ChemBoxes

- How large is the Wikipedia chem safety information universe?
- How high quality is it?
Objective is to analyze how many chemicals in the PubChem LCSS have GHS and NFPA data in the Wikipedia Chembox.

Download list of Chemicals from PubChem Containing LCSS

Use Excel to filter the download of information to only PubChem CIDs

Use Google Sheets to pull NFPA 704 data from Wiki Chembox for Wiki Names

Use Google Sheets to convert InChI Keys to Wiki Names

Use PubChem Identifier Exchange to convert List of CIDs to InChI Keys

Use Google Sheets to pull GHS Hazard Statements from Wiki Chembox for Wiki Names

Use Google Sheets to analyze data

Future Goals to compare Safety Information of Wikipedia to PubChem
LCSS Data to InChI Key

1. PubChem Website
2. Obtain List of Chemicals with LCSS
3. Excel or Google Sheets
   - Filter down data from PubChem into a List of PubChem CIDs
4. PubChem Identifier Exchange
   - Convert List of PubChem CIDs to InChI Keys
InChI Key to Safety Information

1. Convert InChI Key to Wiki Name using Importxml function
2. Convert Wiki Name to Safety Information using Importhtml function
3. Analyze Data
The Early Statistics

- PubChem has an LCSS view for about 3500 (2000 more to come soon) chemicals; Wikipedia has Chemboxes for about 10,000 chemicals
- Of those in the PubChem LCSS collection, about 30% have an entry in Wikipedia
- 4% of the PubChem collection has GHS information; 12% of the PubChem collection have NFPA diamond information

<table>
<thead>
<tr>
<th></th>
<th>Not in Wikipedia</th>
<th>In Wikipedia</th>
<th>GHS Hazard Statement</th>
<th>NFPA 704</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>2441</td>
<td>1038</td>
<td>157</td>
<td>431</td>
<td>3486</td>
</tr>
<tr>
<td>%</td>
<td>70.02%</td>
<td>29.78%</td>
<td>4.50%</td>
<td>12.36%</td>
<td></td>
</tr>
</tbody>
</table>
Future Directions

- Understand the Wikipedia Chembox structure to collect information more efficiently.
- Develop a Wikipedia – PubChem data link that can provide chemical safety information with provence data to the ChemBox.
- Consider what chem safety data makes sense to put in the Wikipedia Chembox and what can be linked to there.

The hazard portion of the acetone chembox:

- **Hazard Data Sheet**: See: data page
- **GHS Pictograms**
  - GHS signal word: DANGER
  - GHS hazard statements: H225, H319, H336
  - GHS precautionary statements: P210, P261, P305+351+338
- **EU classification (DSD)**: Flame, Fire, Exp
- **R-phrases**: R11, R36, R66, R67
- **S-phrases**: S2, S9, S16, S26
- **NFPA 704**
  - Flash point: -20 °C (-4 °F; 253 K)
  - Autoignition temperature: 465 °C (869 °F; 738 K)
  - Explosive limits: 2.6–12.8%[10]
  - Threshold Limit Value: 1185 mg/m³ (TWA), 2375 mg/m³ (STEL)
  - Lethal dose or concentration (LD, LC):
    - **LD₅₀ (Median dose)**: 5800 mg/kg (rat, oral) 3000 mg/kg (mouse, oral) 5340 mg/kg (rabbit, oral)[11]
    - **LC₅₀ (Median concentration)**: 20,702 ppm (rat, 8 hr)[11]
Closing Thoughts

**The Medium is the Message:**
the form of a medium embeds itself in the message, creating a symbiotic relationship by which the medium influences how the message is perceived. (Wikipedia)

"Wikipedia is the last refuge of the Internet optimist"
Christopher Lydon

![Ethanol diagram](image)