ELEMENTA
Science of the Anthropocene

Kokai, A, et al. 2020. Building shared information infrastructure for chemical alternatives assessment. *Elem Sci Anth,* 8: 26. DOI: https://doi.org/10.1525/elementa.422

PRACTICE BRIDGE

Building shared information infrastructure for chemical alternatives assessment

Akos Kokai*, Ann Blake†, Michel Dedeo‡ and Tom Lent‡

The substitution of hazardous substances with safer alternatives is being driven by policy pressures and business demands. As a result, scientific techniques for chemical alternatives assessment (CAA) have been established and communities of practice are emerging. Interest in safer chemical substitution is widely shared throughout a range of stakeholder groups across science, industry, public policy, and advocacy. Yet there is an unmet need for intentionally designed public information infrastructure to support the highly knowledge-intensive nature of CAA. We report here on the process of developing the Chemical Hazard Data Commons, an experimental project intended to support a diverse community of practitioners by providing publicly accessible chemical hazard data and tools for understanding it. In an arena where market forces and regulatory regimes have largely failed to generate the necessary knowledge, this project represents a novel application of a commons-based approach emphasizing building shared intellectual and technical capacity for CAA. The Data Commons—now a part of the related Pharos Project—includes an online portal providing simultaneous access to many different sources of information and enabling effective interactions with it. Foremost among these interactions are search and retrieval of hazard information about chemical substances, uniform display of the most relevant information, and the ability to automatically screen substances against consistent and transparent hazard-based criteria. We describe the motivation for the project and report on the principles and key considerations that guided its design as a participatory information infrastructure. We present our approach to organizing chemical information; the process of community engagement and planning; and how we constructed the system to provide functional tools. We discuss the outcomes of the project and highlight important challenges—such as fostering active participation and planning for long-term governance. With this article, we hope to inform future efforts for the collaborative development of knowledge resources for chemical alternatives assessment.

Keywords: Green chemistry; Alternatives assessment; Information infrastructure; Knowledge commons; Shared resources; Tools

Introduction

The persistent problem of toxic substances in our industrially transformed environment cannot be solved using the same approach that produced it. Material and product design often fails to consider human and environmental health; manufacturers typically aim to comply with regulatory requirements or—at most—to avoid limited "red lists" of the most harmful substances. These approaches fail to respond to early indicators of harm, and do nothing to guide the development of safer alternatives. Risk assessment and management strategies, when invoked, favor minimizing risk by managing exposure to toxics rather than through the design of inherently safer material technologies.

Faced with these dilemmas, decision-makers are increasingly seeking a different path. New strategies for protecting public health and the environment are gravitating toward the paradigm of alternatives assessment (O'Brien 2000; Ashford 2005)—a problem-solving approach that encourages minimizing harm by comparing several possible options and incorporating a wide range of knowledge, seeking to facilitate decision-making despite uncertainty. The goal of using alternatives assessment is to enable the informed substitution of hazardous chemicals and materials with safer alternatives (Toxics Use Reduction Institute 2006; Zimmerman and Anastas 2015). This goal has been integrated into a wide range of efforts in policy, design, engineering, and business. Chemical alternatives assessment (CAA) refers to techniques for comparing characteristics of substances to select the safest option, and to identify data gaps and further research needs (Lavoie

* Department of Environmental Science, Policy, and Management, University of California, Berkeley, California, US
† Environmental and Public Health Consulting, Alameda, California, US
‡ Healthy Building Network, Berkeley, California, US
Corresponding author: Akos Kokai (akokai@berkeley.edu)

et al. 2010; Harrison and Hester 2013; National Research Council (US) 2014). Chemical hazard assessment (CHA) refers to techniques for systematically evaluating the types and severity of harm a substance might cause, as well as evaluating the strength of evidence and the sources of uncertainty in the data (Heine and Franjevic 2013). These two sets of techniques are related, with CHA informing and being a necessary tool for CAA (Geiser 2015, 257).

Scaling up CAA requires developing information resources and tools to address key knowledge challenges. A growing range of professionals—such as designers, engineers, and architects—find themselves facing demands to make informed decisions about chemicals (Scruggs 2013; Logan 2016). This could bring great diversity of thought to one of the major sustainability challenges that we face as a society. At the same time, decision-makers face practical challenges in obtaining, evaluating, and operationalizing information about chemicals and their health impacts. Data about chemical substances—their hazard traits, technical uses in products, and sources of human exposure—are often insufficient due to a lack of scientific research and public disclosure (Wilson and Schwarzman 2009). Existing data may come from numerous disparate sources with varying standards of organization and interpretation. Many decision-makers engaged with chemical substitution problems aren't experts in chemistry and toxicology, or they don't have the organizational capacity to access and use detailed information about chemicals. Using CAA therefore tends to be technically demanding, resource-intensive, and costly.

Much duplicative work has been done in the parallel efforts of firms, government agencies, and non-governmental organizations to satisfy the interrelated knowledge needs of many stakeholders. There are dozens of different guidelines, ecolabels, chemical "red lists," and databases, but they lack consistency and mutual compatibility. The high cost and proprietary nature of professional chemical hazard assessment services create barriers to accessing and distributing much of the most actionable and valuable knowledge. As a result, decision-makers may find that there are several tools to address the same problem, but few ways to share information, access a pool of existing knowledge, or reconcile inconsistencies and contradictions between different resources.

Taken together, this array of challenges and systemic inefficiencies motivated us to work on improving publicly accessible information infrastructure[1] for chemical alternatives assessment. We report here on the development of the Chemical Hazard Data Commons, a project intended to support a diverse community of practitioners by providing publicly accessible chemical hazard data and tools for understanding it. This project represents a novel application of a collaborative, commons-based approach, and the first effort to create an open, participatory information infrastructure for CAA. In an arena where decades of adversarial policymaking and activism around toxic chemicals have produced intensely contested scientific knowledge, the Data Commons approach emphasizes the value of shared information resources and collective capacity-building. The project aimed to elicit knowledge, expertise, and collaborative peer-review from its participants. The project, led by US-based NGO Healthy Building Network (HBN), began in 2013 and concluded in 2019 with the integration of the Data Commons' newly developed tools and resources into HBN's ongoing Pharos Project (Healthy Building Network 2019d). Pharos, which is accessible online with free user registration, now serves the intended purpose of the commons: a publicly accessible resource for the chemical alternatives assessment community (Healthy Building Network 2020b).

We do not claim that the work reported here is, or will be, the single or best realization of its goals. Rather, we hope to contribute to a larger ongoing enterprise—the collaborative development of knowledge resources to support CAA.

Overview of the Data Commons project

The Chemical Hazard Data Commons project aimed to freely and openly provide robust, collaboratively-curated information resources for CAA. We intended to do this by compiling, organizing, and linking together existing publicly available information using open standards; developing online tools for making use of this wide array of information; and coordinating community efforts to increase the knowledge base available to participants and to the public. In essence, the commons has two key elements: networked knowledge resources and a community, bound by common principles, that collaborates to support the development, improvement, and assessment of these resources.

A central component of the commons is a web-accessible library of hundreds of thousands of chemicals with known hazards categorized in terms of 22 specific human and ecological health endpoints (see **Table 1**). Hazard information is drawn from over 40 authoritative sources representing rigorous assessments of scientific evidence, and from other publicly available sources such as professional hazard assessment reports, health studies, and community-vetted datasets. Information about the functional uses[2] of substances is drawn from government and industry sources, along with other relevant information—including physicochemical properties, industrial releases, and "preferred" lists of safer substances. **Tables 2, 3** and **4** include representative examples of data sources (for a full list, see the Pharos system documentation: Healthy Building Network 2019c). The system is searchable by chemical names, synonyms and identifiers[3] as well as by keywords indicating product categories and functional use. All data from external resources are integrated or linked in a transparent way, such that users can easily access the sources of cited information.

Besides serving as an up-to-date reference source and knowledge repository, the Data Commons was designed as a practical tool to support CAA. The online system provides rich hazard profiles for chemical substances and classes, and it enables side-by-side comparisons between substances on the basis of hazard. It also provides a screening tool to identify known high-hazard substances based on aggregated information sources. The technical basis for organizing, evaluating, and presenting hazard information is derived from the GreenScreen for Safer Chemicals, an open-source method for chemical hazard

Table 1: Taxonomy of hazard endpoints used in the Chemical Hazard Data Commons[11]. DOI: https://doi.org/10.1525/elementa.422.t1

Grouping	Endpoints
Group I human	Carcinogenicity Mutagenicity Reproductive toxicity Developmental toxicity Endocrine activity
Group II human (single exposure)	Acute mammalian toxicity Systemic toxicity/Organ effects, single exposure Neurotoxicity, single exposure Eye irritation Skin irritation
Group II* human (repeated exposures)	Systemic toxicity/Organ effects, repeated exposures Neurotoxicity, repeated exposures Respiratory sensitization Skin sensitization
Ecotoxicity	Acute aquatic toxicity Chronic aquatic toxicity Terrestrial ecotoxicity
Physical hazard	Flammability Reactivity
Environmental fate	Persistence Bioaccumulation Persistent bioaccumulative toxicants (PBT) Global warming Ozone depletion

assessment developed by Clean Production Action and widely recognized throughout the CAA practitioner community (Heine and Franjevic 2013; Clean Production Action 2020a). GreenScreen's rich underlying technical framework ensures that chemical hazard profiles can be evaluated systematically and compared on an endpoint-by-endpoint basis, making it a valuable standard for CAA. Incorporating elements of the GreenScreen framework into the Data Commons allows the system itself to perform some of the knowledge management and computation work that is involved in applying the GreenScreen method.

Finally, the Data Commons was designed as a place for community interaction around CHA and CAA topics. Discussion forums enable a lively exchange of ideas and expertise, with users posting questions and answers on a variety of technical, policy, and other issues. Discussion posts can be linked to specific chemical substance records, allowing users to "annotate" the information presented in the commons. Collaborative projects intended for community participation also have their own forums. These current forms of community interactivity are a first step toward building a living, community-managed knowledge hub.

Background and rationale

The Data Commons was driven by the needs of a community of people who practice CHA and CAA. The knowledge users and producers that we refer to as *practitioners* include a range of people with research, design, and decision-making roles in policy, business, and advocacy contexts. They have a variety of professional roles, such as: advocates analyzing chemical production and use; government scientists researching substances or products for regulatory or administrative decision-making; consultants and chemical "profilers" doing background work to support detailed chemical or product sustainability assessments; company staff who manage chemical use in manufacturing and supply chains; and designers researching the material hazards of products and materials.

In consultation with this practitioner community we identified a number of key problems, outlined here, which motivated our efforts to build the Data Commons (Lent et al. 2014).

Knowledge challenges

Historically, many kinds of knowledge that are now recognized as crucial for CHA and CAA have been under-produced, contested, and carefully guarded by industry and governments. This includes scientific data about chemical properties, hazards, human exposures, environmental emissions, and the specific ways that chemicals are used in industry and consumer products. This is partly a result of the public policy and intellectual property regimes operating throughout the 20th century. For a long time, industrial chemical manufacturers had neither requirements nor incentives to test chemicals for toxicity, and they have claimed much of what is known as trade secrets—cutting off the flow of information to the public domain and even to downstream businesses (Wilson and Schwarzman 2009; Scruggs et al. 2014). Scientists have found significant gaps in knowledge about chemical hazards and human exposures (Judson et al. 2009; Egeghy et al. 2012), and intellectual property rights continue to obscure much of the regulatory data collected by public agencies (Gilbert 2016; Schwarz and Denison 2018). Toxicity data may exist for many chemicals but are mainly circulated among industry, private labs, and government agencies. Similarly, information about how chemicals are used—in what products and for what functions—is largely kept confidential or collected only in proprietary databases.

There are many publicly available *sources* of information about chemicals, but also many barriers to their use. One of the project goals was to make these public sources more useful to decision-makers. Individually, each of these sources generally covers only a small proportion of the chemicals in commerce, and each source typically provides only a very limited kind of knowledge (for example: identifying substances associated with one type of health hazard, out of many). Techniques for CAA require practitioners to retrieve and organize many diverse sources and heterogeneous forms of scientific and technical data. Publicly available data sources—such as government lists, databases, and scientific literature—are numerous and appear in an array of digital formats without any consistency. These separate "data silos" are not easily accessible and searchable as a whole. Gathering, transforming, and integrating such information into comprehensive and searchable datasets is tedious and resource-intensive. The

underlying data are constantly being updated, making the task a never-ending maintenance process.

A related problem is the lack of standardization and compatibility between information systems, tools, and other resources for chemical alternatives assessment. For example, there are many ways to define and categorize hazards; many ways to describe the functions of chemicals within products or formulations; and different ways to integrate diverse types of data into evaluative frameworks (Harrison and Hester 2013; Jacobs et al. 2016). For non-experts, this makes it difficult to select, interpret, and compare information. In general, inconsistencies between systems make them less interoperable and less cohesive as a toolbox of knowledge resources for CAA (a problem that has led to efforts for "harmonization," for example in the building industry: L. Heine et al. 2013; Bobenhausen 2016). Similarly, there has been a tendency for database- and tool-building efforts to proliferate rather than converge. For example, multiple organizations have more or less duplicated each other's work in compiling and organizing public chemical hazard information, or producing lists of chemicals of concern—including government agencies, NGOs, product design companies, and firms offering the resulting information resources as for-profit services (Stone and Delistraty 2010; Scruggs 2013).

Examining all these issues in a discussion paper written at the outset of the Data Commons project, we concluded that there was a need for "more efficient, affordable, effective, and consistent" tools for chemical hazard assessment (Lent et al. 2014). The knowledge challenges we have highlighted make CHA and CAA costly and resource-intensive, setting up barriers to safer chemical substitution. If information infrastructure could more thoughtfully and effectively meet the needs of the practitioner community, then these costs and barriers could be reduced.

Chemical hazard data as a commons

We framed our work as building a commons, focusing on how the community of chemical alternatives assessment practitioners can build more effective ways to use, manage, and share knowledge. A commons is a collectively-owned resource, managed or produced by the community of people who also benefit from its use (Hess and Ostrom 2007; Bollier 2014). In a science/policy arena that doesn't typically encourage collaboration, the commons invites us to pursue collective solutions. While no single resource can satisfy all stakeholders' knowledge needs, what if many independent resources could interoperate smoothly? What if disparate information sources were linked, eliminating the redundant work of compiling them? What if many sources of data could be understood together under a common framework, enabling more consistent and comprehensive approaches to hazard assessment and alternatives assessment? What if people from a range of historically separate stakeholder groups (such as product designers and NGOs) could collaboratively develop the intellectual and technical capacity to advance CAA?

A commons approach could reinforce the collaborative ethos of the practitioner community. Already, many people and organizations contribute to making knowledge and tools available to support better decision-making about chemicals—and, likewise, many people and organizations benefit from this knowledge. A commons, by design, can harness the collective expertise and effort of many individuals and organizations. In imagining the Data Commons as a participatory and community-driven project, we took inspiration from open-source software commons (Schweik and English 2012) and from numerous contemporary ideas that highlight the generative results of participatory knowledge communities: peer production (Benkler 2006), user-centered innovation (von Hippel 2006), and networked science (Nielsen 2012).

The commons framing also reinforces a widely-held principle central to effective chemicals policy and product stewardship: that information about the toxicity and environmental hazards of chemical substances should be publicly available in reliable and accessible forms (e.g. Guth, Denison, and Sass 2007; Geiser 2014; Lent and Stamm 2019). Put another way, everyone who uses or makes decisions about chemical substances has the *right to know* if there may be health risks, and how to reduce those risks. This principle is enshrined in, for example, the Dubai Declaration on International Chemicals Management (United Nations Environment Programme 2006, 9); California's Proposition 65, which was passed by 2/3 of the state's voters in 1986 (OEHHA 2015); and the U.S. Emergency Planning and Community Right-to-Know Act, which created one of the most effective information-based regulatory regimes for pollution prevention (United States Congress 1986; Fung and O'Rourke 2000). There is real and strong demand from a diverse array of stakeholders for relevant and credible scientific knowledge that can inform safer chemical substitution (Scruggs and Ortolano 2011). Finally, chemical hazard information is basic scientific knowledge and arguably should be considered a public good (Stiglitz 1999)—in one way or another, it is part of a global common pool of knowledge that should be available for new science and innovation to build upon.[4]

Project history

The concept of a chemical hazard data commons emerged from the work of participants in a 2012 conference called Building a Chemical Commons, organized by BlueGreen Alliance, Clean Production Action, and the Lowell Center for Sustainable Production.[5] These participants included government agency staff, academics, environmental consultants, labor union representatives, consumer product manufacturers, philanthropic funders, and environmental NGOs. One outcome of this conference was the Commons Principles for Alternatives Assessment (BizNGO 2013), a consensus statement of the role of CAA in safer substitution. Subsequently, a Data Commons working group formed to pursue one of three streams of work originating from the conference—the others focusing on developing CAA methods and organizing the community of practice. Healthy Building Network (HBN) secured funding from the Forsythia Foundation to carry out an initial scoping project to explore what it would take to develop a chemical data commons. Begun in 2013, this project resulted in a set of discussion papers (Lent et al.

2014; Lent 2014b, 2014a; Blake 2014; Kokai, Lent, and Dedeo 2014). HBN has since led the development of the Data Commons, supported by the funding sources listed at the end of this article.

Initial development of the Data Commons' web and data infrastructure by HBN began after collecting feedback on the discussion papers (Healthy Building Network 2014b, 2014a). In 2016–2017, HBN invited 175 participants to join a pilot project in which they could use an early development version of the Data Commons and give feedback through a community survey (Healthy Building Network 2016). The systems and data were then developed continuously, with ongoing use and testing by this initial group. In 2017, the commons was opened to the public; development has continued since then. Approximately 1500 users registered before the Data Commons was merged into the Pharos Project, which itself has several thousand users.

Design of the Data Commons

We aimed to overcome pervasive "information silos," enable simultaneous access to many different sources of information, and facilitate more effective user interactions with those combined sources. Foremost among these interactions are: search and retrieval of relevant chemical hazard information; clear and succinct display of the information most relevant to hazard assessment; and the ability to automatically screen a substance—or many substances simultaneously—against consistent and transparent hazard-based criteria. Furthermore, the system would have to be broadly accessible to a wide range of users, participatory, and able to sustain a community.

Hazard as an organizing principle

Organizing principles dictate how resources—such as data and visualized information—are intentionally arranged to enable particular kinds of interactions (Glushko 2013). For example, chemical identity is an organizing principle of chemical information systems, meaning that those systems may enable retrieving, categorizing, and comparing information by substance.[6] Chemical hazard is an organizing principle of the Data Commons information system. This is fundamental to how it can support CAA, and sets it apart from other public information systems in the domains of chemistry, sustainability, and environmental health sciences.

Hazard—the potential for a chemical substance to harm human or environmental health—encompasses multiple "endpoints" (outcomes) for a variety of species and systems. Accordingly, information system design should reflect this complexity (Lavoie et al. 2010; Whittaker and Heine 2013; Jacobs et al. 2016). At the same time, the system should make indicators of hazard understandable for users with a range of expertise in CHA. For example, users should be able to distinguish between chemicals of high, moderate, and low concern; they should be able to tell why substances may be of concern (i.e. for what hazard properties and endpoints) and based on what specific evidence. Furthermore, users should be aware of uncertainties and gaps in the evidence of hazard. In other words, the system should provide varying levels of data interpretation, using consistent and transparent methods.

Realizing this design calls for a number of elements: a taxonomy of hazard endpoints, systematic ways to evaluate and represent the severity of hazard and the strength of scientific evidence across endpoints, and ways of classifying the overall level of hazard for a substance. Several existing chemical assessment frameworks provide these elements: the GreenScreen for Safer Chemicals (Clean Production Action 2020a), the Globally Harmonized System (United Nations 2019), the US EPA Safer Choice Criteria (US Environmental Protection Agency 2012), and the Cradle to Cradle Certified Material Health Assessment Methodology (Cradle to Cradle Products Innovation Institute 2019). These frameworks are comparable and interrelated, but we chose to adopt the GreenScreen as an underlying organizing system for hazard data.[7] The taxonomy of hazard endpoints is shown in **Table 1**. GreenScreen was selected for its "open-source" and peer-reviewed technical methodology; evaluative system that provides substance-level and endpoint-level indicators of hazard; and widespread adoption across government, industry, research, and advocacy communities—especially in our practitioner community. Furthermore, GreenScreen includes a widely used protocol called the GreenScreen List Translator (GSLT), for identifying high-hazard substances based on information from authoritative lists, such as those included in the Data Commons (**Table 2**).

The GreenScreen framework became the information infrastructure that shaped the Data Commons' selection, organization, interpretation, and visual display of hazard data. Each substance profile presents an interactive hazard summary table (**Figure 1**) that shows an overall substance-level hazard score, as well as the endpoint-specific indicators that underlie that substance-level score. The endpoint-specific indicators use consistent visual conventions to communicate characterizations of hazard levels, data availability (or data gaps), and strength of evidence. If a full GreenScreen assessment is publicly available for the substance, then its results are shown, including the GreenScreen Benchmark score (a roll-up of the full assessment). If there is no public full assessment, then the table shows endpoint-specific hazard levels determined from authoritative public data sources using the GSLT protocol, as well as the overall GSLT score for the substance. Extending this paradigm to enable more useful interactions for CAA practitioners, the Data Commons also provides a hazard comparison tool. Users can create, share, import, and export customized sets of substances for automated screening and side-by-side comparison—the hazard profiles are displayed together in a matrix that enables comparative analysis of safer alternatives and data gaps (**Figure 2**).

Incorporating functional use

Knowing the specific technical role or purpose a substance serves in a material, product, or manufacturing process is essential for identifying feasible technological alternatives (Tickner et al. 2015). We aspired to provide well-organized functional use information alongside hazard information

Table 2: Representative chemical hazard and restricted substances information sources included in the Chemical Hazard Data Commons. DOI: https://doi.org/10.1525/elementa.422.t2

Title	Hazard endpoints	Source
GreenScreen hazard assessments	All	Various (primary sources)
IARC Monographs on the Evaluation of Carcinogenic Risks to Humans	Carcinogenicity	International Agency for Research on Cancer
IRIS Carcinogen assessments	Carcinogenicity	US Environmental Protection Agency, Integrated Risk Information System Database (IRIS)
EU Priority Endocrine Disruptors	Endocrine activity	European Commission, EU Community Strategy for Endocrine Disruptors
AOEC Exposure Codes – Asthmagen List	Respiratory sensitization	Association of Occupational and Environmental Clinics
EU Scientific Committee on Consumer Safety – Fragrance Allergens	Respiratory sensitization	European Commission
GHS classifications published by national-level agencies	Multiple endpoints	Multiple sources, including the European Chemicals Agency and the Governments of Japan, Denmark, New Zealand, Korea, Malaysia, and Australia
Stockholm Convention – Persistent Organic Pollutants	PBT	United Nations Environment Programme, Stockholm Convention on Persistent Organic Pollutants
IPCC Global Warming Chemicals	Global warming	Intergovernmental Panel for Climate Change (IPCC) Third Assessment Report (2001)
Ozone-Depleting Substances (ODS) Class I & Class II	Ozone depletion	US Environmental Protection Agency
Substitute It Now (SIN) List	Multiple endpoints	International Chemical Secretariat (ChemSec)
Substances restricted under REACH	Multiple endpoints; Restricted substances list	European Chemicals Agency
RoHS Annex II	Restricted substances list	European Commission

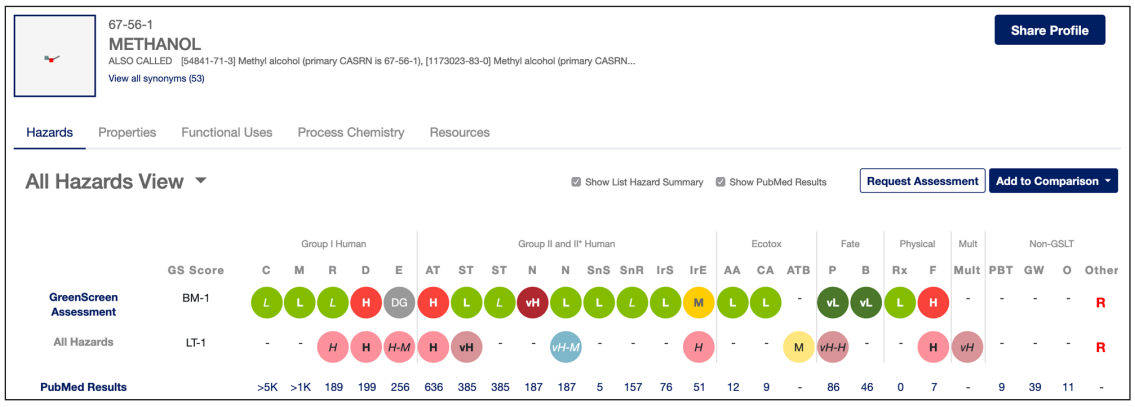

Figure 1: Hazard summary table displayed in the Pharos substance profile for methanol. Each chemical profile shows the overall substance-level hazard score (GreenScreen Benchmark or GSLT score) as well as endpoint-specific indicators that underlie the hazard assessment. The table includes hazard information based on full a GreenScreen assessment, if available, as well as hazard indicators derived from authoritative public data sources. The display uses consistent visual conventions to communicate characterizations of hazard levels, data availability (or data gaps), and strength of evidence. It also includes links to scientific literature searches about the substance in PubMed, showing the number of hits for each endpoint. These information elements are dynamically generated and therefore always up-to-date in relation to the resources they reference. DOI: https://doi.org/10.1525/elementa.422.f1

in the Data Commons. However, the CAA practitioner community lacks a standard for describing or categorizing the functional uses of chemicals. Instead, we found that there are many diverse and incommensurate systems for categorizing functional uses, including industry-specific vocabularies and regulatory classification systems (Blake 2013, 2014).

In the Data Commons, we understand functional use as determined by varying sets of chemical properties in a wide range of technological contexts, for which no single

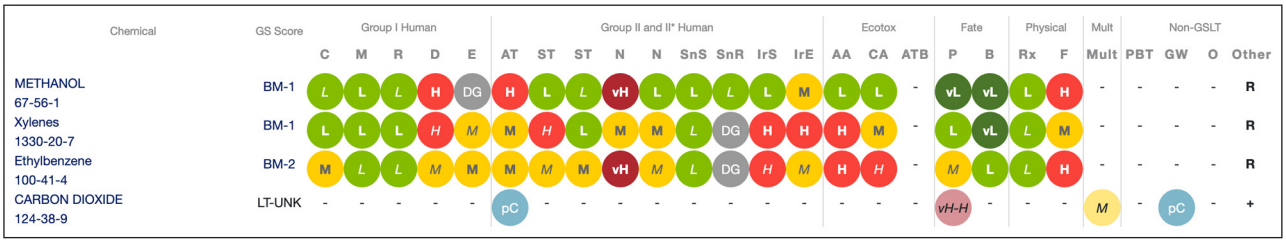

Figure 2: A hazard comparison table displayed in Pharos. The hazard comparison tool supports CAA practitioners by enabling them to create, share, import, and export customized sets of substances for automated screening. The hazard profiles of each substance are displayed together in a matrix that enables comparative analysis of safer alternatives and data gaps. DOI: https://doi.org/10.1525/elementa.422.f2

Table 3: Representative sources of functional use information included in the Chemical Hazard Data Commons. DOI: https://doi.org/10.1525/elementa.422.t3

Title	Source
CosIng Cosmetic Ingredient Database	European Commission
REACH Registered Substances: Article Categories	European Chemicals Agency
Hazardous Substances Databank (HSDB)	US National Institutes of Health
Safer Chemical Ingredient List (SCIL)	US Environmental Protection Agency
US EPA Registered pesticides	US Environmental Protection Agency
Chemical and Products Database (CPDat)	US Environmental Protection Agency
HBN Common Products Database	Healthy Building Network

overarching classification system is necessarily appropriate. For example, researchers at the US EPA integrated many sources of data on chemical ingredients in consumer products to produce CPDat, a database that links thousands of chemicals to a harmonized set of functional use categories (Dionisio et al. 2018). While CPDat provides an excellent source of information (and is used in the Data Commons), it still emphasizes certain consumer product sectors and therefore certain types of functions—is not a "universal" classification system for functional use. For functional use descriptors, we therefore adopted a tagging-based system that is able to include terms from many data sources (see **Table 3**) even if they are based on conceptually disparate systems. This is in contrast to how hazard is understood, where the use of a single classification system enables consistent comparisons.

Compiling and linking data

Rather than serving as a container of information, the Data Commons was designed as a portal that makes accessible information from many sources. Using the paradigm of linked data (Heath and Bizer 2011), the system consumes and integrates data from several high-quality open chemical data sources and provides links for users to navigate directly to specific records in the source systems. This allows the system to tap into a rich, distributed landscape of chemical information without reproducing the work or taking on the costs of providing and maintaining all of that information directly. Examples of linked data sources are included in **Table 4**. For example, physical property data and molecular structure graphics displayed in the Data Commons are drawn from PubChem. Hazard summary tables include linked scientific literature searches for each specific substance-endpoint combination via PubMed, showing the number of hits and enabling users to instantly delve into the search results (see **Figure 1**). Because these elements are dynamically generated, they are always up-to-date in relation to the resources they reference.

Yet, much of the publicly available information relevant for CAA cannot easily be collected from open data sources. A central part of the Data Commons was borrowed from (and is still part of) HBN's Pharos project: the Chemical and Material Library (CML), which compiles and cross-references hazard associations from authoritative public-domain sources such as those in **Table 2** (for full documentation, see Healthy Building Network 2019c). This compilation takes substantial effort: HBN curates and regularly updates the library from government and scientific sources that encompass a great variety of original publication formats.

Leveraging open standards and frameworks

Open standards and frameworks play important roles in CAA: promoting consistency in how science is used to inform decisions, and ensuring mutual compatibility of work products and practices in the field. The Data Commons was designed to both leverage and reinforce open standards to support CAA. Two notable examples are the GreenScreen for Safer Chemicals (discussed above) and the Health Product Declaration (HPD). HPD is an open standard for disclosure and communication of the material content and associated hazards of building materials (HPD Collaborative 2020), which has been instrumental in extending

Table 4: Representative linked data sources included in the Chemical Hazard Data Commons. DOI: https://doi.org/10.1525/elementa.422.t4

Title	Source	Type of data
PubChem	US National Institutes of Health	Physical properties, molecular structure graphics
ChemIDplus	US National Institutes of Health	Physical properties, synonyms
Hazardous Substances Data Bank	US National Institutes of Health	Literature summaries
PubMed	US National Institutes of Health	Literature search results
CompTox Chemistry Dashboard	US Environmental Protection Agency	Synonyms, database IDs
ToxCast Dashboard	US Environmental Protection Agency	Toxicity data
Toxics Release Inventory	US Environmental Protection Agency	Pollution and waste (USA)
REACH Registration Dossiers	European Chemicals Agency	Manufacturer-submitted data (EU)
OECD eChemPortal	Organisation for Economic Co-operation and Development	Various
International Toxicity Estimates for Risk (ITER)	Toxicity Excellence for Risk Assessment (TERA)	Toxicity data

material hazard considerations into sustainable design in the building sector. As detailed above, the Data Commons employs GreenScreen to organize information; but it also helps to operationalize the GreenScreen method by automatically computing the GSLT score for any chemical. This is possible because the system integrates all of the data sources that the GSLT protocol requires for hazard evaluation. Similarly, the Data Commons was designed to support the use of HPD by integrating all "priority hazard lists" referenced in the hazard screening component of the standard, thus giving manufacturers and designers the ability to instantly evaluate chemical and product hazards as specified by the standard.

In collaboration with Clean Production Action, the Data Commons has also contributed to new standards intended to make list-based hazard screening more consistent, including standards for updating screening list databases and for interpreting the hazard properties of groups and classes of chemicals.[8]

Community engagement and knowledge sharing

When we first articulated the idea of the Chemical Hazard Data Commons, we understood it as "both a tool-building project and a community-building project" (Lent et al. 2014). A global community of experts and practitioners was already engaged in collective work to make chemical information actionable for CAA. For example, organizations in North America and Europe—such as ChemSec and the Business-NGO Working Group for Safer Chemicals and Sustainable Materials—were already working to advance CAA across government, business, and advocacy sectors. The role we envisioned for the Data Commons was to link this community together with shared and interactive resources that filled key knowledge needs.

In an effort to build the commons, we engaged an initial core group of people who expressed high interest in the project goals, sought to create opportunities for knowledge sharing and collaboration, and set up a basic framework for fostering and governing community interactions. We reached out to professionals located primarily in the US and Europe, and representing approximately 40 organizations in the following sectors: (1) government agencies involved with chemical assessment, regulation, or public chemical information services; (2) NGOs and labor unions active on environmental health issues; (3) academic institutions; (4) firms and independent consultants in the chemical hazard assessment consulting sector; and (5) companies in the manufacturing and retail sectors. Throughout the development of the Data Commons, over 50 community members had direct input into the design and priorities of the commons at the scoping stage via discussions, interviews, and written comments on the discussion papers (Healthy Building Network 2014b, 2014a). Over 100 community members helped to shape the early development of the web system through their responses to a detailed online survey, pilot testing, and participation in the development of the infrastructure (Healthy Building Network 2016). The resulting system design reflected our thinking and the feedback we received from participants.

We included online discussion forums, intending for these to facilitate knowledge sharing among the commons participants who have diverse professional backgrounds; they include chemists, toxicologists, and environmental health scientists, but also engineers, public health practitioners, civil servants, architects, and industrial designers. Given this heterogeneity, simply inviting open discussion among participants seemed like a good starting point. Discussion threads are organized into broad categories of topics and tagged with detailed topic descriptors, and they can be linked to specific substance data records.

Actively facilitating and coordinating collaborative efforts among participants was a more ambitious design goal. Such efforts could include working together on new knowledge resources or collectively evaluating scientific evidence about chemical hazards. Designing systems to foster direct collaboration between participants remains a priority for future development of the commons. Still, we did include some elements (currently included in Pharos) that were intended to provide opportunities for collaborative work. For example:

- Substance comparison sets can be shared—publicly or among limited groups—for collaborative curation, discussion, or CAA work.
- Discussion forums include special sections dedicated to ongoing collaborative projects.
- Participants can signal their demands for full GreenScreen assessments of particular chemicals by clicking a "Request Assessment" button, which allows the system to aggregate community interest in developing new knowledge.

Governance

According to David Bollier (2014), "a commons is *a resource + a community + a set of social protocols*." Social protocols are what govern the processes of sharing and management through which the commons can provide ongoing benefits to the community (Frischmann, Madison, and Strandburg 2014). We made several key decisions about these social protocols, or the governance of the commons.

In the Data Commons, governance relates to what participants *bring in*. To the extent that anyone can actively contribute knowledge, ideas, and discourse, they have an influence on what others can take away from the commons. We identified two potential problems[9] to which the Data Commons could be vulnerable: false contributions of knowledge, such as misleading claims or inaccurate data; and discourse that seeks to undermine the goals, methods, or practices of CAA. Both of these problems have precedents in the often contentious public discourse around chemical safety and risk.

Whereas CAA seeks to reduce harm by substituting toxic substances with safer alternatives, the dominant industry and government discourse in chemicals policy is based on the idea of acceptable risk—a notion that emerged as a compromise between industry interests and public health protection in the conflicted history of environmental policy-making (Boyd 2012). As exemplified by contemporary controversies about substances like bisphenol A and glyphosate, representatives of the chemical industry have historically used scientific discourse to advocate for the use of data, technical criteria, and assumptions that make health risks seem more acceptable (Vogel 2009; Montenegro de Wit and Iles 2015). The science and methods of CAA appear to be under scrutiny from advocates of risk-based approaches (Palmer 2016). For example, after the publication of the Commons Principles for Alternatives Assessment (BizNGO 2013), a coalition of industry advocates created their own competing set of completely different principles a year later (American Chemistry Council 2014). It was reasonable to anticipate that the participatory features of the Data Commons could also be used to undermine the goals of safer chemical substitution.

To address these potential problems, we instituted the following basic governance mechanisms, which set boundaries on what kinds of knowledge and discourse can be contributed to the commons and expectations about how people should conduct themselves as participants.

- A Code of Conduct (COC), which all participants must agree to (Healthy Building Network 2019a). The COC provides guidelines for how community members should participate, and a standard to which community members can hold each other (see below). It also defines sanctions that can be brought against participants who violate the code, as well as an appeals process.
- Terms of Service (TOS), which contain official rules and the consequences of breaking them in precise legal language (Healthy Building Network 2019b). These rules may be invoked in cases of clear and severe breach of the community guidelines set out in the COC.
- Technical systems and requirements to support the social protocols set out in the COC and TOS. Site registration is required before contributing content, enabling the commons administrators to identify and address any offenses. Users with administrative privileges also have control over contributed content.

The norms and guidelines expressed in the COC represent what we viewed as reasonable basic principles that our community would readily accept and abide by—such as being considerate, respectful, constructive, and avoiding advertising or political campaigning. The COC generally posits that the Data Commons should be used for reasoned discussion of science pertaining to CAA, and codifies those concepts via a mission statement and conflict of interest (COI) guidelines. The mission statement links the goals of the Data Commons to the concepts of hazard assessment and alternatives assessment, and Data Commons participants are asked to recognize and support it:

> The mission of the Chemical Hazard Data Commons is to advance science, policy, and business efforts to identify hazardous substances through chemical hazard assessment and find safer substitutes through alternatives assessment. We aim to do this by developing community-accessible tools and information resources. Your engagement in this project should be grounded in your support of these goals. (Healthy Building Network 2019a)

Establishing a shared understanding of the meaning and value of these concepts addresses the contested nature of the field of alternatives assessment as a science/policy field. While the Data Commons community thrives on open, constructive criticism of methods and science, we wanted it to be clear that actively undermining these goals is an unwelcome form of participation.

The COC also asks participants to be transparent about potential COI. A COI may arise when someone contributes to the Data Commons on a topic that has bearing on their own financial interests or those of their affiliates. We did *not* seek to discourage participants with COI from contributing, because there are potential COI situations where participation would still be highly valuable—for example, users who work with companies seeking to commercialize safer substances and products. Rather, we ask for awareness and transparency of COI so that the community can understand the full context of all the information in the commons.

Discussion

The design considerations and development process reported here represent our efforts to plan, initialize, and formally constitute a knowledge commons for chemical hazard assessment. Although this is still a partially realized and evolving project, it has already produced a valuable public resource at the same time as it has surfaced some important challenges and limitations.

Outcomes

Feedback from a wide range of Data Commons participants indicates that the project has yielded a valuable and novel public resource. By enabling easy access and reference to the most commonly needed information and tools, the Data Commons benefited practitioners across the field of chemical alternatives assessment. It was the first open-access system to link information about chemical substances with hazard properties in a consistent way—thus facilitating direct comparisons between substances on the basis of hazard—and also to link substances to functional use classifications. The development of the Data Commons provided two particularly useful functions for chemical assessors: automated chemical screening using the GreenScreen List Translator protocol, and hazard-based comparisons. Having free access to these resources has been particularly helpful for public-sector and civil society organizations, small enterprises (such as environmental consultants), and educational institutions, because they may not have access to equivalent paid services. Several academic and NGO research projects have reported using the Data Commons as a resource. It has also helped green chemistry education: for several years, students at the University of California, Berkeley have used the Data Commons (and Pharos) to learn and carry out CAA in real applied projects (Schwarzman and Buckley 2019).

Moreover, participants can join or interact with a community of practice, and keep themselves updated on developments in the community or broader field. The Data Commons has hosted a variety of discussions and debates in an open forum—now continued in Pharos. Some of our participants have noted that the Data Commons consolidated, and in some cases replaced, previous informal person-to-person channels for knowledge sharing. One of the collaborative efforts initiated in the Data Commons—the systematic identification of substances belonging to known hazardous chemical classes, mentioned above—became a significant and thoroughly-debated ongoing project involving multiple organizations (Healthy Building Network 2020a). To date, HBN staff have collaborated with various organizations to integrate findings from their projects into the commons, making participants among the first to benefit from the work. This has resulted in new information resources being incorporated into the system, such as data about chemicals found in plastic packaging materials (in collaboration with Food Packaging Forum: Groh et al. 2018) and hard-to-find regulatory agency documents obtained via FOIA requests (in collaboration with the Natural Resources Defense Council).

Challenges

The experimental participatory project begun by the Data Commons leaves many opportunities for advancement, and its development faced a number of challenges—particularly in the areas of fostering participation and collaboration, and planning for long-term governance and viability.

Participation and collaboration

Ideally, the Data Commons could productively harness some of the community's attention, directing a little effort from many participants toward collective tasks. The commons is designed to do this by enabling multi-directional interactions like commenting on resources, asking and answering questions, or contributing to open projects. However, as a new entrant into the field of tools and collaborative practices for chemical alternatives assessment, the commons faced a paradox in which potential participants would be more likely to engage if there were already more engaged participants. Even after several years, most participants seemed to prefer using the Data Commons as a tool rather than to interact with others in an active exchange. Providing tools was indeed one of the goals of the commons, but it does not reflect the full participatory ambition of the project.

Another part of the aspirational vision of the Data Commons was to aggregate community expertise to accelerate progress on shared goals—such as improved CAA tools, methods, and data. We imagined, for example: collaboratively-authored summaries of scientific studies on substances of interest, or community peer-reviewed methods for filling data gaps in chemical assessments. The CAA community already engages in a number of collaborative efforts, such as tackling shared methodological problems.[10] We hoped to extend collaborative practices into the realm of data. There are many excellent precedents for collective work in knowledge commons, such as open-source software (Schweik and English 2012), Wikipedia (Benkler 2006), and new scientific research practices that Nielsen (2012) refers to as "networked science." For these networked efforts to be successful, there must be self-sustaining participatory systems that efficiently direct the attention and expertise of participants where it is most needed and best matched to their individual interests and skills. For example, very large projects can be coordinated through modular organization, or incrementally accomplished through many accumulated "microcontributions"—also known as crowdsourcing (Nielsen 2012).

Initializing active collaborations in the Data Commons has been challenging, and crowdsourcing of data resources has generally not occurred. For example, an earlier version of the Data Commons included a shared library of scientific citations and public domain documents to which users could contribute—but few did, and it was eventually discontinued. Two main factors may explain these challenges. First, the commons has not yet evolved sufficiently effective mechanisms to harness the attention and efforts of the community. We could have actively guided and facilitated engagement, which would have come with significant organizational and resource

demands on HBN. Instead, we relied to a large extent on the online platform itself to foster participation. But discussion forums enable only a limited form of interaction, whereas coordinating the collaborative review of a dataset (for example) may require dedicated staff time and far more specialized software tools. In other words, the design and implementation of the participatory infrastructure is still incomplete. Second, there may not be sufficient incentives or appropriate opportunities for participants to get more involved. For example, prospective participants may feel that a project's possible outcomes are not worth their time, that it could get done anyway without their input, or that they can't participate because the project is too technically demanding. Ideal collaborative projects would be those that present a compelling shared benefit and that are easy for many participants to contribute to. Identifying such projects should be a priority for future development of the commons.

Creating a hazard assessment commons

One hope was that the Data Commons could aggregate community resources by producing and distributing a significant number of new GreenScreen hazard assessments. A GreenScreen assessment is a highly valuable work product requiring considerable toxicological knowledge and effort. The vast majority of GreenScreen assessments is produced by a handful of companies with vetted expertise and capacity (Clean Production Action 2020b). Many of these chemical profilers are part of the Data Commons user community, and there was a considerable level of interest in leveraging the commons to "scale" CHA practice.

Nevertheless, facilitating the commons-based production of chemical hazard assessments remains a significant challenge. First, CHA is highly demanding work regardless of the resources provided by the Data Commons. It requires dedicated professional effort and associated expenses. Second, to recover the high costs of CHA, profilers restrict the redistribution of their work using copyright. These intellectual property arrangements are in tension with the commons-based intention to pool assessments for collective re-use—and more generally with the open access goals of the Data Commons project. Some profiler organizations, ToxServices and NSF International, have collaborated to release a limited number of assessments into the public domain, but they cannot be expected to make all of their work freely available.

One approach for creating a common pool of chemical assessments is to abandon open access and switch to a more limited form of sharing that is supported by subscription fees. A newly emerging project, ChemFORWARD (ChemFORWARD 2020), is pursuing this strategy. Another approach is to generate an open-access pool of assessments by funding profilers to work on a set of substances regarded by the community as high priority for assessment. This approach has not been realized, but still holds promise given the high financial stakes of new chemicals regulations (e.g. California Department of Toxic Substances Control 2020) that may require industry to switch to safer alternatives for certain substances in major product categories.

Long-term governance

Another challenge to be faced by the Data Commons is its governance and maintenance in the long term. Part of the rationale for creating a Data Commons was that no single organization could support all of the information needs of the community. Yet, due to the particular circumstances of organizational capacity and philanthropic funding streams, there is currently a single organization (namely the Healthy Building Network) endeavoring to manage and sustain the commons. One consequence of this is that the project has been conceived and executed largely in a US-centric context, privileging Anglophone resources and audiences despite the global nature of the broader community.

At a minimum, sustaining the Data Commons involves development and maintenance of essential infrastructure (software, servers, databases, etc.), financially supporting a free public-access resource, and providing a basic level of coordination for the affairs of the community (e.g., assigning administrative privileges and duties to volunteers). These responsibilities could in principle be shared among a group of organizations and individuals, but there are currently no arrangements for allocating or distributing them in the long term. Likewise, the challenge of long-term financial support may require shifting from philanthropic funding models to community contributions or other models.

Conclusions

We developed the Chemical Hazard Data Commons to fill a need for public information infrastructure to support chemical alternative assessment practice. More fundamentally, we developed the Data Commons as an experiment in collective capacity-building for safer chemical substitution—as an alternative to continued reliance on regulatory and market drivers to furnish this information infrastructure.

The design of the Data Commons considers how chemical hazard information should be organized so that it can be retrieved and used in meaningful ways, particularly through the lens of frameworks and standards prevalent in the practitioner community. These design considerations contributed to the development of new tools and functions based on aggregated and linked data. The commons was also designed to enable interactions among community members, with the aspirational goal of facilitating rich multi-directional flows of knowledge among a diverse set of producers and users of chemical hazard information.

This development process represents a novel application of a collaborative, commons-based approach, and the first effort to create an open, participatory information infrastructure for CAA. In an arena where decades of adversarial policymaking and activism around toxic chemicals have led to intense contestation of scientific knowledge, the Data Commons approach emphasizes the value of shared information resources and collective capacity-building.

The development of the Data Commons, now merged into the ongoing Pharos Project, has yielded an open-access resource available online at https://pharosproject.net/.

The commons exists for members of a global community interested in CAA, and serves as a tool and a space for exchanging knowledge and expertise. We welcome participation in the commons itself, as well as constructive criticism of our efforts to design and formally constitute the commons. While this project is unlikely to singlehandedly solve most of the problems that motivated it, we hope that it will not be the last of its kind. We present our work, and the challenges we encountered, so that our experience may benefit future efforts.

Notes

[1] Infrastructure refers to information technologies that provide services, such as web-based software and databases. It also refers to information systems themselves, and the technical, conceptual, and social agreements that they embody, which serve to enable and coordinate work across many people at once (Bowker and Star 1999).

[2] "Functional use" refers to functions which a substance may perform in a product or formulation, such as oxidant or plasticizer (Tickner et al. 2015).

[3] Including CAS Registry Number, IUPAC InChI and InChIKey, SMILES, and PubChem CID.

[4] Indeed, as a result of the work of many research organizations, there is already a much broader information commons of web-enabled chemical data resources, mostly centered on biosciences and computational chemistry fields (for a holistic view, see Murray-Rust et al. 2011).

[5] The conference was organized by Charlotte Brody (BlueGreen Alliance), Mark Rossi (Clean Production Action), and Joel Tickner (University of Massachusetts, Lowell Center for Sustainable Production).

[6] This seemingly self-evident principle conceals a complex and historically contingent answer to the question of what makes "one" substance unique (Hepler-Smith 2019). The Data Commons identifies chemical substances using Chemical Abstracts Service Registry Numbers (CASRN), PubChem Compound identifiers (CID) (Kim et al. 2016), SMILES molecular representations (Weininger 1988), and IUPAC International Chemical Identifiers (InChI) (Heller et al. 2013).

[7] In principle, the commons could have been designed to use multiple hazard classification and evaluation frameworks in tandem. Using the same set of hazard data, the system could translate between frameworks to accommodate user preferences. The Pharos Project currently takes this approach, allowing users to choose between multiple "views" of hazard information.

[8] For example, the GreenScreen List Translator automator update policy (Clean Production Action 2017) and the GreenScreen chemical groups policy (Clean Production Action 2018; Healthy Building Network 2020a).

[9] Beyond these specific problems, any system on the internet is vulnerable to more general problems like spam (unwanted content) and intellectual property liabilities (e.g. a user posts the entire text of a copyrighted article in a comment).

[10] For example, through formally organized working groups consisting mostly of volunteer participants.

[11] This categorization of endpoints is based on the GreenScreen for Safer Chemicals (Clean Production Action 2020a).

Acknowledgements

We wish to acknowledge Sarah Gilberg, Larry Kilroy, and Jon Stavis, who designed and developed the information systems and software for HBN's Pharos Project and the Data Commons. We also acknowledge the HBN Materials Research Team for their work on researching and compiling the building product-related chemical information that is included in Pharos. We thank Megan Schwarzman, Christine Rosen, Rachel Morello-Frosch, and Alastair Iles for their reading and critical suggestions on the manuscript.

Funding information

The Data Commons project has received funding from the Forsythia Foundation, the Lisa & Douglas Goldman Foundation, the John Merck Fund, and the JPB Foundation. MD and TL were supported, through HBN, by grants from those foundations. AK was partially supported by a National Science Foundation IGERT Fellowship at the Berkeley Center for Green Chemistry. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. AB participated in this work without financial compensation. Publication of this article is made possible in part by support from the Berkeley Research Impact Initiative (BRII) sponsored by the UC Berkeley Library.

Competing interests

AK serves on the GreenScreen for Safer Chemicals Science Advisory Committee, without compensation. The authors declare that no competing interests exist.

Author contributions

- Contributed to conception and design: AK, AB, MD, TL
- Drafted the article: AK
- Critically revised the article: AK, AB, MD, TL
- Approved the submitted version for publication: AK, AB, MD, TL

References

American Chemistry Council. 2014. "Manufacturing Coalition Announces Principles for Effective Alternative Assessments." *ACC News Releases* (blog). March 4, 2014. http://www.americanchemistry.com/Media/PressReleasesTranscripts/ACC-news-releases/Manufacturing-Coalition-Announces-Principles-for-Effective-Alternative-Assessments.html.

Ashford, NA. 2005. "Incorporating Science, Technology, Fairness, and Accountability in Environmental, Health, and Safety Decisions." *Human and Ecological Risk Assessment: An International Journal* **11**(1): 85–96. DOI: https://doi.org/10.1080/10807030590919918

Benkler, Y. 2006. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. New Haven: Yale University Press.

BizNGO. 2013. "The Commons Principles for Alternatives Assessment: Addressing Chemicals of Concern to Human Health or the Environment." http://www.bizngo.org/alternatives-assessment/commons-principles-alt-assessment.

Blake, A. 2013. "Examining Existing Systems for Functional Use Categorization."

Blake, A. 2014. "Chemical Hazard Data Commons Working Paper: Finding Alternatives Through Categorizing Functional Use." Healthy Building Network. https://s3.amazonaws.com/hbndatacommons/ChemDat-WorkPaper-FunctionalUse.pdf.

Bobenhausen, C. 2016. "Harmonization in Green Building: Tools for Assessing Chemical Hazards in Building Products." *The Synergist*, April 2016. https://synergist.aiha.org/201604-harmonization-in-green-building.

Bollier, D. 2014. *Think Like a Commoner: A Short Introduction to the Life of the Commons*. Gabriola, British Columbia: New Society Publishers.

Bowker, GC and **Star, SL.** 1999. *Sorting Things Out: Classification and Its Consequences*. Cambridge, Mass.: MIT Press. DOI: https://doi.org/10.7551/mitpress/6352.001.0001

Boyd, W. 2012. "Genealogies of Risk: Searching for Safety, 1930s–1970s." *Ecology Law Quarterly* **39**(4): 895–987. DOI: https://doi.org/10.15779/Z38QG3T

California Department of Toxic Substances Control. 2020. "Safer Consumer Products (SCP)." 2020. https://dtsc.ca.gov/scp/.

ChemForward. 2020. "ChemForward." 2020. https://www.chemforward.org.

Clean Production Action. 2017. "GreenScreen List Translator Automator Update Policy (Version 1.0)." http://www.greenscreenchemicals.org/method/method-documents.

Clean Production Action. 2018. "GreenScreen for Safer Chemicals Chemical Groups Policy (Version 1.0)." http://www.greenscreenchemicals.org/method/method-documents.

Clean Production Action. 2020a. "GreenScreen for Safer Chemicals." 2020. https://www.greenscreenchemicals.org/.

Clean Production Action. 2020b. "GreenScreen Professionals & Service Providers." *GreenScreen for Safer Chemicals*. 2020. https://www.greenscreenchemicals.org/assess/gs-professionals.

Cradle to Cradle Products Innovation Institute. 2019. "Material Health Assessment Methodology." *Cradle to Cradle Products Innovation Institute*. January 2019. https://www.c2ccertified.org/resources/detail/material_assessment_methodology.

Dionisio, KL, Phillips, K, Price, PS, Grulke, CM, Williams, A, Biryol, D, Hong, T, and **Isaacs, KK.** 2018. "The Chemical and Products Database, a Resource for Exposure-Relevant Data on Chemicals in Consumer Products." *Scientific Data* **5**(July): 180125. DOI: https://doi.org/10.1038/sdata.2018.125

Egeghy, PP, Judson, R, Gangwal, S, Mosher, S, Smith, D, Vail, J and **Hubal, EAC.** 2012. "The Exposure Data Landscape for Manufactured Chemicals." *Science of the Total Environment* **414**(January): 159–66. DOI: https://doi.org/10.1016/j.scitotenv.2011.10.046

Frischmann, BM, Madison, MJ and **Strandburg, KJ.** (eds.). 2014. *Governing Knowledge Commons*. Oxford; New York: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199972036.001.0001

Fung, A and **O'Rourke, D.** 2000. "Reinventing Environmental Regulation from the Grassroots up: Explaining and Expanding the Success of the Toxics Release Inventory." *Environmental Management* **25**(2): 115–27. DOI: https://doi.org/10.1007/s002679910009

Geiser, K. 2014. "The New Chemicals in Products Information Transparency." *GBIG Insight Blog* (blog). January 13, 2014. http://insight.gbig.org/the-new-chemicals-in-products-information-transparency/.

Geiser, K. 2015. *Chemicals Without Harm: Policies for a Sustainable World*. Urban and Industrial Environments. Cambridge, Mass.: MIT Press. DOI: https://doi.org/10.7551/mitpress/9780262012522.001.0001

Gilbert, N. 2016. "Legal Tussle Delays Launch of Huge Toxicity Database." *Nature*, February. DOI: https://doi.org/10.1038/nature.2016.19365

Glushko, RJ. 2013. *The Discipline of Organizing*. Cambridge, Mass.: MIT Press. DOI: https://doi.org/10.1002/bult.2013.1720400108

Groh, KJ, Backhaus, T, Carney-Almroth, B, Geueke, B, Inostroza, PA, Lennquist, A, Leslie, HA, Maffini, M, Slunge, D, Trasande, L, Michael Warhurst, A and **Muncke, J.** 2018. Overview of known plastic packaging-associated chemicals and their hazards. *Science of The Total Environment*. DOI: https://doi.org/10.1016/j.scitotenv.2018.10.015

Guth, JH, Denison, RA and **Sass, J.** 2007. "Require Comprehensive Safety Data for All Chemicals." *New Solutions* **17**(3): 233–58. DOI: https://doi.org/10.2190/NS.17.3.g

Harrison, RM and **Hester, RE.** (eds.). 2013. *Chemical Alternatives Assessments. Issues in Environmental Science and Technology* **36**. Cambridge: Royal Society of Chemistry. DOI: https://doi.org/10.1039/9781849737234

Healthy Building Network. 2014a. "Chemical Hazard Data Commons White Paper Reviewer Commentary." https://s3.amazonaws.com/hbndatacommons/ChemDat-ContentFeedback.pdf.

Healthy Building Network. 2014b. "Chemical Hazard Data Commons White Paper Survey Results." https://s3.amazonaws.com/hbndatacommons/ChemDat-SurveyResults.pdf.

Healthy Building Network. 2016. "Community Survey Response Analysis." https://drive.google.com/file/u/1/d/0ByChXzZXqy9WRWNJTWYweEZldTQ/view?usp=drive_open&usp=embed_facebook.

Healthy Building Network. 2019a. "Chemical Hazard Data Commons Code of Conduct." *Chemical Hazard Data Commons*. 2019. https://commons.healthymaterials.net/code.

Healthy Building Network. 2019b. "Chemical Hazard Data Commons Service Terms and Conditions

of Use." *Chemical Hazard Data Commons*. 2019. https://commons.healthymaterials.net/terms.

Healthy Building Network. 2019c. "Chemical & Material Library (CML) Full System Description."

Healthy Building Network. 2019d. "Combining the Powers of Pharos and Data Commons." August 2019. https://www.pharosproject.net/merger.

Healthy Building Network. 2020a. "Compound Group Population Project." *Pharos*. 2020. https://pharosproject.net/compound-groups.

Healthy Building Network. 2020b. "Pharos Project." *Pharos*. 2020. https://www.pharosproject.net/.

Heath, T and **Bizer, C.** 2011. "Linked Data: Evolving the Web into a Global Data Space." *Synthesis Lectures on the Semantic Web: Theory and Technology* **1**(1): 1–136. DOI: https://doi.org/10.2200/S00334ED1V01Y201102WBE001

Heine, LG and **Franjevic, SA.** 2013. "Chemical Hazard Assessment and the GreenScreen for Safer Chemicals." In *Chemical Alternatives Assessments*. Harrison, RM and Hester, RE (eds.), 129–56. *Issues in Environmental Science and Technology* **36**. Cambridge: Royal Society of Chemistry. DOI: https://doi.org/10.1039/9781849737234-00129

Heine, LG, Kausch, M, Klosterhaus, S, Glass, S and **Lent, T.** 2013. "Material Health Evaluation Programs: Harmonization Opportunities Report." *US Green Building Council*.

Heller, S, McNaught, A, Stein, S, Tchekhovskoi, D and **Pletnev, I.** 2013. "InChI – the Worldwide Chemical Structure Identifier Standard." *Journal of Cheminformatics* **5**(1): 7. DOI: https://doi.org/10.1186/1758-2946-5-7

Hepler-Smith, E. 2019. "Molecular Bureaucracy: Toxicological Information and Environmental Protection." *Environmental History* **24**(3): 534–60. DOI: https://doi.org/10.1093/envhis/emy134

Hess, C and **Ostrom, E.** (eds.). 2007. *Understanding Knowledge as a Commons: From Theory to Practice*. Cambridge, Mass.: MIT Press. DOI: https://doi.org/10.7551/mitpress/6980.001.0001

HPD Collaborative. 2020. "Health Product Declaration Collaborative." 2020. https://www.hpd-collaborative.org.

Jacobs, MM, Malloy, TF, Tickner, JA and **Edwards, S.** 2016. "Alternatives Assessment Frameworks: Research Needs for the Informed Substitution of Hazardous Chemicals." *Environmental Health Perspectives* **124**(3). DOI: https://doi.org/10.1289/ehp.1409581

Judson, R, Richard, A, Dix, DJ, Houck, K, Martin, M, Kavlock, R, Dellarco, V, Henry, T, Holderman, T, Sayre, P, Tan, S, Carpenter, T and **Smith, E.** 2009. "The toxicity data landscape for environmental chemicals." *Environmental Health Perspectives* **117**(5): 685–95. DOI: https://doi.org/10.1289/ehp.0800168

Kim, S, Thiessen, PA, Bolton, EE, Chen, J, Fu, G, Gindulyte, A, Han, L, He, J, He, S, Shoemaker, BA, Wang, J, Yu, B, Zhang, J and **Bryant, SH.** 2016. PubChem substance and compound databases. *Nucleic Acids Research* **44**(D1): D1202–D1213. DOI: https://doi.org/10.1093/nar/gkv951

Kokai, A, Lent, T and **Dedeo, M.** 2014. "Chemical Hazard Data Commons Working Paper: Getting Hazard Right Through Managing Chemical IDs." *Healthy Building Network*. https://s3.amazonaws.com/hbn-datacommons/ChemDat-WorkPaper-ChemID.pdf.

Lavoie, ET, Heine, LG, Holder, H, Rossi, MS, Lee, RE, Connor, EA, Vrabel, MA, DiFiore, DM and **Davies, CL.** 2010. "Chemical Alternatives Assessment: Enabling Substitution to Safer Chemicals." *Environmental Science & Technology* **44**(24): 9244–9. DOI: https://doi.org/10.1021/es1015789

Lent, T. 2014a. "Chemical Hazard Data Commons Working Paper: Facilitating Decision Making with Assessment Summaries." *Healthy Building Network*. https://s3.amazonaws.com/hbndatacommons/ChemDat-WorkPaper-AssessmentSummaries.pdf.

Lent, T. 2014b. "Chemical Hazard Data Commons Working Paper: Speeding & Improving Assessment Through Data Sharing." *Healthy Building Network*. https://s3.amazonaws.com/hbndatacommons/ChemDat-WorkPaper-DataSharing.pdf.

Lent, T, Kokai, A, Blake, A, Walsh, B, Kilroy, L and **Dedeo, M.** 2014. "Toward Safer Products: Accelerating Change with a Chemical Hazard Data Commons." *Healthy Building Network*. https://s3.amazonaws.com/hbndatacommons/ChemDat-WhitePaper.pdf.

Lent, T and **Stamm, R.** 2019. "The Solution Is Transparency: Public Disclosure of Product Content and Hazard Is Critical to Solving Material Health Challenges." *Healthy Building Network* (blog). January 2019. https://healthybuilding.net/blog/499-the-solution-is-transparency-public-disclosure-of-product-content-and-hazard-is-critical-to-solving-material-health-challenges.

Logan, K. 2016. "Healthy Materials: Sweating the Small Stuff." *Architectural Record*, June 1, 2016. http://www.architecturalrecord.com/articles/11711-healthy-materials.

Montenegro de Wit, M and **Iles, A.** 2015. "Getting Past Scientized Scrutiny." *Earth Island Journal*, April 29, 2015. https://www.earthisland.org/journal/index.php/articles/entry/getting_past_scientized_scrutiny/.

Murray-Rust, P, Adams, SE, Downing, J, Townsend, JA and **Zhang, Y.** 2011. "The Semantic Architecture of the World-Wide Molecular Matrix (WWMM)." *Journal of Cheminformatics* **3**(1): 42. DOI: https://doi.org/10.1186/1758-2946-3-42

National Research Council (US). 2014. *A Framework to Guide Selection of Chemical Alternatives*. Washington, DC: The National Academies Press. https://www.nap.edu/catalog/18872/a-framework-to-guide-selection-of-chemical-alternatives.

Nielsen, MA. 2012. *Reinventing Discovery: The New Era of Networked Science*. Princeton, N.J.: Princeton University Press. DOI: https://doi.org/10.2307/j.ctt7s4vx

O'Brien, M. 2000. *Making Better Environmental Decisions: An Alternative to Risk Assessment*. Cambridge: MIT Press.

OEHHA. 2015. "About Proposition 65." Text. OEHHA. January 12, 2015. https://oehha.ca.gov/proposition-65/about-proposition-65.

Palmer, S. 2016. "GreenScreen, SciVera Hit Back at ACC on Hazard Assessment Tools." *Chemical Watch* (blog). May 2016. https://chemicalwatch.com/47353/greenscreen-scivera-hit-back-at-acc-on-hazard-assessment-tools.

Schwarz, S and **Denison, RA.** 2018. "No Justification: Substantiations for Rampant New Chemical CBI Claims Are Deficient or Lacking Altogether." *EDF Health* (blog). February 16, 2018. http://blogs.edf.org/health/2018/02/16/no-justification-substantiations-for-rampant-new-chemical-cbi-claims-are-deficient-or-lacking-altogether/.

Schwarzman, MR and **Buckley, HL.** 2019. "Not Just an Academic Exercise: Systems Thinking Applied to Designing Safer Alternatives." *Journal of Chemical Education*, September. DOI: https://doi.org/10.1021/acs.jchemed.9b00345

Schweik, CM and **English, RC.** 2012. *Internet Success: A Study of Open-Source Software Commons*. Cambridge, Mass.: MIT Press. DOI: https://doi.org/10.7551/mitpress/9780262017251.001.0001

Scruggs, CE. 2013. "Reducing Hazardous Chemicals in Consumer Products: Proactive Company Strategies." *Journal of Cleaner Production* **44**(April): 105–14. DOI: https://doi.org/10.1016/j.jclepro.2012.12.005

Scruggs, CE and **Ortolano, L.** 2011. "Creating Safer Consumer Products: The Information Challenges Companies Face." *Environmental Science & Policy* **14**(6): 605–14. DOI: https://doi.org/10.1016/j.envsci.2011.05.010

Scruggs, CE, Ortolano, L, Schwarzman, MR and **Wilson, MP.** 2014. "The Role of Chemical Policy in Improving Supply Chain Knowledge and Product Safety." *Journal of Environmental Studies and Sciences* **4**(2): 132–41. DOI: https://doi.org/10.1007/s13412-013-0158-4

Stiglitz, JE. 1999. "Knowledge as a Global Public Good." In *Global Public Goods: International Cooperation in the 21st Century*. Kaul, I, Grunberg, I and Stern, MA (eds.), 308–25. New York: Oxford University Press.

Stone, A and **Delistraty, D.** 2010. "Sources of Toxicity and Exposure Information for Identifying Chemicals of High Concern to Children." *Environmental Impact Assessment Review* **30**(6): 380–87. DOI: https://doi.org/10.1016/j.eiar.2009.11.002

Tickner, JA, Schifano, JN, Blake, A, Rudisill, C and **Mulvihill, MJ.** 2015. "Advancing Safer Alternatives Through Functional Substitution." *Environmental Science & Technology* **49**(2): 742–49. DOI: https://doi.org/10.1021/es503328m

Toxics Use Reduction Institute. 2006. "Five Chemicals Alternatives Assessment Study." Lowell, MA: University of Massachusetts Lowell. https://www.turi.org/TURI_Publications/TURI_Methods_Policy_Reports/Five_Chemicals_Alternatives_Assessment_Study._2006.

United Nations. 2019. *Globally Harmonized System of Classification and Labelling of Chemicals (GHS)*. 8th Edition. Geneva. http://www.unece.org/index.php?id=51896&L=0.

United Nations Environment Programme. 2006. *Strategic Approach to International Chemicals Management: SAICM Texts and Resolutions of the International Conference on Chemicals Management*. Geneva: United Nations. http://www.saicm.org/Portals/12/documents/saicmtexts/SAICM-publication-EN.pdf.

United States Congress. 1986. *Emergency Planning and Community Right-to-Know Act. U.S. Code* **42** Chapter 116.

US Environmental Protection Agency. 2012. "Safer Choice Master Criteria for Safer Chemical Ingredients." *US EPA*. September 2012. https://www.epa.gov/saferchoice/safer-choice-master-criteria-safer-chemical-ingredients.

Vogel, SA. 2009. "The Politics of Plastics: The Making and Unmaking of Bisphenol A 'Safety'." *American Journal of Public Health* **99**(Suppl 3): S559–S566. DOI: https://doi.org/10.2105/AJPH.2008.159228

von Hippel, E. 2006. *Democratizing Innovation*. 1. MIT Press paperback ed. Cambridge, Mass.: MIT Press.

Weininger, D. 1988. "SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules." *Journal of Chemical Information and Modeling* **28**(1): 31–36. DOI: https://doi.org/10.1021/ci00057a005

Whittaker, MH and **Heine, LG.** 2013. "Chemicals Alternatives Assessment (CAA): Tools for Selecting Less Hazardous Chemicals." In *Chemical Alternatives Assessments*, Harrison, RM and Hester, RE (eds.), 1–255. *Issues in Environmental Science and Technology* **36**. Cambridge: Royal Society of Chemistry. DOI: https://doi.org/10.1039/9781849737234-00001

Wilson, MP and **Schwarzman, MR.** 2009. "Toward a New U.S. Chemicals Policy: Rebuilding the Foundation to Advance New Science, Green Chemistry and Environmental Health." *Environmental Health Perspectives* **117**(8): 1202–9. DOI: https://doi.org/10.1289/ehp.0800404

Zimmerman, JB and **Anastas, PT.** 2015. "Toward Substitution with No Regrets." *Science* **347**(6227): 1198–9. DOI: https://doi.org/10.1126/science.aaa0812

How to cite this article: Kokai, A, Blake, A, Dedeo, M and Lent, T. 2020. Building shared information infrastructure for chemical alternatives assessment. *Elem Sci Anth,* 8: 26. DOI: https://doi.org/10.1525/elementa.422

Domain Editor-in-Chief: Anne R. Kapuscinski, University of California, Santa Cruz, US

Associate Editor: Kevin Dooley, Department of Supply Chain Management, Arizona State University, US

Knowledge Domain: Sustainability Transitions

Part of an *Elementa* Forum: Multi-Stakeholder Initiatives for Sustainable Supply Networks

Submitted: 11 December 2019 **Accepted:** 28 May 2020 **Published:** 26 June 2020

ELEMENTA Science of the Anthropocene

Elem Sci Anth is a peer-reviewed open access journal published by University of California Press.

OPEN ACCESS