

<sup>1</sup> National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA  
<sup>\*</sup> Current affiliation: Infinitia ML, Durham, NC, USA  
<sup>2</sup> Kelly Services, Research Triangle Park, NC, USA

<sup>3</sup> International Agency for Research on Cancer (IARC) Monographs Programme, World Health Organization, Lyon, France  
<sup>4</sup> Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, TX, USA

## Abstract

Decision-makers and researchers frequently conduct literature-based assessments of the potential for chemicals or other exposures to pose a threat to human health. Such assessments typically consist of a critical review of a literature corpus to identify adverse health effects, to extract data for exposure-response relationship modeling, and/or to elucidate toxicity mechanisms. The systematic review methodology increases the transparency and objectivity in an evaluation by using a pre-defined, multistep process to identify, critically assess, and synthesize evidence. In addition to extraction of data, systematic review may also include an assessment of potential bias in a body of literature. A clear and detailed presentation of problem formulation, analysis and outputs, as well as properly documented search strategies and intermediate decisions, are critical to ensure transparency of the process.

We address these challenges by creating a modular, web-based content-management system to synthesize multiple data sources into overall human health assessments of chemicals. This free, open-source web-application, HAWC (Health Assessment Workspace Collaborative, <https://hawcproject.org/>), integrates and documents the overall workflow from literature search, literature screening, risk of bias assessment, data extraction, dose-response analysis using EPA benchmark dose modeling software (BMDs), and data synthesis by enabling creation of customizable visualizations of evidence and risk of bias.

Each HAWC assessment can be composed of some of all of these steps, based on the goals of the assessment, and at the discretion of assessment owners. User access is assessment-specific; project-managers can create public or private assessments, and can share with their team during development and ultimately release publicly as supplemental information to final reports (e.g., the US National Toxicology Program (NTP) monograph of **immunotoxicity associated with PFOA/PFOA exposure**, or the National Academy of Science's report on low-dose **toxicity from endocrine active chemicals**). All data and figures are exportable in user-friendly formats. To date, nearly 500 assessments have been created by users, and has been adopted for use by the NTP, the US EPA, TCEQ, and 34 assessments to date by the WHO IARC Monographs program.

Crucial benefits of such a system include improved integrity of the data and analysis results, greater transparency, standardization and consistency in data collection and presentation.

## Assessment and permissions

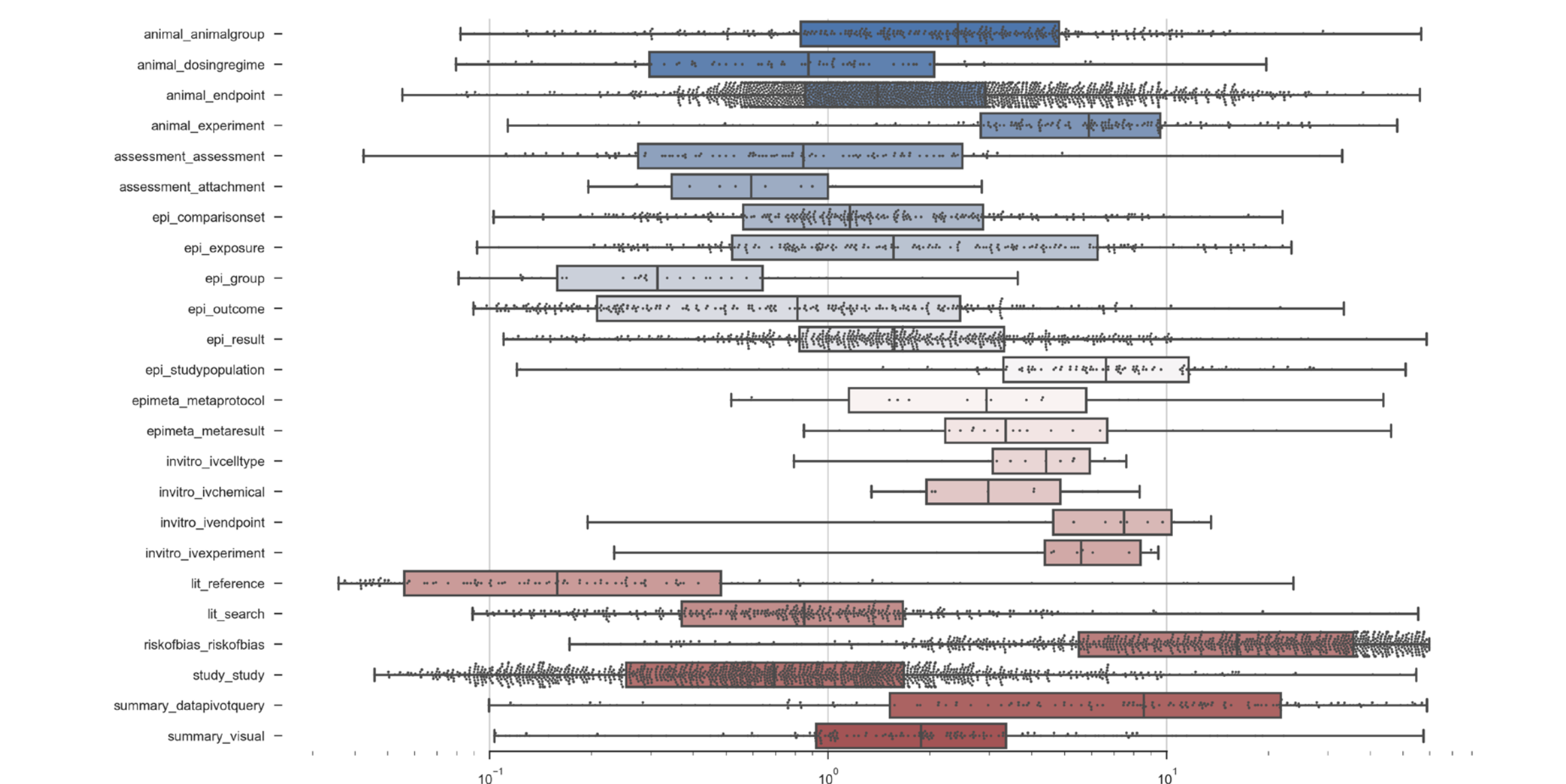
### Levels of access:

- Project managers:** change permissions settings, including who can edit assessment content and which modules are enabled
- Team members:** add, edit, and delete content
- Reviewers:** view assessment before assessment is public
- Public:** if an assessment is made public, the general-public can view

- ### Other assessment features
- Project managers can control which modules to enable/disable, settings are project specific
  - Team members can assign tasks for risk of bias assessment or data extraction tasks

HAWC is a web-based content management system to create, store, share, and display data and results in order to conduct human health assessments

## Time spent extracting data and risk of bias



Analytics tracking the amount of time spent editing content in HAWC has been collected since June 21, 2017. While the results are still early, over 8,700 items of content have been edited in HAWC to date. In the future, these data may be informative in estimating the total time/cost of data extraction in systematic reviews, and may help drive cost-benefit analyses for machine learning and semi-automated extraction methods.

## Application Modules

### LITERATURE REVIEW

**Malathion, cancer in humans**

**Description:** Search for epidemiologic studies of cancer in humans exposed to malathion

**Search type:** Search

**Search database:** Pubmed

**Search text:** ("malathion" [All Fields] OR malathion OR World OR "malathion" [MeSH] AND ("cancer" [MeSH] OR neoplasm OR cancer OR carcinogen) OR hazard AND ("epidemiology" [MeSH] OR "epidemiologic studies" [MeSH] OR epidemiology OR case-reports OR "occupational exposure" [MeSH] OR cases OR cohort)

**Created:** Oct 25, 2014, 10:51 a.m.

**Last updated:** Oct 25, 2014, 10:51 a.m.

**Results from queries:**

Date last executed	Total references found	References added	References removed
April 3, 2014, 7:14 a.m.	21	5	1
March 2, 2014, 3:32 a.m.	17	0	0
Feb. 11, 2014, 9:37 a.m.	17	0	0
Oct. 25, 2014, 10:51 a.m.	17	17	0

**Inclusion:** 1,3-butadiene, 4-Aminodiphenyl, Aftatoxin, Benzene, Human, In vitro, DNA methylation, In vivo, DNA methylation, Histone methylation, microRNA, lncRNA

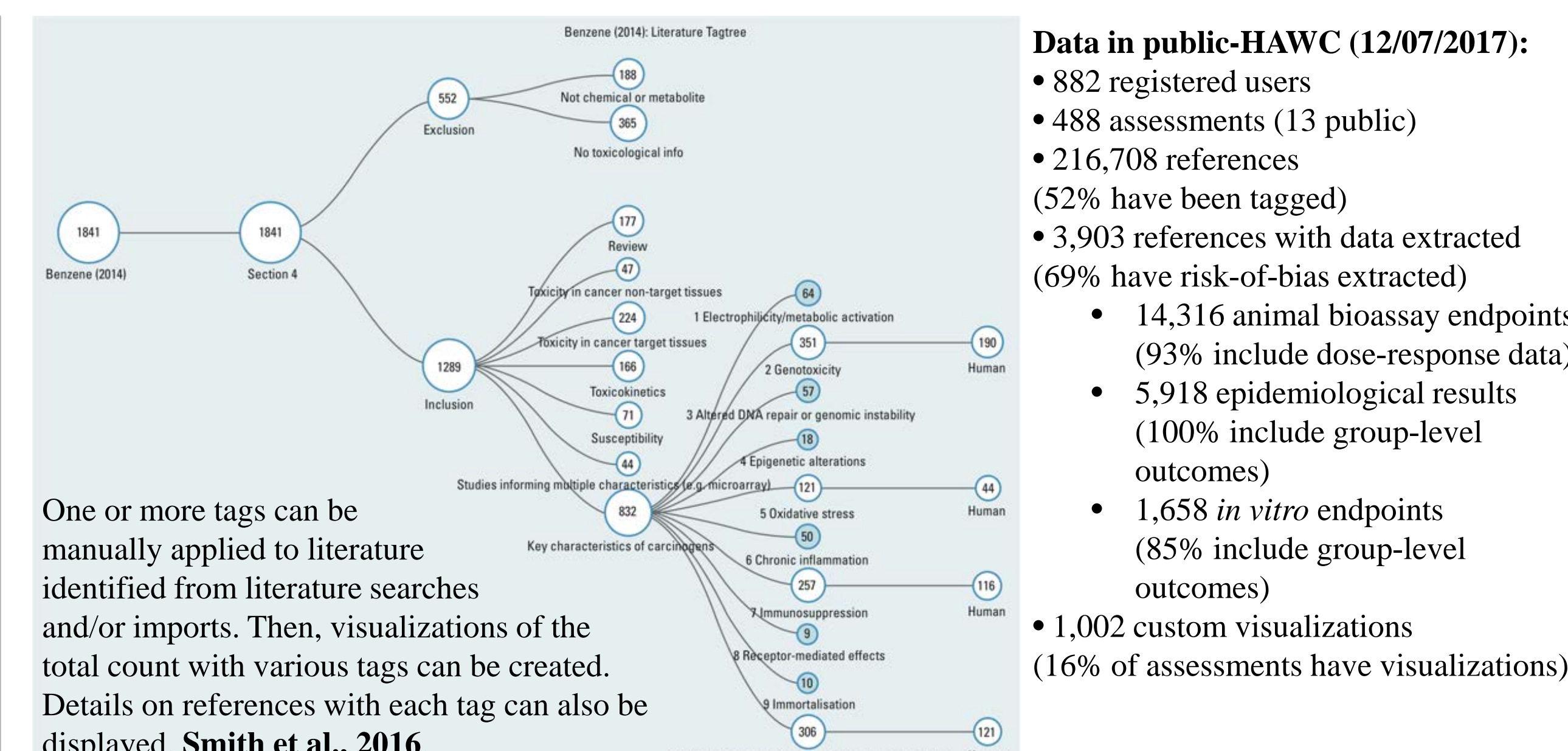
**Exclusion:** no significance, y-HA, Histone-carbonium binding, not chemical of interest, cannot access full text, review or comment, other, carcinogen binding to histone, methylated cytosines, carcinogen binding and resultant chromatin structure

**References can be searched in Pubmed, or imported using IDs from PubMed, HERO, or an entire reference database can be imported from Endnote or RefManager.**

**The example above shows a PubMed query from the IARC Monographs. The query was executed four times, with the final time finding five new references and removing one.**

**Literature tags are fully customizable for each assessment; they are also nested and counts roll-up in visualizations and other presentations.**

**Chappell et al., 2016**



### ENDPOINT EXTRACTION

novel object test (time exploring object 2, training phase)

**Endpoint Details:** Shows a table of study details and a line graph of results for 'Time exploring object 2' across different groups. The graph shows a clear upward trend in the experimental group compared to the control.

**Forest plot:** A forest plot showing the mean and standard deviation for 'Time exploring object 2' across different groups and studies, with a clear separation between the experimental and control groups.

**Results by group:** A table summarizing the mean values and standard deviations for the experimental and control groups.

**Extraction of animal bioassay endpoints, including group-level summary data if available, are extracted for different animal groups. An animal group is defined as group of related animals for which dose-related response differences are compared. NTP Fluoride 2016.**

### RISK OF BIAS

**NTP Fluoride 2016**

**Visualizations of the total count with various tags can be created. Details on references with each tag can also be displayed. Smith et al., 2016.**

**Risk of bias questions can be defined for each assessment and are answered for each study. Multiple reviewers may be required for a response with conflict resolution, or a single reviewer may be used for each study. Heatmaps of responses or strengths/weaknesses across the evidence can be created, to view areas of potential bias across a body of literature.**

**NAS 2017**

## PROJECT MANAGEMENT TOOLS

**Task assignments:** A screenshot of the task assignment interface showing a list of tasks with columns for status, assigned to, and due date.

**Effect data cleanup:** A screenshot of the effect data cleanup interface showing a table of effects with columns for name, organ, effect, and status.

**Tasks can be assigned to team-members and completion of tasks can be tracked by the team. Tasks can be sorted and filtered by study type. Simple task-summary charts and graphs can be presented for project tracking and analytics for completion.**

## DATA ANALYSIS AND PRESENTATION

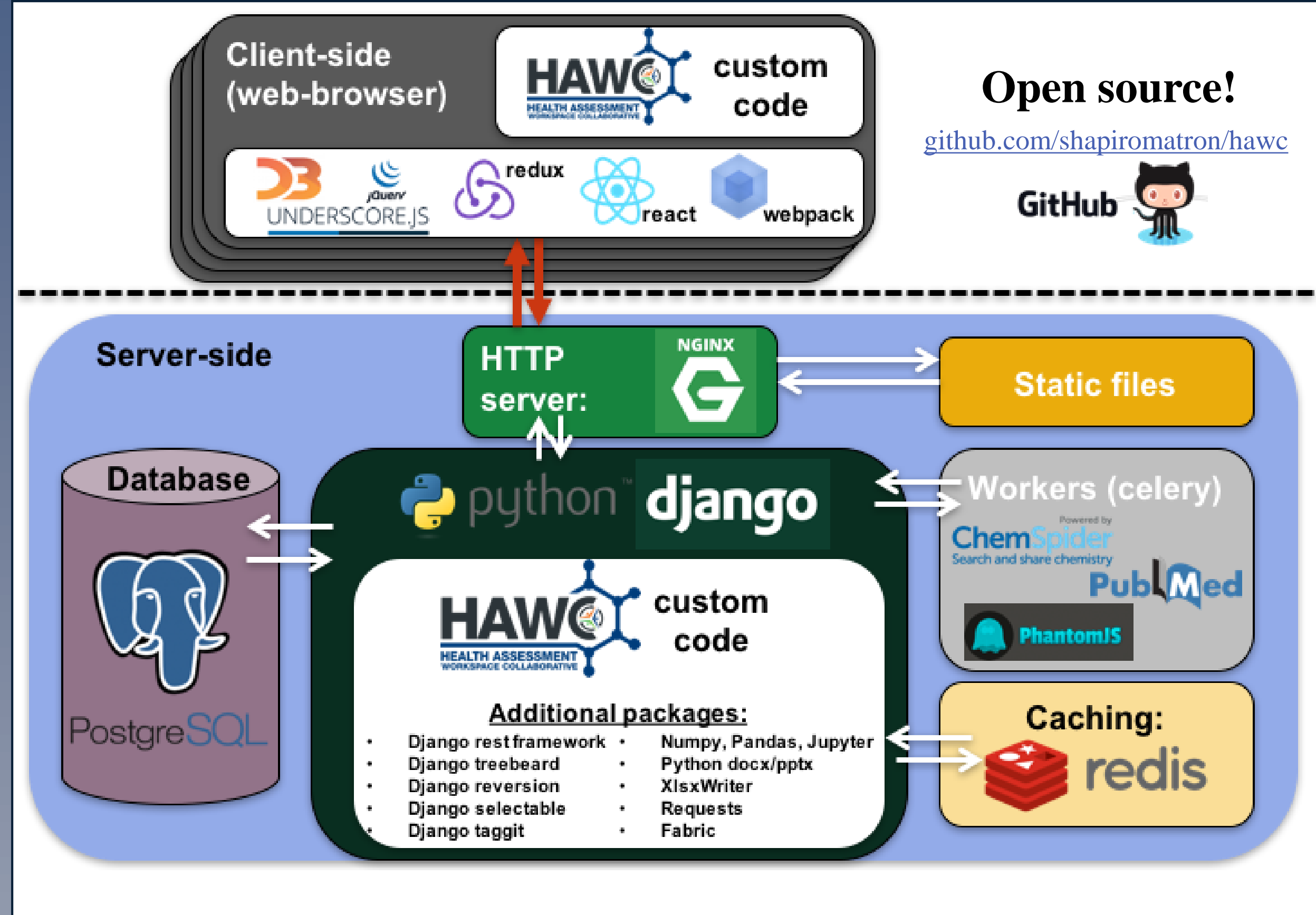
**Benchmark dose modeling using the EPA Benchmark Dose Modeling software. HAWC allows users to customize which models to apply to their datasets; models are executed and model outputs are stored. This example shows the selected model (Log-logistic) and a different model (Weibull) demonstrating slightly different curve fits on a dichotomous dataset in HAWC.**

**A data pivot created to compare similar endpoints from multiple studies. Data presented is customizable (using sort/filter rules), along with which columns of text to present, and rules for how to apply styles to points can be set manually or with conditional formatting. Here significant response are shown in red. NAS 2017.**

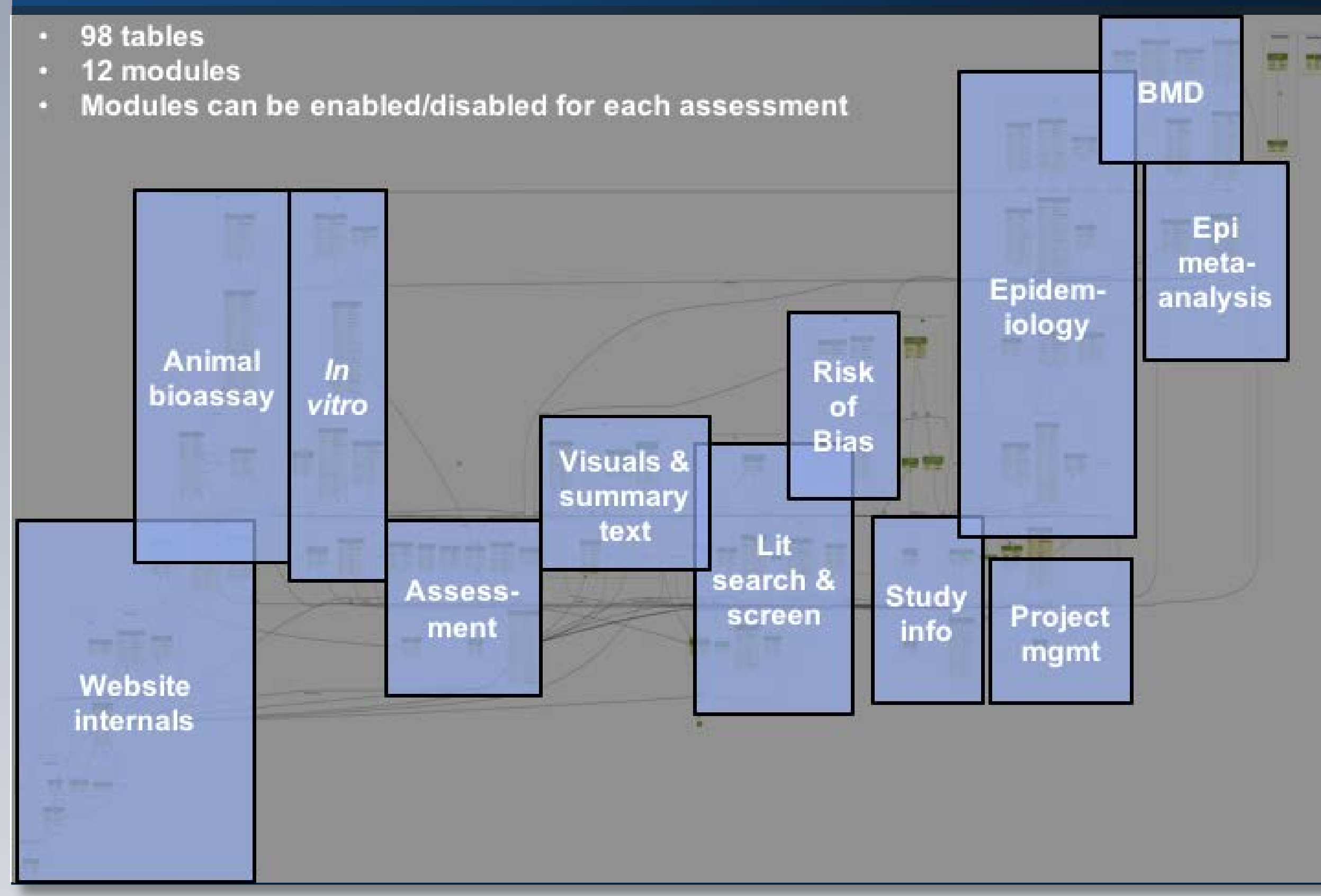
**Analytics tracking the amount of time spent editing content in HAWC has been collected since June 21, 2017. While the results are still early, over 8,700 items of content have been edited in HAWC to date. In the future, these data may be informative in estimating the total time/cost of data extraction in systematic reviews, and may help drive cost-benefit analyses for machine learning and semi-automated extraction methods.**

**A third data pivot showing one row for each in vitro endpoint. Instead of presenting mean response values, red triangles represent response significantly higher than control, and blue triangles represented response significantly lower than control. Pelch et al. (2017).**

## Technical implementation



## Data model



## Example published projects which used HAWC

- Publications citing HAWC include peer-reviewed publications, or final reports from organizations which indicated use of HAWC during analysis or methods guidelines and may not be a complete list:
  - Boyles AL, Blain RB, Rochester JR, Avanas J, Goldhaber SB, McComb S, Holmgren SD, Masten SA, Thayer KA. Systematic review of community health impacts of mountaintop removal mining. *Environ Int.* 2017;107:163-172. [10.1016/j.envint.2017.07.002](https://doi.org/10.1016/j.envint.2017.07.002)
  - Chappell G, Pogribny IP, Guyton KZ, Rusyn I. Epigenetic alterations induced by genotoxic occupational and environmental human chemical carcinogens: A systematic literature review. *Mutat Res Rev Mutat Res.* 2016;768:27-45. [10.1016/j.mrev.2016.03.004](https://doi.org/10.1016/j.mrev.2016.03.004)
  - Guha N, Guyton KZ, Loomis D, Banual DK. Prioritizing Chemicals for Risk Assessment Using Chemoformatics: Examples from the IARC Monographs on Pesticides. *Environ Health Perspect.* 2016; 124(12):1823-1829. [10.1289/EHP186](https://doi.org/10.1289/EHP186)
  - IARC Monographs. 2017. Instructions to Authors for the Preparation of Drafts for IARC Monographs. [http://monographs.iarc.fr/ENG/Preparation/Instructions\\_to\\_Authors.pdf](http://monographs.iarc.fr/ENG/Preparation/Instructions_to_Authors.pdf)
  - Molander L, Hanberg A, Rudén C, Agnerstrand M, Beronius M. Combining web-based tools for transparent evaluation of data for risk assessment; developmental effects of bisphenol A on the mammary gland as a case study. *J Appl Toxicol.* 2017; 37(3):319-330. [10.1002/jat.3363](https://doi.org/10.1002/jat.3363)
  - National Academies of Sciences, Engineering, and Medicine. 2017. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. Washington, DC: The National Academies Press. [10.17226/24758](https://doi.org/10.17226/24758)
  - National Toxicology Program. 2015. Handbook for Conducting a Literature-Based Health Assessment Using OHAT Approach for Systematic Review and Evidence Integration. [https://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookjan2015\\_508.pdf](https://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookjan2015_508.pdf)
  - National Toxicology Program. 2015. Monograph on Identifying Research Needs for Assessing Safe Use of High Intakes of Folic Acid. 2015. Research Triangle Park, NC: National Toxicology Program. [https://ntp.niehs.nih.gov/ntp/ohat/folicacid/final\\_monograph\\_508.pdf](https://ntp.niehs.nih.gov/ntp/ohat/folicacid/final_monograph_508.pdf)
  - National Toxicology Program. 2016. NTP Research Report on Systematic Literature Review on the Effects of Fluoride on Learning and Memory in Animal Studies. Research Triangle Park, NC: National Toxicology Program. [https://ntp.niehs.nih.gov/ntp/results/pubs/rrreports/01/fluoride\\_508.pdf](https://ntp.niehs.nih.gov/ntp/results/pubs/rrreports/01/fluoride_508.pdf)
  - National Toxicology Program. 2016. Monograph on Immunotoxicity Associated with Exposure to Perfluorooctanoic acid (PFOA) and perfluorooctane sulfonate (PFOS). Research Triangle Park, NC: National Toxicology Program. [https://ntp.niehs.nih.gov/ntp/ohat/pubs/pfos/pfos\\_monograph\\_508.pdf](https://ntp.niehs.nih.gov/ntp/ohat/pubs/pfos/pfos_monograph_508.pdf)
  - Pelch, KE, Wignall, JA, Goldstone, AE, Ross, PK, Blain, RB, Shapiro, AJ, Holmgren, SD, Hsieh, J-H, Svoboda, D, Auerbach, SS, Parham, FM, Masten, SA, Thayer, KA. 2017. NTP Research Report on Biological Activity of Bisphenol A (BPA) Structural Analogues and Functional Alternatives. NTP RR 4. Research Triangle Park, NC: National Toxicology Program. (4): 1-78. [https://ntp.niehs.nih.gov/ntp/results/pubs/rrreports/rr04\\_508.pdf](https://ntp.niehs.nih.gov/ntp/results/pubs/rrreports/rr04_508.pdf)
  - Smith MT, Guyton KZ, Gibbons CF, Fritz JM, Portier CJ, Rusyn I, DeMarini DM, Caldwell JC, Kavlock RJ, Lambert PF, Hecht SS, Bucher JR, Stewart BW, Baan RA, Coglianor VJ, Straif K. Key Characteristics of Carcinogens as a Basis for Organizing Data on Mechanisms of Carcinogenesis. *Environ Health Perspect.* 2016; 124(6):713-21. [10.1289/ehp.1509912](https://doi.org/10.1289/ehp.1509912)
  - Texas Commission on Environmental Quality. 2016. Ethylene Glycol Development support document. <https://www.tceq.texas.gov/assets/public/implementation/tox/dsd/final/fcb/6EG.pdf>
  - Texas Commission on Environmental Quality. 2017. White Paper: TCEQ Guidelines for Systematic Review and Evidence Integration. [https://www.tceq.texas.gov/assets/public/implementation/tox/dsd/whitepaper/Proposed%20Systematic\\_Review.pdf](https://www.tceq.texas.gov/assets/public/implementation/tox/dsd/whitepaper/Proposed%20Systematic_Review.pdf)

## Contact information:

Andy Shapiro ([shapiromatron@gmail.com](mailto:shapiromatron@gmail.com))  
 Ivan Rusyn ([irusyn@tamu.edu](mailto:irusyn@tamu.edu))

