Sigfried Gold

Medical informaticist, healthcare data analytics researcher and developer

sigfried at sigfried dot org – web – github – UMD iSchool – OHDSI

Sigfried is a medical informaticist, researcher, applications architect, and research programmer specializing in the design and development of tools and infrastructure facilitating the exploratory visualization and analysis of real-world patient data. In recent years they have studied and supported the analysis, construction, and reuse of clinical code sets in computational phenotypes.

Work History

Johns Hopkins University, Biomedical Informatics and Data Science, General Internal Medicine Sr Research Informatics Applications Architect, June 2021—present. Support the National COVID Cohort Collaborative (N3C), Clinical Data to Health (CD2H) coordination project by leveraging implementation and subject matter expertise in biomedical, translational, clinical data standards and Real-World data (RWD) utilization. Support program investigator analyses, data alignment, and quality infrastructure teams. Design and develop software to support the analysis and development of code sets (aka, value sets, concept sets) for use in RWD studies.

ICF, subcontracting to the National Library of Medicine Expert consultant, data science, December 2018–September 2019: Analysis and programming on distributed HIV studies, including a diverse set of clinical trials and a real-world-evidence research network using the OMOP/OHDSI data model. Mostly in R and PostgreSQL.

National Socio-Environmental Synthesis Center Research Assistant, June 2018–October 2018: Data science, visualization consultant to NSF-funded research center of the University of Maryland.

Columbia University, Department of BioMedical Informatics Consultant, April 2017–March 2018: Visualization consultant on NSF grant, "Integration of Environmental Factors and Causal Reasoning Approaches for Large-Scale Observational Health Research," integrating OHDSI patient data with environmental exposure data.

University of Colorado, Health Data Compass Consultant, September 2016–June 2017: Built innovative visualization tools for researchers using OHDSI/OMOP warehouse of Children's Hospital and UCHealth data. Drug Exposure/Gap Explorer. Vocabulary/Population Navigator. Served as visualization consultant and designed and prototyped new data quality visualization tools for PCORI grant.

Johnson & Johnson, Janssen Research Consultant, March 2016–October 2016: Built visualization tools into ATLAS, the primary Web interface for OHDSI and the OMOP Common Data Model: patient profile, graphical display of cohort definitions, generalized scatterplot module. Also added many utilities to aid in further ATLAS development.

FasterCures Consultant, 2015: Software architecture consulting.

Gryphon Scientific for NIAID/NIH Consultant, 2015: Technical consulting, software design, software development. React, D3.js, Node, Python. Built interface for querying and annotating publication references from PubMed.

Social & Scientific Systems, Inc. Interactive Information Visualization Designer, October, 2011–Feb 2014. Designed static and interactive information visualization displays for scientific and operational data in the context of large clinical trials research.

Oracle Corporation (previously Phase Forward) Senior Systems Consultant, June 2008–September 2011. Designed and prototyped research software for discovering and validating drug safety signals in longitudinal healthcare data. Worked with DoD TRICARE data warehouse of 12 million patient medical records. Was Oracle's principal investigator for a $100K grant to UMD/HCIL EventFlow project.

Columbia University Medical Center Systems Consultant, January 2003–March 2008. Worked with researchers, neonatologists and administrators at the New York Presbyterian Hospital Regional Perinatal Centers (RPC) on quality assurance and research applications using neonatal intensive care unit data. The RPC has New York State-mandated data oversight responsibilities for 15 NICUs around New York City. Projects involved collection of data from multiple sources, including New York State's Perinatal Database System, New York City's and New York State's electronic birth certificates, Vermont Oxford, and the NICUs themselves.

OJC Technologies/Pixo Founder, CIO, August 1998–January 2003. As founder, led all aspects of soliciting and executing contracts: sales, proposals, business analysis, project management, software architecture, implementation. Managed significant projects for dozens of clients including UIUC/EvaluationOnline, Morgan Stanley Dean Witter, Imagine Media/Daily Radar, MortgageSelector, Trimark Securities, United States Geological Survey, Stats Inc., Fitch IBCA Duff & Phelps, McGraw-Hill, Illinois Department of Children and Family Services, Illinois Math and Science Academy, Music Browser Inc., Provena Covenant Medical Center, Psychics.com. Company grew from two employees in August 1998 to twenty in 2000. Pixo remains the leading web application development firm in Champaign-Urbana, Illinois, providing services to many departments and offices of the University of Illinois and other local business and organizations.

Prior to 1998: Tradetech Securities, Sybase, Allstate Insurance Company, MetLife, Dean Witter Reynolds, Victim Services Agency.

Prior to first programming job: Secretary, bookstore clerk, tutor, construction worker, doorman, security guard, driver, clown, bouncer, barista, busboy, short order cook, dishwasher, corn detassler.

Education

University of Maryland College Park, PhD candidate. iSchool, HCIL, COMBINE Fellow (2017 – present)

Columbia University, M.A. in Biomedical Informatics (2008)

Sarah Lawrence College, M.F.A. in Creative Writing (1987)

Goddard College, B.A. in Literature (1984)

Fellowships and Awards

University of Maryland COMBINE (Computation and Mathematics for Biological Networks) Fellow - **National Science Foundation Research Traineeship** (Award DGE-1632976) 2018 – 2022. Certificate awarded 8/19/22.

Best Poster. Observational Health Data Sciences and Informatics Symposium, Washington, DC. **Gold**, Blacketer, Sena, and DeFalco. CHRONOS: Cohort exploration through individual patient profiles. 2016.

Best Paper. International Conference on eHealth, Telemedicine, and Social Medicine: eTELEMED. Zhu, **Gold**, Lai, Hripcsak and Cimino. Using Timeline Displays to Improve Medication Reconciliation. 2009.

Best Paper. American Medical Informatics Association Annual Symposium. Washington, DC. **Gold**, Elhadad, Zhu, Cimino and Hripcsak. Extracting Structured Medication Event Information from Discharge Summaries. 2008.

Peer Reviewed Publications and Presentations

S Gold, T Zhang, RL Zhu, S Hong, HP Lehmann…"ICD10–SNOMED mapping pitfalls: Post-coordinated expressions and concept sets". OHDSI. 2022 Oct 21; https://www.ohdsi.org/wp-content/uploads/2022/10/21-Gold-USE-abstract.pdf

Pfaff, Girvin, Gabriel, Kostka, Morris, Palchuk, **Gold**, et al. Synergies between centralized and federated approaches to data quality: a report from the national COVID cohort collaborative. Journal of the American Medical Informatics Association. 2021 Nov 2; https://doi.org/10.1093/jamia/ocab217

Alper BS, Flynn A, Bray BE, Conte ML, Eldredge C, **Gold S**, et al. "Categorizing metadata to help mobilize computable biomedical knowledge." Learning Health Systems. 2021;e10271. https://doi.org/10.1002/lrh2.10271

Fung, Xu, **Gold**. "The Use of Inter-terminology Maps for the Creation and Maintenance of Value Sets." AMIA Annual Symposium. American Medical Informatics Association, Washington, D.C. 2019. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7153132/.

Mayer C, Williams N, **Gold S**, Fung KW, Huser V. "Sharing of Individual Participant Data from Clinical Trials: General Comparison and HIV Use Case." AMIA Annual Symposium, 2019. (Short paper, preprint)

Huser V, Mayer C, Williams N, **Gold S**, Fung KW. "Set of derived metadata elements for comparing cohorts from human clinical trials datasets and EHR." OHDSI Symposium 2019. (Poster)

Mayer C, Williams N, **Gold S**, Fung KW, Huser V. "Validation of Real World Data: Case study in hepatitis C." OHDSI Symposium 2019. (Poster)

Williams N, Huser V, Mayer C, **Gold S**, Fung KW. "Case Studies Beware! Rx and Labs return minor improvements in HIV prevalence capture for Medicare Parts D and B for 2012 to 2016." OHDSI Symposium 2019. (Poster)

* Gold, Batch, McClure, Jiang, Kharrazi, Saripalle, Huser, Weng, Roderer, Szarfman, Elmqvist, Gotz. "Infrastructures and Interfaces to Encourage Value Set Reuse for Health Data Analytics." AMIA Annual Symposium, 2018. (Full paper)

Gold, Hochheiser, Boyce, Elmqvist. "P3 — Partition, Pivot, and Prune: Aggregated Semantic Graphs for Analyzing Medical Terminologies." Visual Analytics in Healthcare, San Francisco, November 2018. (Poster)

Weng, Mosa, Ahuja, Szarfman, Solomonides, Huser, **Gold**, Zozus. "Community Strengthening and Knowledge Sharing Towards Systematic and Scalable Clinical Data Quality Assessment." Panelist for AMIA Annual Symposium, 2018. (Panel)

* Gold, Blacketer, Sena, and DeFalco. "CHRONOS: Cohort exploration through individual patient profiles." Observational Health Data Sciences and Informatics Symposium, Washington, DC. 2016. (**Best poster**)

Gove, Saxe, **Gold**, Long, and Bergamo. "SEEM: A Scalable Visualization for Comparing Multiple Large Sets of Attributes for Malware Analysis." In Proceedings of the 11th International Workshop on Visualization for Cyber Security. Paris, France. 2014. (Paper)

Meyer, Monroe, Plaisant, Lan, Wongsuphasawat, Coster, **Gold**, Millstein, Shneiderman. "Visualizing Patterns of Drug Prescriptions with EventFlow: A Pilot Study of Asthma Medications in the Military Health System." Proc: Workshop on Visual Analytics in Healthcare. 2013. (Paper)

* Gold, Xie, Taylor, Szarfman and Coster. "A Sorting and Classification Method for Interactive Visualization of Drug/Event Patterns across Many Patient Timelines." 26th International Conference on Pharmacoepidemiology and Therapeutic Risk Management. Brighton, UK. 2010.

* Zhu, Gold, Lai, Hripcsak and Cimino. "Using Timeline Displays to Improve Medication Reconciliation." International Conference on eHealth, Telemedicine, and Social Medicine: eTELEMED. 2009. (**Best paper**)

* Gold, Elhadad, Zhu, Cimino and Hripcsak. "Extracting Structured Medication Event Information from Discharge Summaries." American Medical Informatics Association Annual Symposium. Washington, DC. 2008. (**Best paper**, code)

Other Publications and Presentations

Alper, Bray, Conte, Eldredge, Flynn, **Gold**, Greenes, et al. 2020. "Developing Metadata Categories as a Strategy to Mobilize Computable Biomedical Knowledge." https://deepblue.lib.umich.edu/handle/2027.42/155655.

Richesson, Wiley, **Gold**, Rasmussen. 2020. "Electronic Health Records-Based Phenotyping." Rethinking Clinical Trials: A Living Textbook of Pragmatic Clinical Trials. https://rethinkingclinicaltrials.org/chapters/conduct/electronic-health-records-based-phenotyping/electronic-health-records-based-phenotyping-introduction/.

Gold, Elmqvist. "Inconsistent and Overlapping Ontologies for Coded Data Analytics." 30th Human-Computer Interaction Lab Symposium. University of Maryland. 2018.

DeFalco, **Gold**, Klebanov, Evans, OHDSI Development Architecture. OHDSI Symposium, Bethesda, MD. 2017. (Tutorial)

Gold. Exploratory visualization tools for health records research, and an exciting detour into infrastructural support for health records research at UMD. HCIL BBL, 2017.

Gold. "Interactive Visualization of Complex Clinical Research Metadata." 30th Human-Computer Interaction Lab Symposium. University of Maryland. 2013.

Gold. "Using Visualization to Explore Claims and EHR Data for Signal Strengthening." Drug Information Association Annual Meeting, Chicago. 2011.

Gold and Coster. "Visual Representation of Exposure Patterns in Drug Safety Research." 28th Human-Computer Interaction Lab Symposium, Electronic Health Record Informatics Workshop. University of Maryland. 2011.

Gold, Lehmann, Lutters Schilling. "Practices, norms, and aspirations regarding the construction, validation, and reuse of code sets in the analysis of real-world data." Journal of Biomedical Informatics. 2021.

Kubick and **Gold**. "Safety in Numbers." Good Clinical Practice Journal. 2008.

Service

Journal of Biomedical Informatics, 2022, Reviewer
JAMIA, 2018, 2019, Reviewer
Visual Analytics in Healthcare, 2018, Program Committee, Reviewer
AMIA Annual Symposium, 2017, 2018, Reviewer
OHDSI 2017, 2018 Symposium Planning Committee
OHDSI 2017 Symposium Tutorial Co-Leader
OHDSI Architecture Working Group (2016-2017)
OHDSI GIS Working Group (2017)
OHDSI Vocabulary Visualization Working Group, Leader (2016)

OHDSI Patient Visualization Working Group (2015-2016)
Data Quality Collaborative Code-A-Thon (Nov 2015) and follow-up participation. ([Code](#))

Expertise in tools for software development, data science, research, and informatics

Medical informatics tools and terminologies: OHDSI/OMOP, N3C, FHIR, VSAC, ICD-9/10, SNOMED, MedDRA, UMLS, RxNorm, 3M's HDD, FDB NDDF+, LOINC.

Data science: Python, R, SQL (PostgreSQL, Oracle, SQL Server, MySQL, Sybase, DB2, etc.), Apache Spark, MongoDB, CouchDB, ElasticSearch, D3.js, Excel, Jupyter, Pandas, SQLAlchemy, ggplot, Matplotlib, Seaborn, graph-tool, Palantir Foundry, Tableau, Spotfire, GraphQL, Sigma.js, Cytoscape, Graphviz, Gephi, Qualtrics.

Web front end: JavaScript, ES6, HTML, CSS, SVG, React, jQuery, Lodash, Redux, R Shiny, Angular 1, Jasmine, Mocha, Vows, Bootstrap, Material, Webpack, RxJs, Meteor, etc., etc.

Web back end: Node, Tornado, Flask, MongoDB, Elasticsearch, Loopback, Perl, Java, Postico.

Languages: JavaScript, Python, R, C/C++, bash, Perl, VB.

OS and misc: Unix/Linux (Ubuntu, REHL/CentOS, etc.); Vi/Vim, Git, Regex.

Productivity, publishing, design, and wireframing: MS Office, Visio; LaTeX, Markdown, Pandoc, Zotero, BibTeX; Roam, Obsidian, Good Notes, various open-source tools for zettelkasten, notetaking, blogging, document management, and publishing; Adobe Illustrator, Procreate, Balsamic, Sketch.