

Charles T. Gray, PhD

Applied Mathematician | Data Scientist | Analytics & Data Engineer | Statistical Software Specialist

Canberra (ACT) · Australian Citizen

[GitHub](#) – [softloud](#) · [Publications](#) · charlestigray@gmail.com · [Data storytelling gallery](#)

Profile

Applied mathematician and senior scientific programmer with over a decade of experience designing, building, and maintaining analytical systems across research, industry, and policy contexts.

I specialise in **scientific programming, analytics engineering, and data visualisation**, with a strong focus on reproducibility, validation, and governance. I design analytical systems that can be trusted: auditable, constraint-aware, and observable from ingestion through to decision-making.

My work spans statistical modelling, simulation, and large-scale data pipelines, with particular strength in transforming complex, heterogeneous data into clear, verifiable analytical outputs. I am recognised for high-quality **data visualisation and data storytelling**, ensuring results remain interpretable and defensible for technical and non-technical stakeholders.

I bring particular depth in **graph-structured and entity-aware data modelling**, complementing strong capability across statistical analysis and analytics engineering. This allows me to reason precisely about identity, equivalence, uncertainty, and dependency in complex analytical workflows.

I work effectively across disciplines, value clear scope and delivery standards, and contribute calmly to process-oriented teams, particularly in regulated or high-assurance environments.

Core Specialisations

- Scientific programming (Python, R, SQL)
- Analytics engineering and workflow orchestration
- Statistical modelling, simulation, and uncertainty analysis
- Data visualisation and data storytelling
- Reproducible analytics and FAIR (findable, accessible, interoperable, reusable) data systems
- Data validation, testing, and auditability
- Graph-based and entity-aware data modelling
- Data governance, lineage, and quality frameworks
- Technical documentation and stakeholder translation
- Secure and regulated data environments

Professional Experience

Independent Data Architect, Strategist, Engineer

Australia · Denmark · UK | 2018 – Present

Technologies: dbt · Dagster · Python · R · SQL · Snowflake · AWS Redshift · DuckDB · Fivetran · Tableau · PowerBI · pandas · scikit-learn · numpy · ggplot2 · Shiny · Git · GitHub · GitLab · CI/CD · Agile · JIRA · ClickUp

Design and implement analytical and data engineering systems within large, distributed ecosystems, with an emphasis on reliability, validation, and maintainability.

- Built end-to-end analytical pipelines integrating SQL and Python, orchestrated with dbt and Dagster
- Designed validation frameworks covering data quality, lineage integrity, and analytical assumptions
- Architected relational and graph-based data models to preserve identity and semantics across sources
- Led data migration, governance, ESG, and documentation in regulated and high-risk environments
- Developed interpretable machine-learning workflows with strong focus on validation and leakage prevention
- Produced clear technical documentation and supported teams in workflow and delivery standards

Clients included a national telecommunications provider, research groups, a neuromarketing firm, and a video-game studio.

Impact: Delivered auditable analytical systems handling millions of daily events; reduced technical debt and improved confidence in decision-critical data products.

Research Software Engineer / Data Scientist

Academic & Policy Institutions (Australia & EU) | **2012 – 2020**

Technologies: R · Targets · ggplot2 · Shiny · Git · GitHub · reproducible research tooling

Developed research-grade analytical software and reproducible data pipelines across biomedical, ecological, and policy domains.

- Designed FAIR, analysis-ready data pipelines with explicit modelling of uncertainty and equivalence
- Developed R packages for simulation, estimation, validation, and visualisation
- Collaborated with multidisciplinary teams to resolve data interoperability challenges
- Published peer-reviewed methods in reproducible data science and evidence synthesis
- Lectured and coordinated subjects including graph theory, linear algebra, automata, linear programming, and generalised linear models

Organisations included the Walter and Eliza Hall Institute, University of Melbourne, La Trobe University, and EU policy research collaborations.

Impact: Delivered validated analytical systems adopted in peer-reviewed research and policy reporting; contributed open-source tools used internationally.

Music Career to STEM Transition

1997 – 2017

Pianist → Applied Mathematician / Data Scientist

Transitioned from professional music performance and pedagogy into mathematics and data science beginning in 2011. This background developed strong communication skills, discipline, and ensemble problem-solving ability, and informed my approach to formal systems, structure, and pattern recognition. Performed as an emerging mathematician at the opening ceremony of the 2018 Heidelberg Laureate Forum.

Education

Qualification	Institution	Focus
PhD (Data Science) , 2020	La Trobe University	Reproducible analytics, statistical simulation, validation frameworks · <i>Toward a Measure of code::proof</i>
BSc (Hons I, Mathematics) , 2015	La Trobe University	Graph theory, universal algebra · Published thesis in <i>Order</i>
BA / BMus , 2007	University of Melbourne	Critical theory, musicology

Academic distinctions: First Class Honours (87% aggregate). Published original mathematical result in quasi-primal algebra. Awarded multiple research grants.

Selected Invited Talks & Outreach

- DBT Meetup Copenhagen (2023–2024) — *A dbt Grrl in a dbt World*
- Copenhagen useR Group (2024) — *Animating a Melody as a Mathematical Object*
- PyData Copenhagen (2023) — Open video-game data analysis
- Heidelberg Laureate Forum (2018) — Opening Ceremony contributor
- Code Like a Girl · AMSI Choose Maths · World Science Festival — National speaking engagements

Selected Publications

- Davey, B.A., Gray, C.T., Pitkethly, J.G. (2018). *The Homomorphism Lattice Induced by a Finite Algebra*. **Order**
- Gray, C.T., Marwick, B. (2019). *Truth, Proof, and Reproducibility*. Springer
- Gray, C.T. (2019). *code::proof*. Springer
- Haddaway, N.R., Gray, C.T., et al. (2021–2022). Evidence synthesis databases, citation graphs, and visualisation tools (*Environmental Evidence, Research Synthesis Methods*)

Referees available on request.