MATAI: A Unified Interactive Platform for AI-Driven Alloy Discovery

Yixuan Li [1] **Yanchen Deng** [1] **Chendong Zhao** [1] **Penghui Yang** [1] **Jianguo Huang** [1] **Yuhao Lu** [1] **Yushan Xiao** [1] **Xuyu Dong** [1] **Zhonghan Zhang** [1] **Cuntai Guan** [1] **Bijun Tang** [1] **Xinrun Wang** [2] **Zheng Liu** [1] **Bo An** [1]

Abstract

AI-driven materials discovery is increasingly bottlenecked not by individual models but by the fragmentation of the surrounding tooling: composition–property databases, property predictors, and constrained optimisers typically live in disjoint projects, forcing researchers to hand-stitch scripts, file formats, and provenance. We present MATAI, a unified, browser-based platform that integrates visual data analytics, composition-aware property prediction, and simulated-annealing-based constrained alloy design into a single closed-loop workflow. A conversational layer exposes the same prediction and design engines through natural-language commands, and a document-ingestion pipeline converts unstructured literature into structured composition–property records. Every prediction is paired with dual-provenance neighbour views and screened by an inline LLM feasibility gate that checks classical metallurgical heuristics before inference. By coupling domain-aware visualisation with provenance-aware inference and constraint-driven design, MATAI lowers the engineering burden of applying AI to materials science and provides a reusable blueprint for closed-loop AI-for-Science systems.

1. Introduction

The design–test loop in structural alloy research is costly and slow: a single cast-and-test iteration can consume weeks and substantial material, and the composition space for even a modest n-ary system is combinatorially large. AI-for-materials-science (AI4MS) promises to shorten this loop by learning surrogate property models from historical data and by posing discovery as constrained optimisation over those surrogates (Raccuglia et al., 2016; Butler et al., 2018; Liu et al., 2017). Yet in practice the loop remains fragmented. Open databases such as Materials Project (Jain et al., 2013) and AFLOW (Curtarolo et al., 2012) provide curated data but no first-class prediction or inverse-design interface; benchmark suites such as Matbench (Dunn et al., 2020) standardise evaluation but are not workflow environments; Bayesian-optimisation toolkits such as COMBO (Ueno et al., 2016) and Dragonfly (Kandasamy et al., 2020) drive experiments but presume the data, features and surrogate are already in hand; and integrated commercial offerings are closed, limiting scientific customisation. Researchers therefore routinely hand-stitch Jupyter notebooks across these silos, repeating feature engineering and losing provenance at every hop.

MATAI addresses this gap with a single browser-based platform that closes the loop from data to candidate. The contributions are:

1. **A unified, web-based AI4MS workflow.** MATAI integrates visual analytics, property prediction, and constrained composition optimisation behind a shared data schema and REST backends, so that a hypothesis formed on a correlation heatmap can be validated by the predictor and deployed as an objective in the designer without leaving the browser.
2. **Domain-aware interactive analytics.** Eight coordinated views (element–element and element–property correlation, co-occurrence network, parallel coordinates, PCA/t-SNE feature maps with K-Means/DBSCAN, and provenance-based neighbour views) are engineered to surface physically meaningful patterns rather than generic plots, and to make every prediction auditable against its nearest data support.
3. **Physics-aware safeguards and natural-language control.** An LLM feasibility gate screens each composition against classical alloy-design heuristics before inference, and a conversational layer dispatches the same Predictor and Designer engines through structured natural-language commands.

[1]Nanyang Technological University, Singapore [2]Singapore Management University, Singapore. Correspondence to: Xinrun Wang <xrwang@smu.edu.sg>, Bijun Tang <bjtang@ntu.edu.sg>.

Accepted to the AI for Science workshop (ICML 2026).

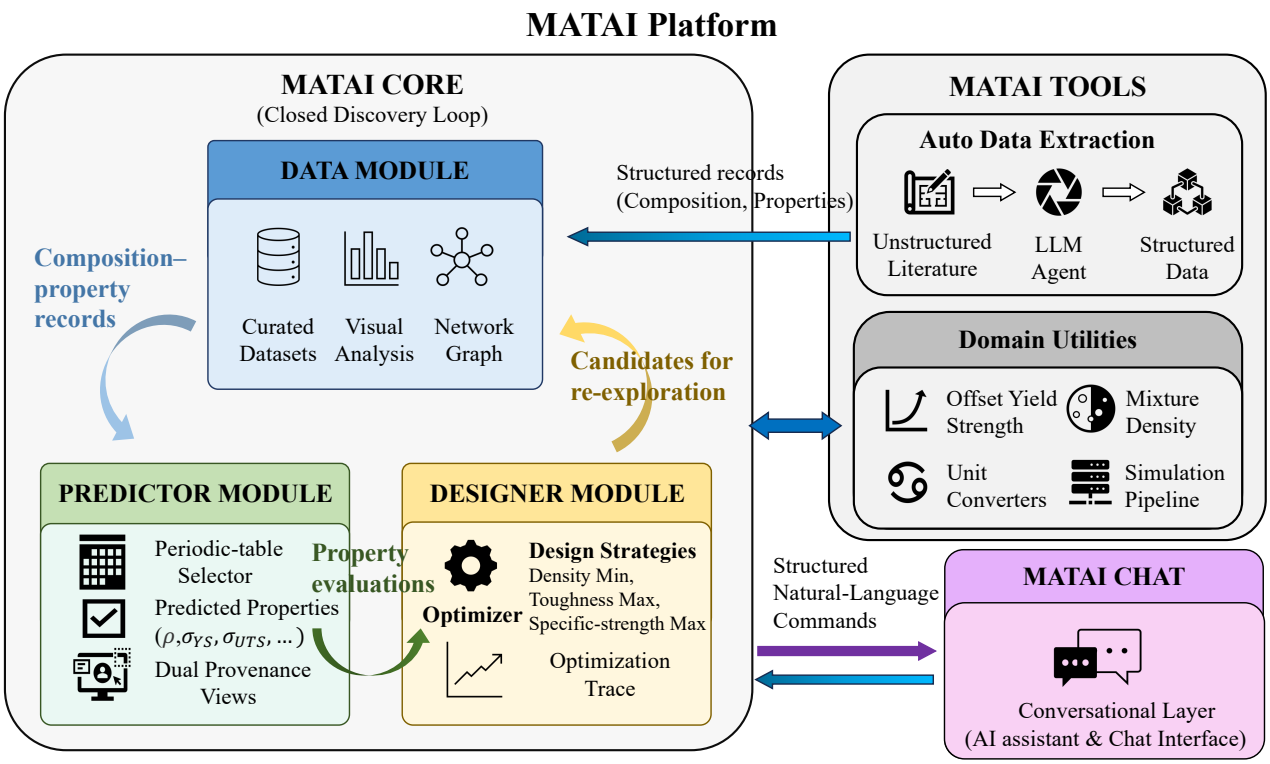

Figure 1. MATAI architecture. The Core modules (Data, Predictor, Designer) share a common composition–property schema and form a closed discovery loop; Tools and Chat extend the loop with literature ingestion and natural-language orchestration.

2. System Architecture and Modules

2.1. Architecture Overview

MATAI is organised as a two-tier architecture (Figure 1). The *MATAI Core* tier hosts three scientifically-motivated modules—Data, Predictor, and Designer—that together realise the discovery loop: Data supports hypothesis formation over a curated composition–property corpus; Predictor validates hypotheses by returning property forecasts together with provenance; Designer converts validated constraints into candidate compositions via constrained optimisation. A second *MATAI Tools* tier wraps experiment-facing utilities, a literature-ingestion pipeline, and a conversational orchestration layer that exposes the Core engines through structured natural-language commands.

The frontend is a single-page React application; the backend is a set of FastAPI microservices for prediction, design, feature analysis, document extraction, and chat, proxied behind a thin Node.js gateway. All modules read from and write to a unified composition–property record format, which is also the export format of the extraction pipeline, so data, predictions, and optimisation outputs are directly interoperable.

2.2. Data Module

The Data module is a visual-analytics surface over a curated composition–property corpus. Eight coordinated views are exposed, organised as a progression from global overview to deep-dive: dataset and element distributions; a scatterplot explorer with bubble encoding and brush-zoom selection; per-element content–property scatter with unit toggles; element– element and element–property correlation heatmaps; a co-occurrence network and linked parallel-coordinate plot; and a machine-learning feature-analysis view. Every view shares a common filtering layer (main-element-group checkboxes, atomic/weight-unit toggle, and optional property-range filters), so a hypothesis formed in one view can be carried verbatim into the next. Beyond the curated corpus, users can upload their own composition–property records that conform to the unified schema; uploaded datasets appear alongside the built-in sources in the dataset selector, so private or in-progress data can be explored, correlated, and fed into the Predictor and Designer through the same views without any code changes.

The **correlation heatmap** computes element–element or element–property association on three mutually exclusive representations (atomic fraction, weight fraction, or binary occurrence) under three metrics (Pearson r, Spearman ρ, and Kendall τ), with element-frequency and property-range

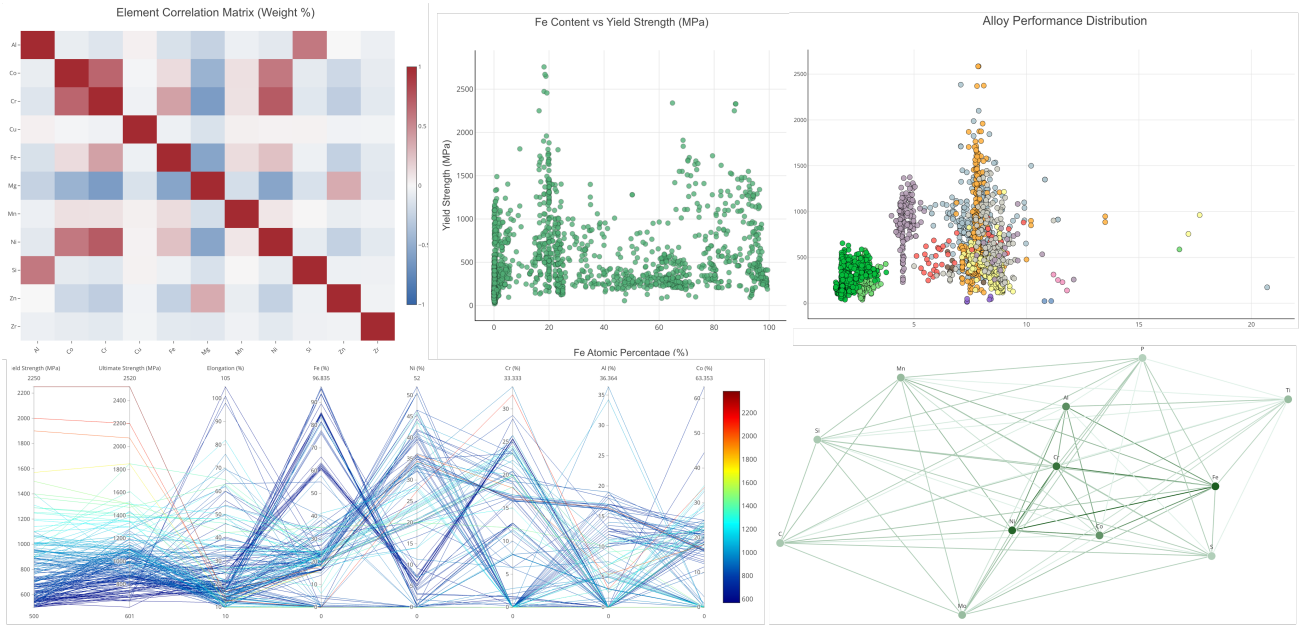

Figure 2. Representative views from the MATAI Data module. *Top-left*: element–element correlation heatmap (weight fractions), revealing co-selection patterns (e.g., Co–Cr–Ni) and mutual exclusions. *Top-middle*: per-element content–property scatter (Fe vs. yield strength). *Top-right*: alloy distribution in density–strength space, coloured by main-element group. *Bottom-left*: linked parallel-coordinate plot coupling mechanical properties to elemental fractions; polyline colour encodes average strength. *Bottom-right*: element co-occurrence network, with node size encoding frequency and edge weight encoding joint occurrence. All views share a common filtering layer so that a hypothesis formed in one view transfers directly to the next.

sliders that restrict the support set. This triple choice is deliberate: occurrence-based Pearson isolates *which elements are co-selected*, whereas atomic-fraction Spearman captures *how their loadings co-vary*—two physically distinct questions often conflated in standard EDA.

The **co-occurrence view** pairs a weighted network graph (nodes sized by frequency, edges weighted by joint occurrence) with a linked parallel-coordinate plot (Inselberg, 1985), in which the first block of axes carries mechanical properties and the second block carries elemental contents. Axis brushing filters samples across both views in real time, revealing bundles that correspond to known alloy families and isolating bridge elements that connect otherwise disjoint chemistries—natural candidates for hybrid compositions.

The **feature-analysis view** row-concatenates one-hot-encoded process and test conditions with continuous composition vectors and projects the resulting feature vector x_i to two dimensions via PCA (Abdi & Williams, 2010) or t-SNE (van der Maaten & Hinton, 2008), clustering the embedding with K-Means or DBSCAN (Ester et al., 1996). Toggling feature blocks on and off exposes whether a given cluster is composition-driven or process-driven—a standard but operationally tedious experiment that is reduced to a checkbox here.

2.3. Predictor Module

The Predictor is a composition-aware inference engine over an interactive periodic-table selector. Users pick 3–7 supported elements (57 in the current release) and enter atomic or weight fractions; a sum-to-100% constraint is enforced client-side. The backend returns core mechanical properties—density ρ, yield strength σ_{YS}, ultimate tensile strength σ_{UTS}, and elongation ε—as well as bulk and Young's moduli, Poisson's ratio, and top-ranked phase volumes. Out-of-coverage properties are explicitly flagged rather than silently extrapolated.

LLM-based feasibility gate. To guard against physically implausible queries before inference, the Predictor interposes a lightweight feasibility gate driven by a large language model. Each submitted composition is evaluated against a structured checklist rooted in classical alloy-design heuristics: the Hume-Rothery rules for solid-solution formation—atomic-size mismatch ($\delta > 15\%$ considered problematic), electronegativity difference, valence-electron concentration, and crystal-structure compatibility—as generalised to multi-principal-element systems by Zhang et al. (2008); the mixing-enthalpy (ΔH_{mix}) and atomic-size-difference criteria of Takeuchi & Inoue (2005), with ΔH_{mix} approximated from Miedema's semi-empirical model (Miedema et al., 1980); proximity to stoichiometries of known brittle intermetallics (e.g., Fe_2Al_5,

$TiAl_3$, Ni_3Sn); and boiling/melting-point conflicts that preclude homogeneous co-melting by conventional routes (arc melting, induction melting, powder metallurgy). Near-equiatomic multi-principal-element inputs are granted additional latitude because configurational entropy can stabilise single-phase solid solutions that otherwise violate binary rules (Yeh et al., 2004). The gate returns a structured verdict—*feasible*, *feasible with warning*, or *infeasible*—surfaced as an inline advisory before the prediction is issued, so users receive a concise diagnostic (e.g., a dominant immiscible binary pair such as Cu–Fe or Al–Pb, a candidate dominated by a brittle-phase stoichiometry, or a melting-point spread exceeding 1500 °C) instead of a silently extrapolated forecast.

A distinguishing feature is **dual provenance visualisation**. For every query, MATAI retrieves the 20 nearest recorded alloys under two complementary metrics. *Composition similarity* operates in log-ratio space via the centred log-ratio (CLR) transform and Aitchison distance (Aitchison, 1982), which is the natural geometry for simplex- valued compositional data and avoids the spurious correlations induced by Euclidean distance on fractions. Given compositions $x, y \in \Delta^{n-1}$ with geometric means $g(\cdot)$,

$$d_A(x, y) \;=\; \left\| \log \frac{x}{g(x)} - \log \frac{y}{g(y)} \right\|_2 . \tag{1}$$

Property similarity, by contrast, is the Euclidean distance in the z-standardised property space $(\rho, \sigma_{\mathrm{YS}}, \sigma_{\mathrm{UTS}}, \varepsilon)$, answering the orthogonal question "which known alloys *behave* like the prediction?". Presenting both lists next to the prediction gives the user an immediate, data-backed assessment of whether the model is interpolating or extrapolating.

2.4. Designer Module

The Designer casts alloy discovery as constrained black-box optimisation over the predictor. Let $x \in \Delta^{n-1}$ denote a candidate composition. A simulated-annealing (SA) search (Kirkpatrick et al., 1983) explores a neighbourhood defined by perturbing k elemental fractions subject to the simplex constraint $\sum_i x_i = 1$, and candidates are scored by calling the predictor. Three preset strategies cover common engineering regimes:

(i) Density minimisation.

$$\min_x \; \rho(x) \quad \text{s.t.} \;\; \sigma_{\mathrm{YS}}(x) \geq 300 \text{ MPa} \tag{2}$$

(ii) Toughness maximisation.

$$\max_x \; \varepsilon(x)\, \sigma_{\mathrm{UTS}}(x) \quad \text{s.t.} \;\; \rho(x) \leq 4.5 \text{ g cm}^{-3} \tag{3}$$

(iii) Specific-strength maximisation.

$$\max_x \; \frac{\sigma_{\mathrm{UTS}}(x)}{\rho(x)} \quad \text{s.t.} \;\; \varepsilon(x) \geq 5\%, \;\; \rho(x) \leq 4.5 \text{ g cm}^{-3} \tag{4}$$

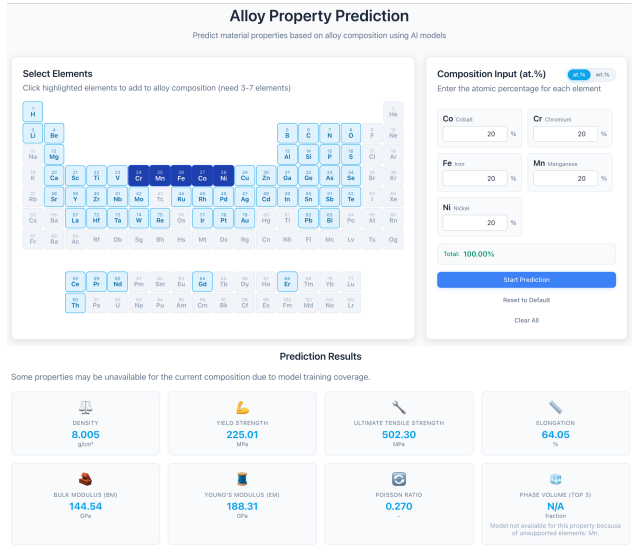

Figure 3. The Predictor module. Users select 3–7 elements via the interactive periodic table (non-supported elements greyed out), enter atomic or weight fractions under a sum-to-100% constraint, and receive predictions for eight mechanical properties. Out-of-coverage properties are explicitly flagged (e.g., Phase Volume).

The constraint editor exposes strategy-specific thresholds alongside shared advanced controls (maximum iterations N_{iter}, number of independent runs N_{runs}, and neighbourhood size k). Results are presented as a top-5 ranked table with predicted $\rho, \sigma_{\mathrm{YS}}, \sigma_{\mathrm{UTS}}, \varepsilon$, and the derived toughness and specific-strength indices, together with an optimisation-trajectory plot that enables the user to diagnose convergence and inspect solution diversity across runs.

2.5. Tools and Conversational Layer

A complementary tools layer streamlines everyday experimental chores and feeds the Core corpus. The *Auto Data Extraction* pipeline converts uploaded PDFs to structured composition–property records using a large- language-model extraction agent with per-user quota management; *Offset Yield Strength* fits the elastic segment of an uploaded stress–strain curve via a Web Worker to recover the 0.2% offset point; rule-of-mixtures *Mixture Density* and bidirectional *Mass/Atomic Converters* handle the unit-conversion overhead that typically blocks a non-specialist from feeding real lab data into a model. A *Simulation-in-the-Loop* workflow generates CALPHAD-verified synthetic records and is surfaced to users through a dedicated simulation-statistics view. Finally, *MATAI Chat* is a domain-tuned conversational agent that accepts structured commands (`PREDICT`, `DESIGN_MIN_DENSITY`, etc.) and dispatches them to the same Predictor and Designer backends, letting natural-language intent drive executable studies.

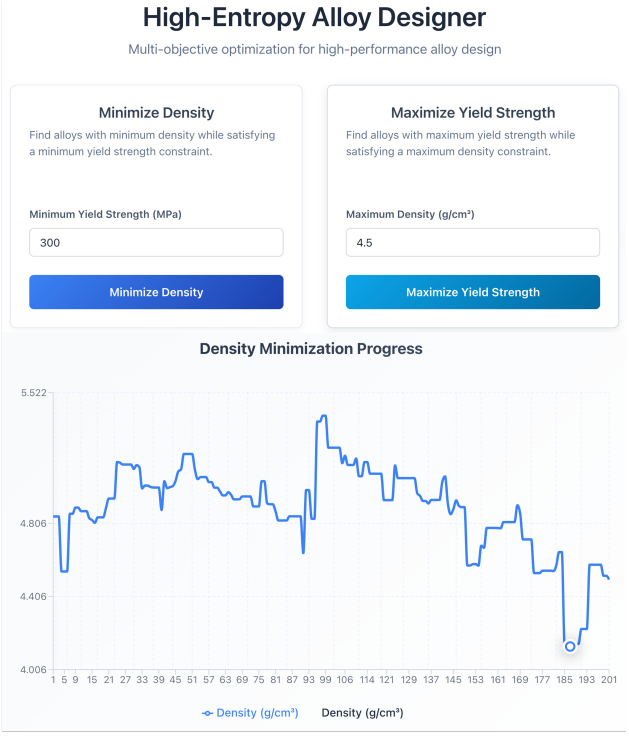

Figure 4. The Designer module. The strategy editor (top) maps user-specified bounds to the SA objective in Equations (2)–(4); here, density minimisation is run under a yield-strength floor of 300 MPa. The optimisation trajectory (bottom) records per-iteration density across 201 steps, visualising the search's coarse-to-fine annealing behaviour.

3. Related Work

Databases and benchmarks. Materials Project (Jain et al., 2013), AFLOW (Curtarolo et al., 2012), OQMD (Saal et al., 2013), JARVIS (Choudhary et al., 2020), and NOMAD (Draxl & Scheffler, 2019) provide large curated corpora, but their interfaces centre on record retrieval rather than interactive hypothesis formation or inverse design. Matbench (Dunn et al., 2020) and Foundry (Blaiszik et al., 2019) standardise datasets and model evaluation but do not expose a design loop. MATAI's Auto Data Extraction complements these corpora with literature-derived structured composition–property records that feed directly into the downstream prediction and design modules.

Property prediction. General-purpose composition- and structure-aware surrogates including CGCNN (Xie & Grossman, 2018), MEGNet (Chen et al., 2019), and Roost (Goodall & Lee, 2020) target property inference in isolation; MATAI complements such models by wrapping inference in a provenance-aware, designer-ready interface. Matminer (Ward et al., 2018) supplies featurisation primitives rather than an end-to-end environment.

Inverse design and Bayesian optimisation. Tools such as COMBO (Ueno et al., 2016), Dragonfly (Kandasamy et al., 2020), and BoTorch (Balandat et al., 2020) provide sequential optimisation libraries for experimentalists but require users to supply data pipelines, surrogates, and interfaces externally. Domain-specific alloy-design frameworks including the generalist constraint-aware predictor–designer of Deng et al. (2025) and the hierarchical LLM-plus-CALPHAD pipeline AutoMAT (Yang et al., 2025) demonstrate strong results on titanium-based systems; MATAI exposes analogous prediction, optimisation, and simulation capabilities through a unified interactive workflow, lowering the barrier for non-specialist users.

Integrated platforms. Citrination (O'Mara et al., 2016) offers an end-to-end informatics stack but is commercial and closed. To the best of our knowledge, MATAI is the first open, browser-based platform *for alloy discovery* that places domain-aware visual analytics, provenance-aware property prediction, and constrained compositional design behind a single, schema-unified interface.

4. Conclusion and Future Work

MATAI demonstrates that a tightly-integrated web platform can convert the traditionally scripted AI-for-alloys workflow into a closed interactive loop. By pairing domain-aware visual analytics with provenance-aware prediction and constraint-driven design, the platform lowers the engineering barrier to data-driven alloy research without sacrificing auditability. Planned extensions include multifunctional composites, 2D materials, and nanoparticle systems, along with tighter coupling to first-principles and CALPHAD simulation loops so that candidates flagged by the Designer can be verified *in silico* before experimental synthesis. The platform is publicly available at `https://matai-sg.org/`.

Acknowledgements

This work was supported by the National Research Foundation Singapore and DSO National Laboratories under the AI Singapore Programme (AISG Award No: AISG2-GC-2023-009-1B). This work was supported in part by the AI2050 initiative at Schmidt Sciences (Grant G-25-68035).

References

Abdi, H. and Williams, L. J. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4):433–459, 2010.

Aitchison, J. The statistical analysis of compositional data. *Journal of the Royal Statistical Society: Series B (Methodological)*, 44(2):139–160, 1982.

Balandat, M., Karrer, B., Jiang, D. R., Daulton, S., Letham, B., Wilson, A. G., and Bakshy, E. BoTorch: A framework for efficient Monte-Carlo Bayesian optimization. In *Advances in Neural Information Processing Systems*, volume 33, pp. 21524–21538, 2020.

Blaiszik, B., Ward, L., Schwarting, M., Gaff, J., Chard, R., Pike, D., Chard, K., and Foster, I. A data ecosystem to support machine learning in materials science. *MRS Communications*, 9(4):1125–1133, 2019.

Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O., and Walsh, A. Machine learning for molecular and materials science. *Nature*, 559(7715):547–555, 2018.

Chen, C., Ye, W., Zuo, Y., Zheng, C., and Ong, S. P. Graph networks as a universal machine learning framework for molecules and crystals. *Chemistry of Materials*, 31(9): 3564–3572, 2019.

Choudhary, K., Garrity, K. F., Reid, A. C. E., DeCost, B., Biacchi, A. J., Hight Walker, A. R., Trautt, Z., Hattrick-Simpers, J., Kusne, A. G., Centrone, A., et al. The joint automated repository for various integrated simulations (jarvis) for data-driven materials design. *npj Computational Materials*, 6(1):173, 2020.

Curtarolo, S., Setyawan, W., Hart, G. L. W., Jahnatek, M., Chepulskii, R. V., Taylor, R. H., Wang, S., Xue, J., Yang, K., Levy, O., et al. Aflow: An automatic framework for high-throughput materials discovery. *Computational Materials Science*, 58:218–226, 2012.

Deng, Y., Zhao, C., Li, Y., Tang, B., Wang, X., Zhang, Z., Lu, Y., Yang, P., Huang, J., Xiao, Y., Guan, C., Liu, Z., and An, B. MATAI: A generalist machine learning framework for property prediction and inverse design of advanced alloys. *arXiv preprint arXiv:2511.10108*, 2025.

Draxl, C. and Scheffler, M. The nomad laboratory: from data sharing to artificial intelligence. *Journal of Physics: Materials*, 2(3):036001, 2019.

Dunn, A., Wang, Q., Ganose, A., Dopp, D., and Jain, A. Benchmarking materials property prediction methods: the matbench test set and automatminer reference algorithm. *npj Computational Materials*, 6(1):138, 2020.

Ester, M., Kriegel, H.-P., Sander, J., Xu, X., et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *KDD*, volume 96, pp. 226–231, 1996.

Goodall, R. E. A. and Lee, A. A. Predicting materials properties without crystal structure: deep representation learning from stoichiometry. *Nature Communications*, 11(1):6280, 2020.

Inselberg, A. The plane with parallel coordinates. *The Visual Computer*, 1(2):69–91, 1985.

Jain, A., Ong, S. P., Hautier, G., Chen, W., Richards, W. D., Dacek, S., Cholia, S., Gunter, D., Skinner, D., Ceder, G., and Persson, K. A. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1(1):011002, 2013.

Kandasamy, K., Vysyaraju, K. R., Neiswanger, W., Paria, B., Collins, C. R., Schneider, J., Poczos, B., and Xing, E. P. Tuning hyperparameters without grad students: Scalable and robust bayesian optimisation with dragonfly. *Journal of Machine Learning Research*, 21(81):1–27, 2020.

Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

Liu, Y., Zhao, T., Ju, W., and Shi, S. Materials discovery and design using machine learning. *Journal of Materiomics*, 3(3):159–177, 2017.

Miedema, A. R., de Châtel, P. F., and de Boer, F. R. Cohesion in alloys—fundamentals of a semi-empirical model. *Physica B+C*, 100(1):1–28, 1980.

O'Mara, J., Meredig, B., and Michel, K. Materials data infrastructure: a case study of the citrination platform to examine data import, storage, and access. In *JOM*, volume 68, pp. 2031–2034, 2016.

Raccuglia, P., Elbert, K. C., Adler, P. D. F., Falk, C., Wenny, M. B., Mollo, A., Zeller, M., Friedler, S. A., Schrier, J., and Norquist, A. J. Machine-learning-assisted materials discovery using failed experiments. *Nature*, 533(7601): 73–76, 2016.

Saal, J. E., Kirklin, S., Aykol, M., Meredig, B., and Wolverton, C. Materials design and discovery with high-throughput density functional theory: the open quantum materials database (oqmd). *JOM*, 65(11):1501–1509, 2013.

Takeuchi, A. and Inoue, A. Classification of bulk metallic glasses by atomic size difference, heat of mixing and period of constituent elements and its application to characterization of the main alloying element. *Materials Transactions*, 46(12):2817–2829, 2005.

Ueno, T., Rhone, T. D., Hou, Z., Mizoguchi, T., and Tsuda, K. Combo: An efficient bayesian optimization library for materials science. *Materials Discovery*, 4:18–21, 2016.

van der Maaten, L. and Hinton, G. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9:2579–2605, 2008.

Ward, L., Dunn, A., Faghaninia, A., Zimmermann, N. E. R., Bajaj, S., Wang, Q., Montoya, J., Chen, J., Bystrom, K., Dylla, M., et al. Matminer: An open source toolkit for materials data mining. *Computational Materials Science*, 152:60–69, 2018.

Xie, T. and Grossman, J. C. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Physical Review Letters*, 120 (14):145301, 2018.

Yang, P. et al. Autonomous multi-objective alloy design through simulation-guided optimization. *arXiv preprint arXiv:2507.16005*, 2025.

Yeh, J.-W., Chen, S.-K., Lin, S.-J., Gan, J.-Y., Chin, T.-S., Shun, T.-T., Tsau, C.-H., and Chang, S.-Y. Nanostructured high-entropy alloys with multiple principal elements: novel alloy design concepts and outcomes. *Advanced Engineering Materials*, 6(5):299–303, 2004.

Zhang, Y., Zhou, Y. J., Lin, J. P., Chen, G. L., and Liaw, P. K. Solid-solution phase formation rules for multi-component alloys. *Advanced Engineering Materials*, 10 (6):534–538, 2008.

A. Case Study: Low-Density, High-Strength Ti-Based Alloys

We illustrate the end-to-end workflow on a canonical design target: a titanium-based alloy with low density, high specific strength, and non-trivial ductility. The full session below was executed inside a single browser tab.

Step 1 – Data exploration. Using the element–property correlation heatmap with Pearson r on atomic fractions, Al shows the expected negative correlation with ρ, while V and Fe correlate positively with σ_{UTS}. Switching the same view to binary occurrence confirms that the effect is not an artefact of a few heavily-loaded records. The co-occurrence network (with element-frequency slider constrained to commonly- used elements and the property filter set to $\sigma_{\mathrm{UTS}} > 800\,\mathrm{MPa}$) isolates Ti–Al–V–Fe as a densely connected sub-network, supporting the hypothesis that this quaternary system is a promising starting point.

Step 2 – Property prediction. A candidate composition of Ti 65–Al 15–Fe 15–V 5 (at.%) is entered via the periodic-table selector. The Predictor returns core mechanical properties and, crucially, two provenance panels: the composition-similarity neighbours (Aitchison distance, Equation (1)) locate the prediction within the recorded Ti–Al–V–Fe design space, while the property-similarity neighbours show experimentally-measured alloys with comparable $(\rho, \sigma_{\mathrm{YS}}, \sigma_{\mathrm{UTS}}, \varepsilon)$ tuples—a consistency check that is unavailable in predictor-only tools.

Step 3 – Constrained design. The same intent is then posed as a specific-strength maximisation (strategy (iii)) with $\varepsilon \geq 5\%$ and $\rho \leq 4.5\,\mathrm{g\,cm^{-3}}$. Across three independent SA runs ($N_{\mathrm{iter}} = 200$, $k = 2$), the optimiser converges in well under 200 iterations per run and returns a diverse top-5 of Ti-rich compositions with Fe and V as the dominant strengtheners—matching the qualitative hypothesis formed in Step 1 while sharpening it into concrete candidates.

B. Data Module: Extended Analytics

The Data module exposes eight coordinated tabs beyond the summary given in the main text: *Overview*, *Scatter Plot*, *Element Analysis*, *Element–Element Correlation*, *Element–Property Correlation*, *Element Co-occurrence*, *Feature Analysis*, and *Simulation Data*. All tabs except *Simulation Data* share a global dataset selector—checkboxes at the top of every page allow users to include or exclude entire source datasets, so that subsequent visualisations remain consistent across the exploration workflow. *Simulation Data* is treated separately because its records originate from the automated simulation workflow rather than experimental measurement.

The *Overview* tab pairs interactive pie charts (data-source distribution and elemental breakdown) with side-by-side histogram/KDE explorers for both element content and material properties. In *Element Mode*, a histogram and Kernel Density Estimate curve visualise the distribution of atomic- or weight-fraction loadings for a chosen element, automatically excluding samples that do not contain it to avoid skewed summaries. In *Property Mode*, the same overlay is applied to four core properties—σ_{YS}, σ_{UTS}, ε, and ρ—to reveal typical performance ranges and variability.

The *Scatter Plot* tab maps each alloy to three dimensions (x-axis, y-axis, bubble size), with points colour-coded by the predominant element to separate single-principal-element alloys from high-entropy alloys at a glance. Brush-zoom selection and persistent filters carry the selected subset into subsequent tabs.

C. Auto Data Pipeline: LLM-Driven Literature Extraction

The Auto Data Pipeline is the corpus-expansion engine exposed through the *Tools* menu. Users upload a research PDF (up to 50 MB per file); an LLM-based extraction agent then (i) generates a structured paper summary covering research background, experimental methods, key findings, and application prospects; (ii) extracts alloy compositions, processing conditions, mechanical properties, and phase-structure descriptors into a tabular record; and (iii) standardises the result into the unified composition–property schema used by the rest of the platform.

Real-time progress is reported through the extraction stages, and any field that cannot be reliably extracted from the source is explicitly marked as "Not extracted" rather than filled with a guess, preserving data transparency. Finalised records can be downloaded as a CSV for statistical analysis or merged into the curated corpus for downstream prediction and design. The pipeline handles complex tables and figure-embedded data and supports per-user quota management to prevent abuse when the underlying LLM is metered.

D. Simulation-in-the-Loop: CALPHAD-Guided Design

The Simulation-in-the-Loop module extends the Designer with a three-layer workflow that interleaves LLM-based ideation, CALPHAD thermodynamic verification, and experimental validation.

Ideation layer. A user-supplied target (e.g., "low density, $\sigma_{\mathrm{YS}} \geq 800$ MPa, avoid high-cost elements, ≥ 4 constituents") is translated into a machine-readable prompt. An LLM scans literature and handbooks to identify alloy families consistent with the target, excludes cost-prohibitive elements, enforces the multi-principal constraint, and nominates a well-documented seed composition as the starting point.

Simulation layer. The seed is refined by AI-guided CALPHAD search in mole-fraction space using a *coarse-to-fine* neighbourhood schedule: a coarse phase with step size $\Delta = 0.5$ and range ± 10 mol% rapidly lowers the primary objective; a fine phase with $\Delta = 0.2$ and range ± 2 mol% performs local convergence. The multi-objective trade-off is collapsed to a scalar proxy

$$S = \frac{\sigma_{\mathrm{YS}}}{\rho} e^{-\rho},$$

which penalises heavier candidates while preserving strength. Each iteration re-evaluates ρ and σ_{YS} via CALPHAD and updates the best composition.

Experimental layer. The final design is cast in the as-cast condition and tensile-tested; XRD, EBSD and TEM confirm phase constitution and the nanoscale features predicted by CALPHAD.

Representative run. Starting from a `Ti-185`-like seed ($\rho \approx 4.70\,\mathrm{g\,cm^{-3}}$), the loop produced a refined $Ti_{81.4}Al_{16.8}V_{1.6}Fe_{0.2}$ candidate with $\rho = 4.36\,\mathrm{g\,cm^{-3}}$, $\sigma_{\mathrm{YS}} = 927$ MPa, and $S = 11.91$. The as-cast ingot measured $\rho = 4.32\,\mathrm{g\,cm^{-3}}$ and $\sigma_{\mathrm{YS}} = 829$ MPa—roughly 8% lighter than the seed with negligible loss in yield strength, and microstructure consistent with an α/β dual-phase structure carrying nano-scale β precipitates. The full loop completed in under seven days at a sustained rate of > 1,000 candidate evaluations per day, compared with the ~2 years a manual CALPHAD study at ~100 evaluations per day would have required for a comparable survey.

E. MATAI Chat: Command Grammar

MATAI Chat is a domain-tuned conversational agent that couples free-form natural-language discussion with a set of structured commands dispatched to the Predictor and Designer backends. Free-form interaction supports materials Q&A, interpretation of uploaded phase diagrams and micrographs, and guided walk-throughs of the platform's modules. Structured commands expose the same engines that the interactive modules use, so a chat session can produce the same predictions and candidate lists as a manual workflow.

The command grammar is intentionally small:

```
PREDICT: {Element1: Percentage1, Element2: Percentage2, ...}
         e.g., PREDICT: {Al: 20.0, Ti: 60.0, Cr: 15.0, Ni: 5.0}

DESIGN_MIN_DENSITY:  min_elongation=X, min_specific_strength=Y
DESIGN_MAX_ELONGATION_UTS:  max_density=X
DESIGN_MAX_SPECIFIC_STRENGTH:  min_elongation=X, max_density=Y
```

Executing a `PREDICT` command returns the same property vector and dual-provenance neighbours that the Predictor module produces; executing any `DESIGN_*` command runs the corresponding constrained simulated-annealing search and returns the top-ranked candidates. The chat layer therefore lets natural-language intent drive the same executable studies as the GUI, with no separate surface for model outputs or constraint editing.

F. Development Roadmap

The platform's planned rollout extends the Core modules beyond alloys while preserving the unified composition–property schema and closed-loop discovery pattern:

- **2H 2025.** Core platform launch covering high-entropy alloys, with the interactive Predictor and Designer described in the main text.
- **1H 2026.** Expansion to multifunctional composites, with extended schema and analytics tuned for composition–structure–function relationships and multi-phase interactions.
- **2H 2026.** Addition of two-dimensional materials, including automated structure generation, property forecasting, and synthesis-parameter recommendation.
- **1H 2027.** Dedicated nanoparticle module covering phase-controlled synthesis pathways and catalytic-performance prediction.
- **2H 2027 and beyond.** Continuous extension to ceramics, polymers, quantum dots, and further emerging classes to cover a broader materials-discovery surface.