

Multi-source integration of geoscientific and economic-mineral knowledge for ranking in mineral exploration

P. X. P. Martin-Cocher

Infomineral

This work proposes the systematic integration of multi-source geoscientific and economic-mineral knowledge assisted by AI, an attempt to make geology more continuous and less discrete in its boundaries.

The framework ingests multiple heterogeneous sources as weighted layers of a single inference system, without manual pre-processing, forced nomenclature standardisation, or sequential steps between domains. Geological sources include: maps and explanatory notes from geological surveys; scientific articles and theses; more than 15,000 technical mineral exploration reports (NI 43-101, JORC, SK-1300) processed by LLM-based pipelines used as data extraction tools; and petrogenetic compilations. Using epistemological weighting mechanism, the system incorporates economic-mineral context: mining legislation and blocked areas; environmental protection zones; political situation and resource sovereignty; commodity price history and projections; profile of active investors in the region; and economic viability parameters (grade cut-off, minimum deposit size, investigation CAPEX). Each source receives weight proportional to its scale, purpose, and evidential quality. The result is a continuous probability field over lithological subtypes and mineral context (suggested name hypergeology), replacing discrete polygons. In practical terms, the unit of evaluation becomes the polygon or mineral title, organised through a ranking system.

The framework presents three key differentials. First, adaptive lithological hierarchy: in data-dense zones (e.g., Borborema Province, NE Brazil), the model descends the arborescent structure, "granite" expands into two-mica granite, rapakivi granite, and associations such as pegmatite and aplite. In data-sparse zones (e.g., Nigerian mobile belts, with analogous geology but far more limited documentation), the model ascends one or two levels (operating on "granitoid") without imposing unsupported subdivisions. Granularity is determined by local observation density, allowing coherent operation across regions with asymmetric data coverage. Second, practical falsifiability: when an area is flagged as favourable, Pit Detect is executed, a routine developed by the author, trained on satellite images of pits derived from OpenStreetMap polygons, capable of identifying artisanal, semi-artisanal, and industrial mining activity. Artisanal miners are historically efficient and low-cost prospectors; their presence is empirical evidence of attractiveness, and their absence in a predicted favourable area is a significant (though not definitive) signal of model failure. Third, investment-decision output: the result is not "where something might exist", but where it is worth investigating, for which investor profile, with which minimum budget, and (critically) what to do to try to break the model, not to confirm it.

The final output includes: a comparative ranking of targets within a budget window; the appropriate investor profile (early-stage explorer, major, State); regulatory and political conditions required for acceptable risk; and the minimum field dataset needed to validate or refute the model's predictions. For geological surveys, the framework enriches legacy vector archives with uncertainty layers and economic context without remapping; for explorers and investors, it concentrates effort where geological probability, economic viability, and the regulatory environment converge.

The model is in a development, evaluation, and testing phase. Results for specific commodities and areas are underway, but the framework does not yet constitute a global or multi-commodity model. This work describes the conceptual architecture, methodological principles, and available preliminary results, with the aim of opening the framework to community discussion and contribution.