Deep-Time Data Infrastructure: Visualizing the Co-evolution of Planets and Life. Michael B. Meyer[1], Robert M. Hazen[1], Ahmed Eleish[2], Daniel R. Hummer[3], Chao Liu[1], Shaunna M. Morrison[1], and the Keck Deep-Time Data Collaboration[1], [1]Geophysical Laboratory, Carnegie Institution for Science, 5251 Broad Branch Road NW, Washington DC 20015; [2]Rensselaer Polytechnic Institute, Troy NY 12180; [3]Department of Geology, Southern Illinois University, Carbondale IL 62901.
Email: mmeyer@carnegiescience.edu or rhazen@carnegiescience.edu

Earth's story is a 4-billion-year epic of transformations driven by physical, chemical, and biological processes. The planet's living and non-living components have co-evolved yet our ability to document, model, and explore these changes is hampered by a lack of data integration among disciplines. The Deep-Time Data Infrastructure (DTDI) is a multi-institution project that seeks to trace the planet's co-evolution by curating and integrating data from diverse fields including mineralogy, paleobiology, tectonics, geochronology, proteomics, and geochemistry. The DTDI's central objective is to document and understand Earth's evolution over 4 billion years.

Comprehensive databases of mineral species (rruff.info/ima) and their geographic localities and co-existing mineral assemblages (mindat.org) reveal patterns of mineral diversification and association through space and time (Fig. 1) and patterns of mineral coexistance pairs appear in chord diagrams (Fig. 2).

Distributions of fossil species (Fig. 3) and minerals (Fig. 4) can be represented by network diagrams. Fossil or mineral species are represented as "nodes" and co-existing species are "links." Filters based on chemistry, age, structural group, and other parameters highlight visually both familiar and new aspects of paleobiology and mineralogy. We quantify networks with statistical metrics, including connectivity (based on the frequency of species co-occurrence), homophily (the extent to which co-existing species share key characteristics), network closure (based on the degree of network interconnectivity), and segmentation (as revealed by isolated "cliques" of species).

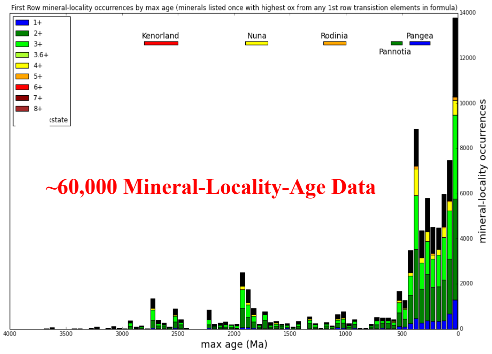

Fig. 1. The temporal distributions of minerals incorporating first-row transition elements reveals episodic mineralization and systematic changes in redox states.

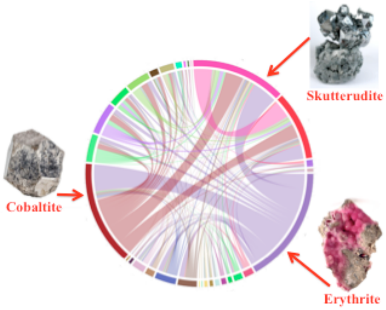

Fig. 2. A chord diagram of cobalt minerals shows coexisting pairs of species. Prominent ties occur between the alteration mineral erythrite and the major ores cobaltite and skutterudite.

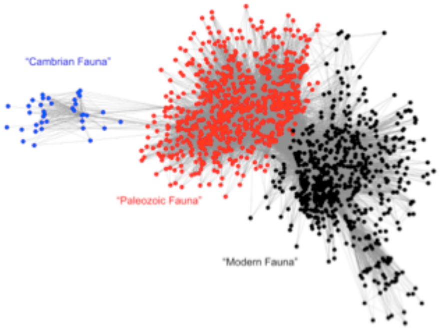

Fig. 3. A network diagram of fossil groups reveals significant changes in fauna through the Phanerozoic, sweeping from left to right. Colors represent fauna from major time intervals: blue = Cambrian; red = later Paleozoic; black = Mesozoic. Major extinction events appear as constrictions in the network.

Fig. 4. A network diagram of common chromium minerals illustrates mineral abundance (corresponding to node size), mode of formation (corresponding to node color: orange = igneous; blue = metamorphic; green = weathering; tan = soil), and patterns of co-occurrence.