

## About “Correlation Technology”

The Correlation Technology Platform implements a unique, new computational model for knowledge discovery and exploration. This model emulates human methods for decomposition, storage, and utilization of data. When used appropriately, this model can deliver practical solutions to a number of complex and previously intractable challenges faced by government, corporate, and cultural entities in the era of “big data”.

The “Correlation Technology” approach starts by recognizing that all human knowledge embedded in text is expressed in the form of “Knowledge Fragments” - each of which encapsulates an “intrinsic relation” and its bound “relata”. There is a substantial cadre of intrinsic relations which are innately comprehended by humans. For example, the fragment, “engine in car”, expresses a “containment” relation. The fragment, “fingers on hand”, expresses an “attachment” relation. The fragment, “portrait of President Obama”, expresses a “representation” relation. Other types of intrinsic relations include extensional relations and intentional relations, class relations, mereological relations, topological relations, existential relations, action relations, transitional relations, causal relations, dependency relations, semiotic relations, mediated relations, conventional relations, and property-based relations. Each of these types and sub-types of relations is expressed using one or more specific words in every natural language. The containment relation expressed by “engine in car” is a valid concept in any language on earth. The natural language terms expressing each and every such relation is then mapped. Unlike any previous technology, Correlation Technology decomposes input resources such as web pages, emails, or text documents by extracting all such relations and their relata in an exhaustive one-way transform.

The resulting Knowledge Fragments are stored in a data store generically referred to as an “infobase”. Each of these Knowledge Fragments is like a “note” a student might take when listening to a lecture. Correlation Technology further acknowledges that successful human reasoning requires that the human mind store such “notes” without any constraint upon subsequent recollection and use. For this reason, the infobase is schema-free, and every Knowledge Fragment is “equally eligible” for use in process. This additional aspect of Correlation Technology is not present in any prior Knowledge Discovery technique.

Humans utilize two methods for reasoning – “Connect the Dots” and “Free Association”. For both these methods, humans iteratively link “note” to “note”, starting with some “origin” note, and build a chain of notes that lead to some “destination” note which serves as conclusion. In response to any stimulus, including an explicit question, humans will construct hundreds or thousands of such chains – all of which are related, and with each note “link” of each chain representing a qualitative, logical “correlation” to every prior “link”. In response to any “N-dimensional query”, the penultimate step in the Correlation Technology process is to build an “answer space” or result set of such qualitative, logical “correlations” for the contents of the infobase. Hundreds or thousands of such correlations are typically constructed from even a relatively small infobase.

The final step for utilizing Correlation Technology is to apply analytics to the “answer space”. Because the correlations are qualitative assertions, the analytical approach may include statistical methods, logical methods, rule-based algorithms, and natural language techniques. For each type of use, the types of analysis required will differ, but the same goal always applies – find actionable insights from overwhelmingly complex data; find important relationships between terms, phrases, concepts or entities where those relationships can not be discovered through entity-linking, navigational, or subjective statistical means.