CLAIMS

What is claimed is:

1. A data analysis system, comprising:

a processor and a memory;

the processor is configured to ingest content from a plurality of data sources with the content comprising a plurality of observables relevant to an entity, wherein each observable from the plurality of observables is related to at least one feature of interest from a plurality of features of interest;

the processor is further configured to tag the observables with an entity identifier in response to the observables referencing at least one of an entity, a tradename associated with the entity, or product associated with the entity;

the processor is further configured to tag an observable from the plurality of observables in response to the feature of interest being related to the observable, wherein an analysis engine measures a materiality of the feature of interest to the entity by counting a number of observables from the plurality of observables tagged with the entity identifier;

the processor is further configured to generate measurability mechanisms based on the tags to create a materiality signature for the entity for a time period during which the plurality of observables were received;

the processor is further configured to create dynamic materiality using the materiality signature for the entity as a function of time; and

the processor is further configured to measure a materiality of a particular event using the dynamic materiality;

wherein the analysis engine is configured to:

cluster the plurality of observables using a clustering technique;

assign a plurality of signatures to the clusters; and

affinitize the signatures using a metric technique.

2. The data analysis system of claim 1, comprising a graphical user interface that is configured to display the materiality of the feature of interest.

3. The data analysis system of claim 1, comprising a graphical user interface that is configured to allow a user to select the feature of interest which causes a process to be performed by the extraction engine or the analysis engine.

4. The data analysis system of claim 1, wherein the clustering technique is an agglomerative clustering, a Euclidean clustering, or a principal component analysis.

5. The data analysis system of claim 1, wherein the metric technique is Levenshtein Distance agglomerative clustering.