

## TYPES OF DATA COLLECTED

### Demographic Criteria

Both the qualitative and quantitative instruments deployed by IMPACTS collect demographic data for the primary purpose of segmenting (“cross-tabulating”) responses by demographic cohorts.

For example, The Ocean Project is able to segment perceptual findings by age, ethnicity, education, gender and household income (among many) cohorts.

The IMPACTS certification process ensures the integrity of the collected data by verifying that respondents to the survey instruments possess the actual demographic characteristics that they claim.

### Scalar Variables

A scalar variable is a proposition that seeks to quantify the relativity of a response within a continuum. These type of propositions inform degrees of agreement with certain statements (e.g., “On a scale of 1-10 with “1” being completely disagree and “10” being completely agree...)

Additionally, scalar variables inform the development of Structural Equation Models. Structural equation modeling is a means of assessing latent constructs – factors that are not observable and, therefore, less readily measured.

In general, the scales utilized within the research processes proposed by IMPACTS seek to *maximize* the variability of responses; thus, the scale deployed within the proposed

qualitative and quantitative survey instruments is “0-100” (as opposed to the more traditional “1-5” or “1-10”).

### Open-End Queries

Open-end queries encourage the sample population to respond in its own words and enable lexical analysis – the results of which underpin many of the network models utilized by IMPACTS.

Perhaps the best means of explaining the lexical analysis process is to provide an example from a precedent implementation. The following queries appeared on an instrument used to inform an organization’s understanding of its visitors’ interpretation of “inspiration”:

Query A. Describe a setting or location where you were inspired?

Query B. What about [INSERT LOCATION] was inspiring to you?

The responding sample population (totaling upwards of 4,000 persons) identified several locations and characteristics as being particularly “inspiring.” Among the most “inspiring” locations were the Parthenon, Great Pyramids, Golden Gate Bridge, National Cathedral, Notre Dame and Omaha Beach. The output of the lexical analysis suggested that a component of the inspiration experienced by visitors to the stated locations was attributable to symmetry.

Thus, after respondents specifically named the Parthenon (or Notre Dame or the Pyramids, etc.) as an inspiring location, they were consequently asked *what* about the cited structure was inspiring.

Respondents generally commented on the beauty and scale of the named locations. A follow-on question delved deeper into what visual components constituted “beauty” – “symmetry” emerged as a key component. Thus, the lexical analysis did reveal “symmetry” in the responses (most frequently cited relating to Notre Dame, Golden Gate Bridge, National Cathedral and Omaha Beach. The output of the lexical analysis was a basic tabular categorization that ranked responses based on frequency.

The analysis contemplated lexical pointers (relationships between words in “synsets” - SYNOnym SETS – sets of words that are generally interchangeable in the prevailing context) and classifies the synsets as holonyms and/or hypernyms. (Holonyms are the whole to which a constituent belongs – “A” is a holonym of “B” if “B” is a part of “A.” Hypernyms are a general classification of words – “A” is a hypernym of “B” if “B” is an example of “A.”)

Specific nouns such as “Parthenon” that do not have obvious synsets are categorized by the synsets of their respective definitions. For example, in the examples of “Parthenon” and “Golden Gate Bridge” (copied verbatim from the lexical analysis):

(n) **Parthenon** (the main temple of the goddess Athena; built on the acropolis in Athens earlier than 400 B.C.; example of Doric architecture)

holonym (n) Athens, Athinai, capital of Greece, Greek capital (the capital and largest city of Greece; named after Athena (its patron goddess))

holonym (n) Greece, Hellenic Republic, Ellas (a republic in southeastern Europe on the southern part of the Balkan peninsula)

hypernym (n) temple (place of worship consisting of an edifice for the worship of a deity)

(n) place of worship, house of prayer, house of God, house of worship (any building where congregations gather for prayer)

(n) building, edifice (a structure that has a roof and walls and stands more or less permanently in one place)

(n) structure, construction (a thing constructed; a complex entity constructed of many parts)

(n) artifact, artefact (a man-made object taken as a whole)

(n) whole, unit (an assemblage of parts that is regarded as a single entity)

(n) object, physical object (a tangible and visible entity; an entity that can cast a shadow)

(n) physical entity (an entity that has physical existence)

(n) entity (that is perceived or known or inferred to have its own distinct existence (living or nonliving))

(n) **Golden Gate Bridge** (a suspension bridge across the Golden Gate)

holonym (n) San Francisco (a port in western California near the Golden Gate that is one of the major industrial and transportation centers; it has one of the world’s finest harbors; site of the Golden Gate Bridge)

holonym (n) California, Golden State, CA, Calif. (a state in the western United States on the Pacific; the 3rd largest state; known for earthquakes)

holonym (n) United States, United States of America, America, the States, US, U.S., USA, U.S.A. (North American republic containing 50 states - 48 conterminous states in North

America plus Alaska in northwest North America and the Hawaiian Islands in the Pacific Ocean; achieved independence in 1776)

hypernym (n) bridge, span (a structure that allows people or vehicles to cross an obstacle such as a river or canal or railway etc.)

(n) structure, construction (a thing constructed; a complex entity constructed of many parts)

(n) artifact, artefact (a man-made object taken as a whole)

(n) whole, unit (an assemblage of parts that is regarded as a single entity)

(n) object, physical object (a tangible and visible entity; an entity that can cast a shadow)

(n) physical entity (an entity that has physical existence)

(n) entity (that is perceived or known or inferred to have its own distinct existence (living or nonliving))

The highest common denominator of the hypernyms (in the above example, the nouns “structure” and “construction”) form the categorization. Thus, the lexical analysis would recognize that both the Golden Gate Bridge and the Parthenon are structures.

A similar process applies to the balance of the user-inputted text. Respondents who associated “Golden Gate Bridge” with the descriptor “symmetry” (or “symmetrical”) would trigger the following analysis:

(n) **symmetry**, symmetricalness, correspondence, balance ((mathematics) an attribute of a shape or relation; exact reflection of form on opposite sides of a dividing line or plane)

hypernym (n) spatial property, spatiality (any property relating to or occupying space)

(n) property (a basic or essential attribute shared by all members of a class)

(n) attribute (an abstraction belonging to or characteristic of an entity)

(n) abstraction (a general concept formed by extracting common features from specific examples)

(n) abstract entity (an entity that exists only abstractly)

(n) entity (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))

Each of the associated descriptors associated with symmetry (i.e., “balance”) have their own hypernym sets. After identifying the multiple levels of hypernyms, a network model “maps” each association and commonality (thereby identifying that respondents who cited “Notre Dame” also cited “symmetry” as a contributing component of inspiration – as did respondents who cited “Golden Gate Bridge.”)

The intent of this process is to gain a greater depth of understanding than the normally stated responses to a survey would otherwise provide about an abstraction such as “inspiration.” By identifying the non-explicit contributors to large volumes of seemingly disparate data points, network models reveal connections and patterns that would otherwise remain hidden and/or isolated.

### **Unstructured Data**

Technology has dramatically altered the means by which the public gains awareness of issues, conducts research, seeks recommendations and offers its feedback.

No organization is immune to the revolutionary impacts of technology. The wide adoption of digital technologies has fundamentally altered an organization's interface with its market. Content-sharing websites such as *flickr.com* and *youtube.com* provide a means of documenting and promulgating experiences. Online communities ("socialhoods") such as *tripadvisor.com* provide forums for affinity communication and testimonials. Blogs offer the validation of publication and the geographic reach of a global audience.

All of the above-described digital technologies represent examples of unstructured data. Unstructured data differs from structured data due to the relative challenge to easily collect, categorize and process it in a consistent, meaningful manner. Recent studies estimate that unstructured data accounts for more than 80% of the world's information supply. As such, an organization's sole reliance on structured data impedes its ability to develop a fully-informed, holistic understanding of its audiences.

When combined with the delivery of predicted intelligence, the analysis of unstructured data enables organizations to approach total information awareness in near real-time.