Efficient Integration of Data

H. Roice Nelson, Jr. & Steven C. Swann
InterLeeV LLC
Cedar City, Utah 84720

Executive Summary

While there are many systems, formulas, indexes, and algorithms to list compile, and sort data, it takes time, effort, and computational work to ensure the right data is presented correctly to accurately determine the results. InterLeeV's System does this faster, easier, and more accurately than anything else available today.

This white paper shows how disparate data from multiple sources can be difficult to glean or interpret accurately. With InterLeev, data can be quicky and accurately indexed and displayed in a more critical and useful way, without a lot of complicated processes or procedures.

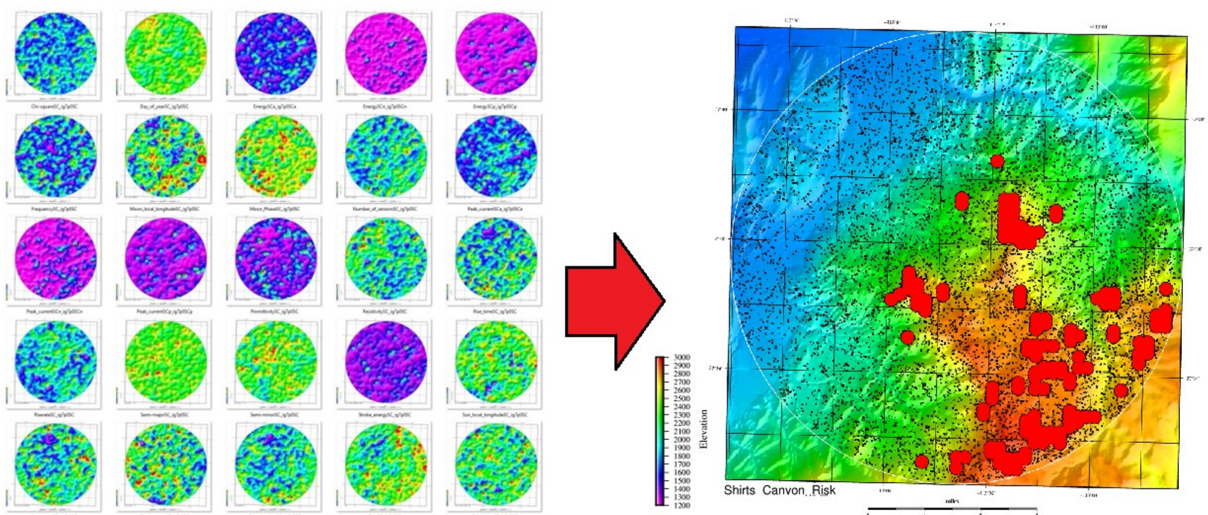

Figure 1. On the left, multiple dimensions of Raw Lightning Strike Data. On the right, the same data consolidated and processed with InterLeeV's indices and system showing historical and projected risk points, or where lightning is likely to strike in the future.

The Product

InterLeeV (ILV) has developed a new efficient way to integrate big data and privileged data with blockchain. From the input side, these tools allow semi-automatic query and indexing of hierarchical, network, relational, and object-oriented databases. Each of the indices is a single 64-bit word, a hashtag. These hashtags are combined into an ILV Master Key, which allows closer-than-ever-before real-time retrieval of data for virtual reality and augmented reality applications. This process is summarized in Figure 2.

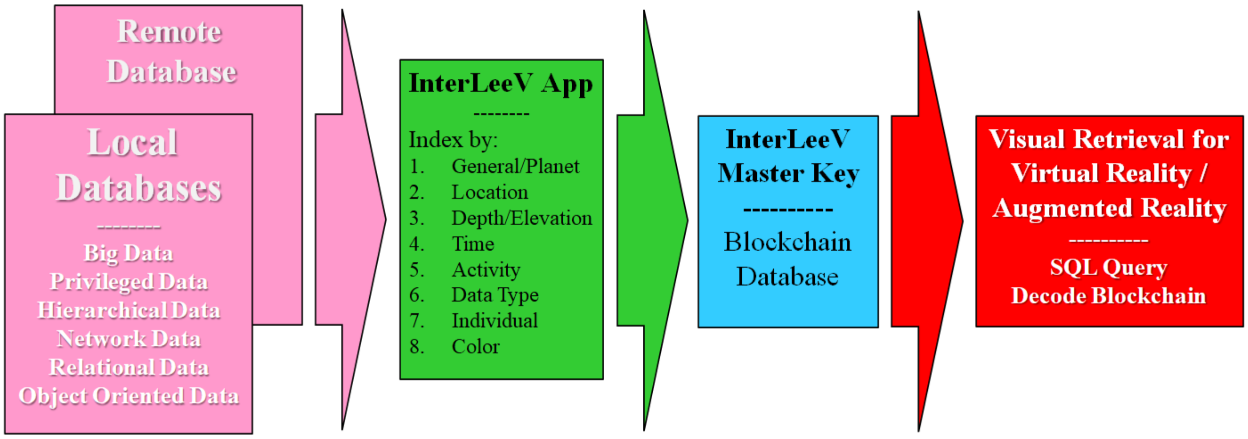

Figure 2. The process of indexing and retrieving data using the ILV App and ILV Master Key.

The Infinite Grid^{sm} – A Spatial Index

Typical projects start with indexing locations within a database. Conventional mapping uses coordinates from a reference datum to draw a vector with direction and magnitude. These vector-based maps are referenced as points, or individual longitudes and latitudes. This keeps file sizes small, but for complicated maps, like a map with all of the telephone poles on the California coast, the maps have long screen drawing times. The Infinite Grid^{sm}, the ILV Location Index, divides longitudes and latitude into tiles of area. All features and information are referenced by area and not by vector. The difference is subtle, but very far reaching.

Using an equidistant longitudinal projection, the highest-level Infinite Grid^{sm} divides the world into 9 x 9 tiles, each tile covering 45º longitude and 22.5º latitude. Cells are numbered from the lower-left, starting with IG1-11 to IG1-99. The United States is covered by IG1-26 and IG1-36. Each IG1 tile is divided into 8 x 8 IG2 5º longitude and 2.5º latitude tiles, giving the Great Salt Lake an IG2-2658 index number. Each IG2 tile is divided into 5 x 5 IG3 1º longitude and 0.5º latitude tiles, meaning that Iron County, Utah is included in IG3-265711-265732 and the northern n part of IG3-265625. Each IG3 tile is divided into 8 x 8 IG4 7.5' longitude and 3.75' latitude tiles, meaning Cedar City, Utah is between IG4-26572183 & IG4-26572184. These IG3 tiles match USGS (United States Geological Survey) Topo Maps. This is illustrated in Figure 3.

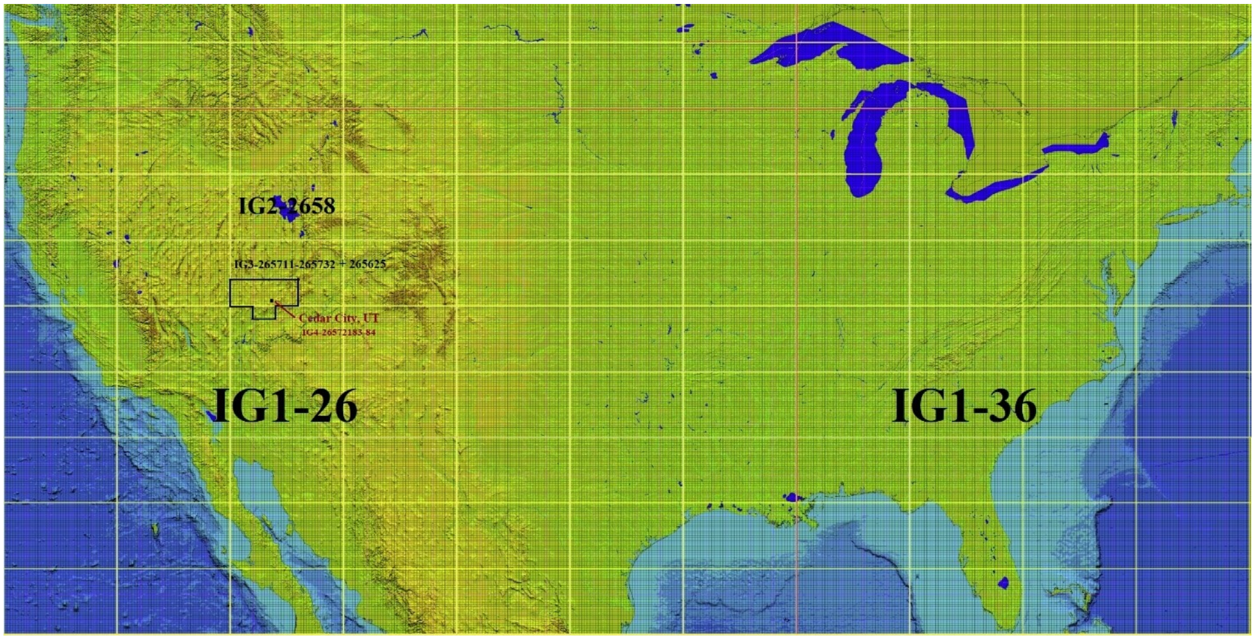

Figure 3. Infinite Grid^{sm} of the United States showing IG2 Great Salt Lake, IG3 Iron County, and IG4 Cedar City, Utah.

This tiling process continue to get infinitely smaller. IG5 is 9 x 9 tiles at 50” (seconds longitude) x 25” (seconds latitude), IG6 is 5 x 5 tiles at 10” x 2”, IG7 is 5 x 5 tiles at 2” x 1”, and IG8 is 8 x 8 tiles at 0.25” x 0.125”. An IG8 tile is 5.9330 meters x 3.8439 meters at 40^{o} latitude, or about the size of 4 mini vans parked close to each other. Each IG8 tile on planet Earth is defined by a single 64-bit word (smaller for Mars, larger for Jupiter), or a hashtag, and provides an ideal spatial index for anyplace on any planet. Adding an additional word, the Infinite Grid^{sm} tiles can be several orders of magnitude smaller that IG8 tiles. Figure 4 shows 0.1 mm tiles across a core slab. Figure 5 shows an IG3 grid across a satellite image of Guatemala processed to show the Agricultural Stress Index (ASI) across the country.

The bottom line is this new IG index provides an easy extensible way to quickly, easily, and securely index anything spatially. Tools are being built which will allow a user to bring up a map of the world, drill down to whatever level they are interested in indexing, and get the IG at that level for that location. There are tools available for taking a table of longitudes and latitudes and converting them to IG indices. As these tools are enhanced, the IG hashtag will be automatically inserted in block chains of the user choice. Since the user is selecting and is the only one who knows the location, this hashtag provides an extra layer of security for the blockchain being added to, whether that blockchain is part of a bitcoin transfer, or some other type of secure blockchain information transfer.

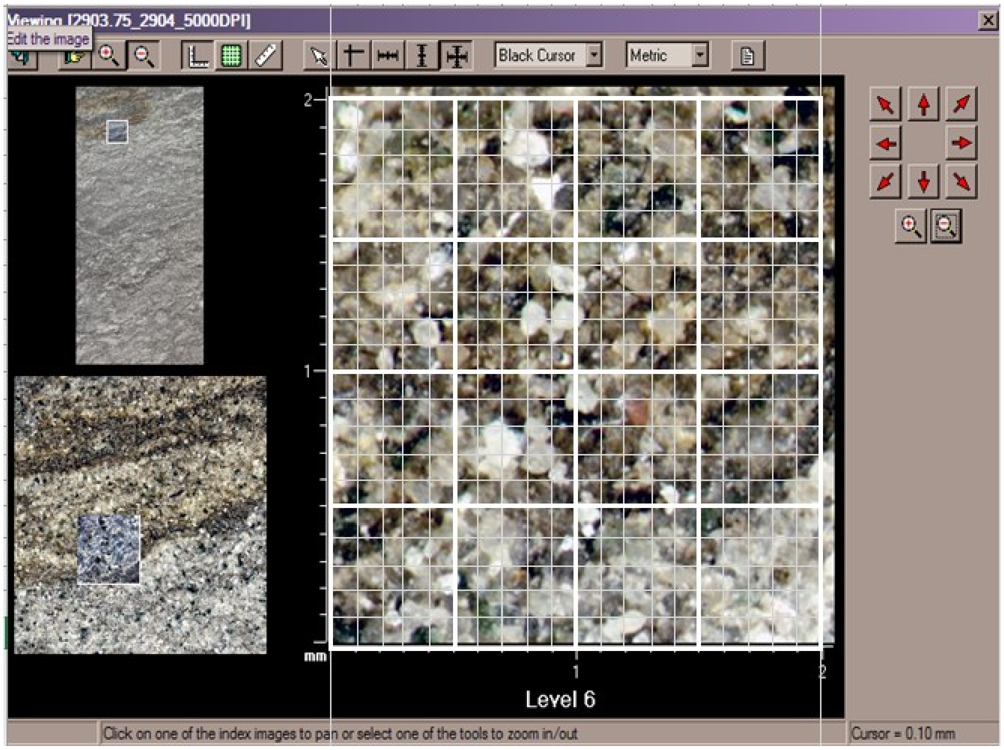

Figure 4. 4-orders-of-magnitude finer IG tiles at 0.1 and 0.5 millimeter on a digital core slab.

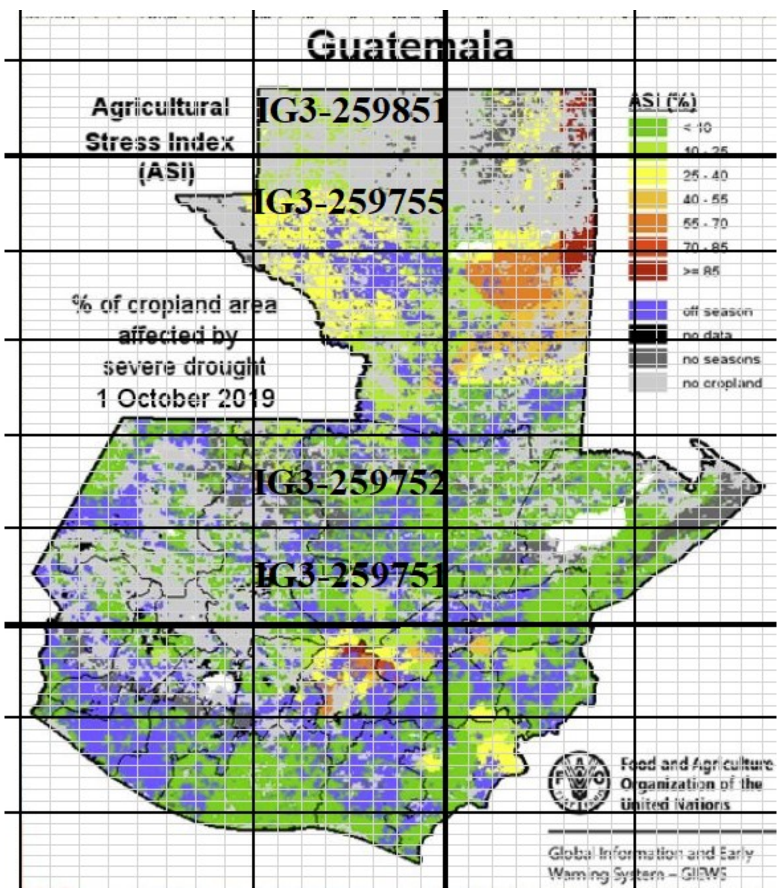

Figure 5. IG4 satellite, data processed to show the Agricultural Stress Index of Guatemala.

Spatial Indexing Options

ESRI (Environmental Systems Research Institute) has a 43% share of GIS (Geographical Information System) software, web GIS, and geodatabase management applications. ESRI's product, Arc/Info, is an existing vector-based system, with small data files, tied to physical things, and universal in some business areas. The system is computationally intensive and is based on database point and polygon vertices coordinates. It is not inherently graphical, is not convenient to search, and has a multitude of incompatible file standards. The Infinite Grid^{sm}, has larger data files and requires image capture or a conversion step from vector files. However, with the IG, it is easy to merge and compare unrelated data. The IG can be used as an interchange format. The IG is easy to visualize and is self-consistent. The IG is dynamically drillable. The IG is area based, rather than point based. And the IG is inherently graphical.

Uber developed an alternative GIS system they call H3. H3 combines the benefits of a hexagonal global grid system and a hierarchical indexing system (https://www.uber.com/blog/h3/). Using a hexagon as the sell shape is key to their application, since there is only one distance between a hexagon center-point and its neighbor. Squares have two distances, and triangles have three distances. Again this is largely a vector problem, tied to the route Uber cars take to pick-up and deliver passengers.

Google has an app called "what3words" where every 3-meter square on the planet is defined by a unique combination of 3 words (https://what3words.com/). To use this application, you start by putting in the longitude and latitude for a place of interest. The app returns the 3-words defining this location, the longitude and latitude, and the longitude and latitude of the southwest an northeast corner of the square at that location. You can translate these three-word addresses into a multitude of languages. You can take a 3-word address and turn it into coordinates. This addressing system is optimized for voice. IG Tile numbers are a much simpler way to index data worldwide.

The Infinite Grid^{sm} is a mapping, as well as a spatial indexing system. Most maps are rectangular. Projections vary, all projections of all historical maps can be transformed, rectified and calibrated against IG grid cells, simply by identifying key points and morphing the image to match IG corners. To illustrate the power of this approach for simple mapping applications, imaging several maps calibrated against the IG. Size all of the maps to the same scale covering the same IG tiles. Place each map as a background image in a standard spreadsheet worksheet. Match the worksheet cells with the IG cells on the background image. Sponge up or type in a number for each cell on each of the worksheets. Then do Boolean calculations between spreadsheets to get desired results. Worksheet 1 could have a "1" in each cell where there is a road. Worksheet 2 could have a "1" in each cell where there is a drainage ravine. Subtracting Worksheet 2 from Worksheet 1 shows "0" where there is a bridge which needs to be maintained. Etc. The IG is a natural way to integrate published maps, like physical maps, topographic maps, political maps, weather maps, economic maps, resource maps, population maps, or world maps, which are almost never hexagonal in shape. Each IG tile knows north, has a unique scale, and has a border. The title, legend, and acknowledgement should be documented for maps sponged from other sources.

Other Indices in the ILV Master Key

The InterLeeV, or ILV Master Key, combines up to 8 different 64-bit word hashtags. An easy way to think of it is to consider a gridded map of topography, where the grids are IG7 tiles (47.464-meter x 30.751-meter cells at 40^{o} latitude). The topography value is the value at the center of each IG7 tile. Merging the IG index and the Depth/Elevation Tag and plotting them as a 3-D grid provides a 3-D bar chart topography map for the area being studied. If the color of a satellite image, where the colors are tuned to show methane seeps in red, is added to color each bar in the bar chart, the result is a 3-D bar chart showing red where methane seeps are occurring along fault scarps. The ILV Master Key has all of this information embedded in a blockchain ledger which can be decoded by the ILV App.

There are up to 8 ILV tags in the ILV Master Key. Some tags, like the Location Tag, or the Time Tag, or the Individual Tag can be repeated with separate locations, or times or individuals in a story referenced by the ILV Master Key. The only required tag for a MLV Master Key is the Location Tag. For the next few years, we anticipate the Planet default tag will be Earth. Some applications will require all of these tags, and possibly multiples of some tags, and most applications will require 1-4 of these tags. Here is a list of the different hashtags:

1. Planet and General Information Tag (P)
2. Location Tag [Infinite Grid^{sm} (IG)]
3. Depth/Elevation Tag (DE)
4. Time Tag [TimeDex^{sm} (TD)]
5. Activity Tag [Knowledge Backbone^{sm} (KB)]
6. Data Type Tag (DT)
7. Individual Tag (ID)
8. Color-Index Tag (CI)

To get a little philosophical, in the long-run, children are more important than parents. Without children there is not a continuation. In the digital world, indices are more important than stories or databases. The digitization of everything is the most significant transformation of our society since the invention of the Gutenberg Press. By the late 1400's, there were so many books printed, no one could find anything, that is until the index was invented. Likewise, today we have so much new data, information, knowledge, and wisdom, we can not find what is relevant with key word searches. We need to index big data and privileged data so we can find relevant material in context and keep secret data secure. InterLeeV's Master Key is designed to do this. The U.S. Patent office issued a Patent to InterLeeV LLC on November 8th, 2022, for the building blocks of the 64-bit ILV hash tags, and ILV's Master Key.

Case History – Lightning Analysis Cedar City, Utah

The basic concepts behind the IG are used by Dynamic Measurement LLC (DML), which is creating a new branch in the geophysical services industry. DML has 2 patents and an exclusive license to the best lightning strike data bases available. DML has discovered that even though lightning is one of the most random events in nature, where lightning strikes the ground is controlled by subsurface, or telluric currents. DML has developed proprietary algorithms and procedures to clean, process, map, and interpret lightning strike databases in order to explore for natural resources. To do this work, we had to develop ways to organize thousands, and even millions of lightning strikes anywhere. DML used the concepts, which became the Infinite Grid[sm], to spatially index each lightning strike within a project area, as well as the concepts behind the TimeDex[sm] to temporally index the time of each strike to the nearest microsecond.

To share a case history example, in 2022 DML was a subcontractor on a project to locate untapped aquifers for Cedar City, Utah in the surrounding mountains. The topography of the area, and the number of lighting strikes in the area studied, are shown on Figure 6.

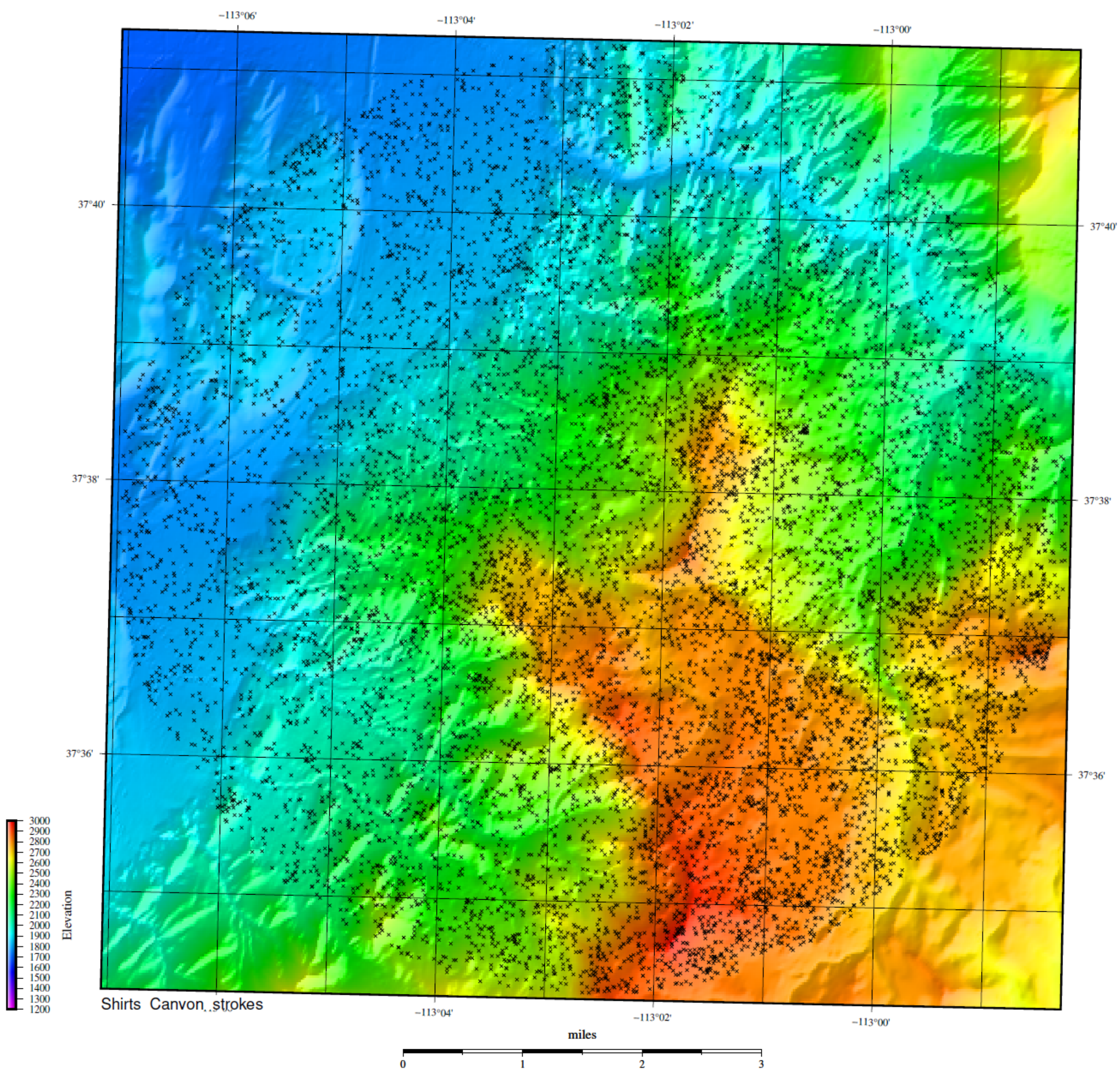

Figure 6. 25-Years of Lightning Strikes within an 8-mile diameter area over Shirts Canyon, south-southeast of Cedar City, Utah.

There were 11,364 lightning stokes recorded over 24.04 years of available data. DML did a data cleaning process. There were 8,589 lightning strokes used for the analysis, after the cleaning was completed. Each stroke was assigned what is now known as an IG7 cell (50.691 meters x 30.743 meters at 35 degrees north longitude). There are typically 1-2 lightning strikes per square kilometer per year in the deserts of Southern Utah. In Florida and Louisiana there can be over 14 lightning strikes per square kilometer per year. When lightning analysis is done over an entire county, it is easy to have millions lightning strikes to manage in a single project. Having each lightning strike indexed by location, elevation, and time are key components of DML Analysis.

DML receives several attributes for each lightning strike from the vendor, including date, time latitude, longitude, peak current, chi (reliability of correlation between different sensors), uncertainty semi-major and semi-minor axes, peak-to-zero time, rise-time, and number of sensors recording the strike. These measured values are stored in IG7 cells to create 25 different maps. Maps include calculated attributes like Total Time (Rise-Time + Peak-to-Zero time), Wavelet Symmetry (Rise-Time / Peak-to-Zero Time), and 2 calculated rock properties (Apparent Resistivity and Apparent Permittivity). The maps for Shirts' Canyon are shown in Figure 7.

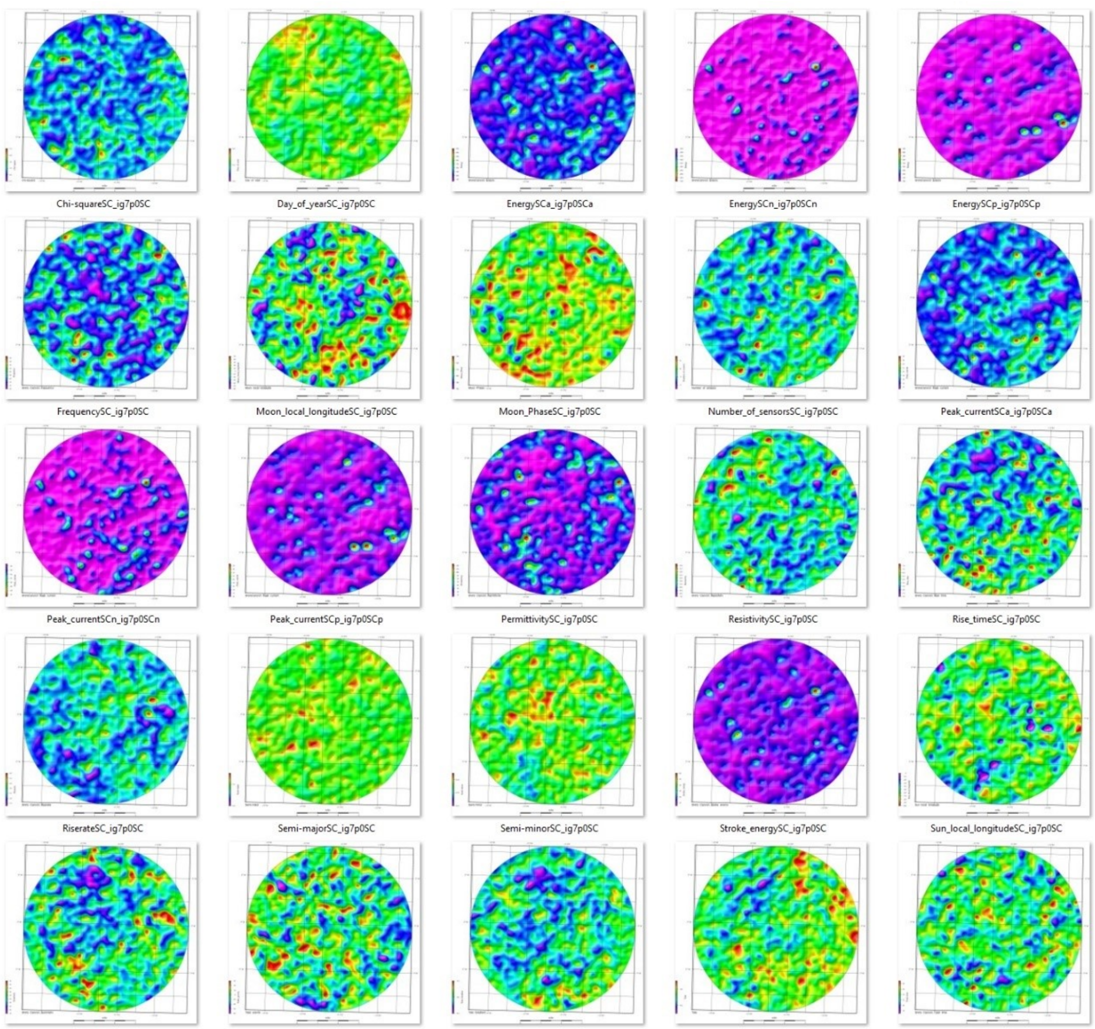

Figure 7. 25 maps generated from lightning strike data for Shirts' Canyon, Utah.

There were also 25 volumes created for the same attributes mapped in Figure 7. Assuming the earth's surface acts as an electromagnetic half-space, the depth to the telluric (earth) currents which are interacting with atmospheric currents, is the same as the height of the lightning stroke. This height is a function of the Peak Current of the lightning stroke, since the stronger the Peak Current, the further the lightning stroke can travel through the atmospheric dielectric (insulating barrier). So the depth of interaction of each stroke is calculated, and the attribute value measured or calculated from each lightning strike is placed at the calculated depth. A 3-dimensional interpolation algorithm is used to create volumes. Then the depth is shifted up by the topographic value for each IG7 cell so the top of the volume matches the topography. These volumes are turned to traces, with one trace for each IG7 tile, exported as SEG-Y data, loaded in a geophysical interpretation workstation, and then interpreted.

Figure 8 shows an example of the results of this process from Cedar City's Shirts' Canyon SPOT^{sm} Lightning Analysis. Note the two high-resistivity lineaments seen on the 1,750 meter depth-slice, and the same lineament on the 1,000 meter depth-slice. An interpretation is these are transverse or strike-slip faults which were caused by compression when the Pacific Plate was pushing against the North American Plate during Cretaceous geologic time. This type of faulting is very deep, and water, with higher resistivity than the surrounding rocks, tends to be pushed up through the skree along this kind of fault plane. The two horizontal and two verticl cross-sections intersect where a higher resitivity anomaly comes close to the surface.

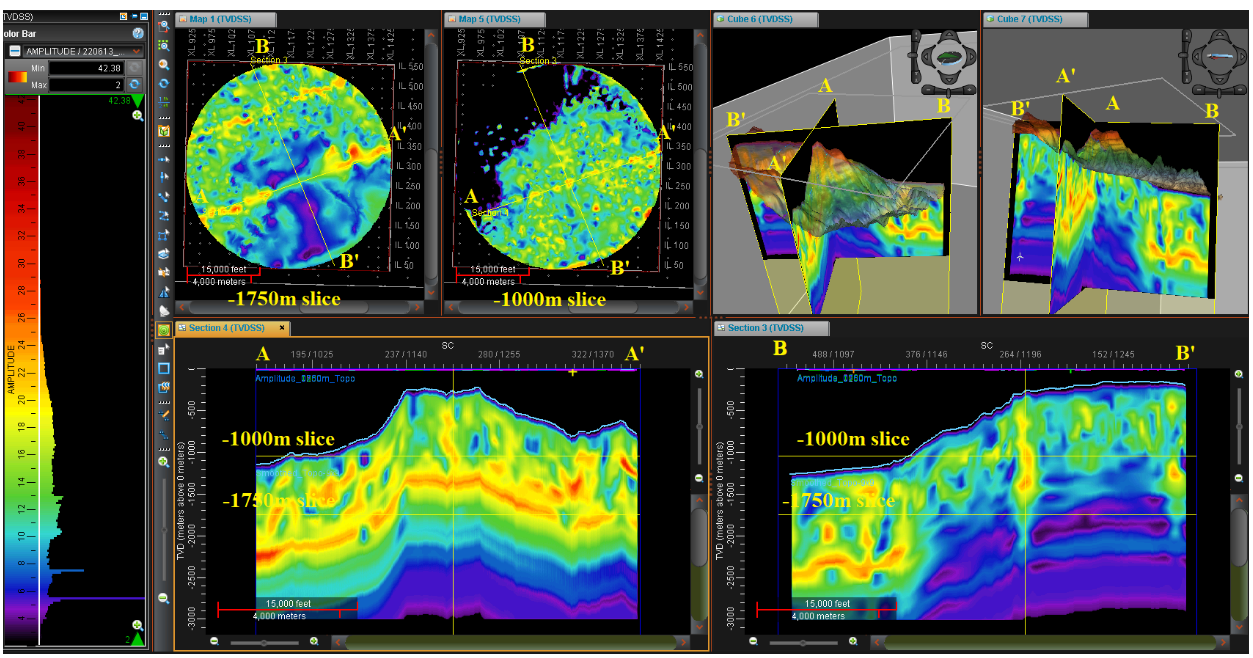

Figure 8. Screen capture of two depth-slices (10,000 meters and 17,250 meters), two rotated perspective displays which include semi-transparent topography and the strike (A-A') and the dip (B-B') cross-sections also displayed on the bottom of the screen.

At IG7 cell resolution, many of the IG7 tiles will include several lightning strikes. DML averages all of the values in each tiles, for each 10 meter increase in depth, from the surface to the deepest value (at Cedar City about 3,000 meters depth or 300 samples). Note in Texas, Louisiana, and Florida there are strikes which travel 9,000+ meters (which converts to about 30,000 feet depth). After 3-D interpolation there were 300 values for each trace. Remember, this interpolation is done for each of the lightning attributes shown on maps, like shown in Figure 7.

Because there are the equivalent of TimeDex^{sm} indices, it is also possible to create lightning attribute volumes for specific time-ranges. DML has done several time-lapse studies, creating volumes of lightning-derived apparent-resistivity, in three-year-intervals. This allows generation of time-lapse volumes to evaluate how apparent resistivity has changed across time. One project was to evaluate how wastewater injections build up subsurface pressures over time in Oklahoma where there was a fluid injection caused earthquake. This is ongoing in-house research.

Another result from the indexing is tied to strike frequency at specific IG7 locations. With a database of 24-years of lightning strike data, DML graphs strike frequency vs the probability of a strike occurring within 200 meters of an IG7 tile (Figure 9). These results are calculated for IG6 and IG7 (as shown at the top of Figure 10) then plotted in map view as Risk Point Map (bottom of Figure 10). Knowing these areas of concern can be of great interest to the public for a variety of reasons including, where to post lightning warning signs, where not to hike or ride a bike during a thunderstorm, and where to monitor for lightning caused forest fires.

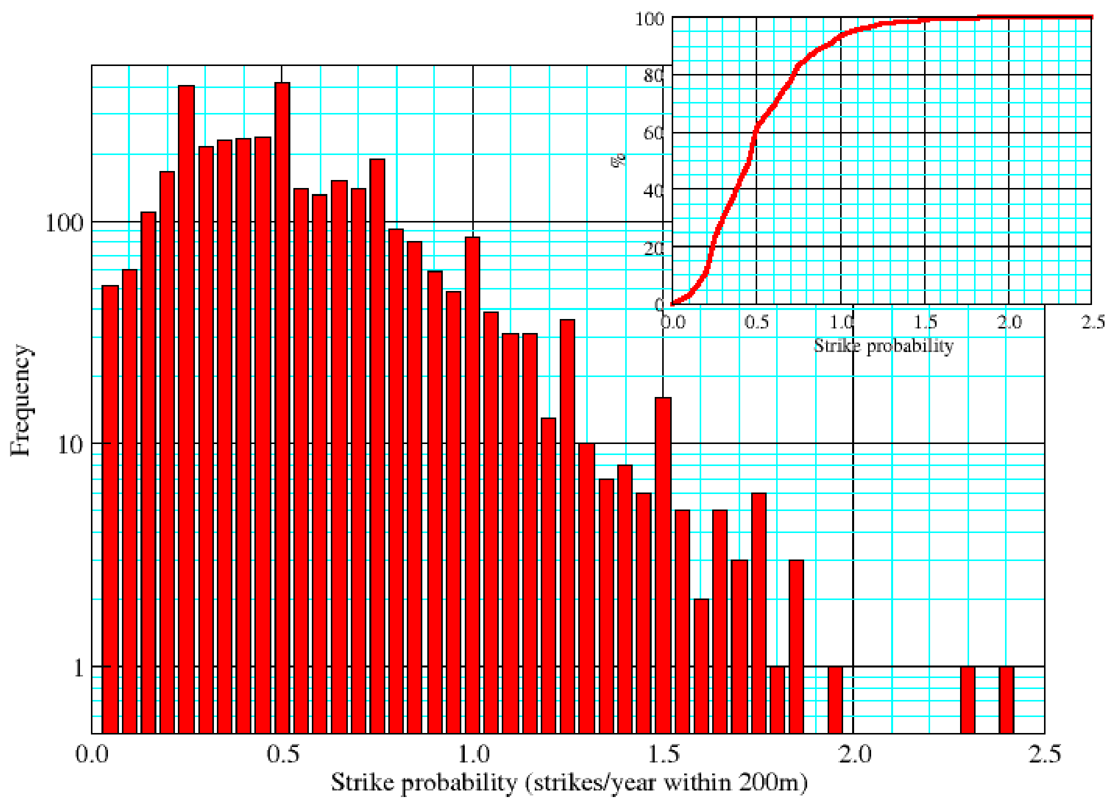

Figure 9. Probability of a lightning strike within 200 meters of an IG7 tile.

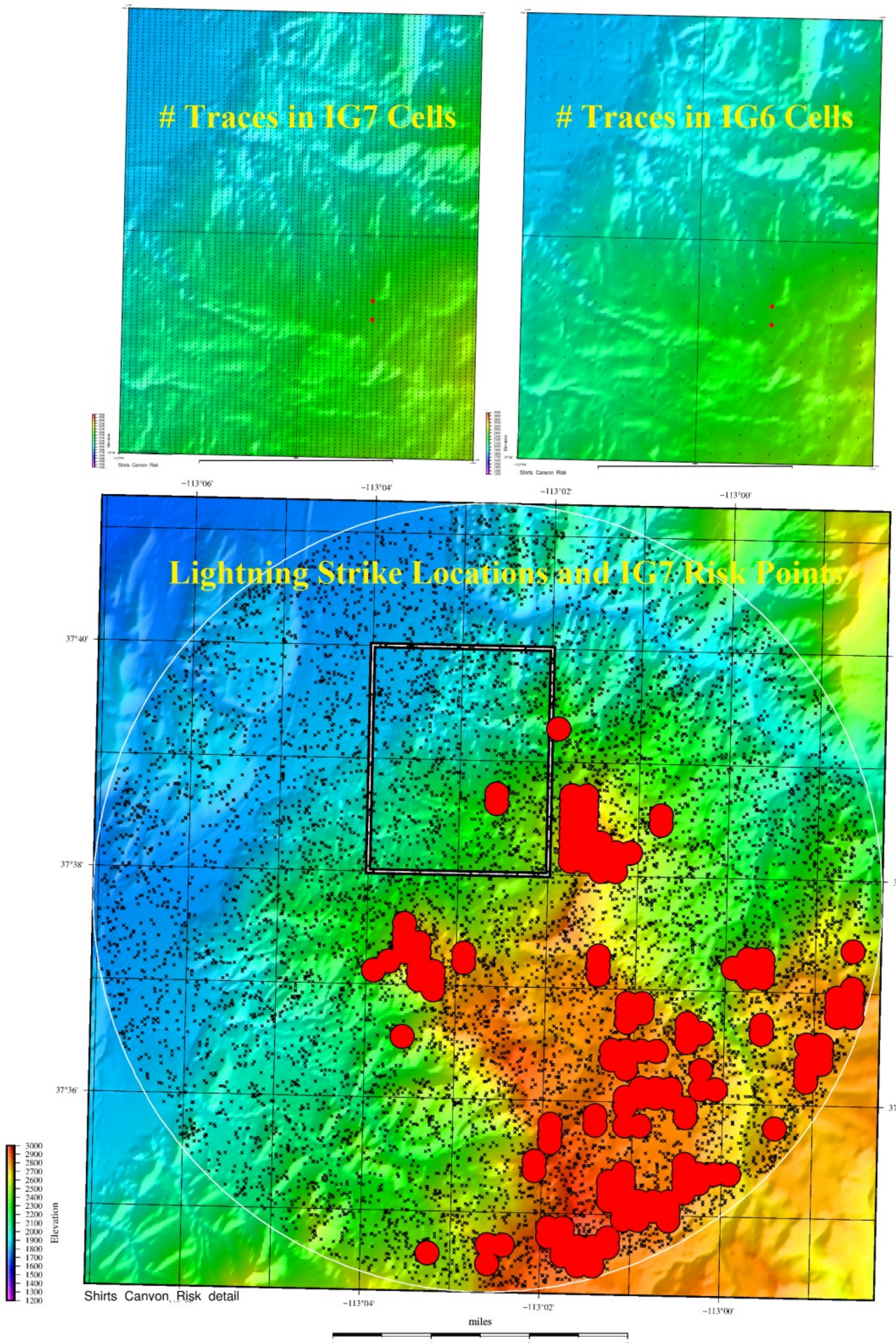

Figure 10. Maps show Shirts' Canyon topography, lightning strike locations, and places where 1.8 or more lightning strikes are likely to occur each year within 200 meters. The top two maps show number of traces in each IG7 and IG6 cell in a the year period (2000-2009), and where the two risk points are in the 2 horizontally divided IG5 cells.

Summary

A digital revolution is occurring, and it is impacting all aspects of society and our personal lives. Metaphors of current computer systems are still from the industrial era: word processors mimic typewriters; spreadsheets mimic ledger sheets; Computer-Aided-Design (CAD) mimics drawing boards; databases mimic storage silos; and digital-retrieval mimics alphabetical card catalogs.

Immense amounts of new multidimensional digital data, information, knowledge, and wisdom are available. Society needs new multi-dimensional enterprise-wide classification approaches to finding relevant digital information in context efficiently. Finding the data will enable easier and better integration of this digital knowledge into wisdom-based virtual reality planning and computer-in-hand augmented reality implementations.

Like using lightning strike databases to explore for new natural resources. The block-chain hashtag master-key approaches described in this white paper are logical first steps on this journey through an unfamiliar environment. A digital environment with an unfamiliar ecology, and untested ecosystems. Building the new digital onion starts with efficient integration of data.

This white paper introduces a logical and systematic block-chain-based indexing and retrieval approach. This is an approach which can be applied to big data and particularly spatial data, immediately. The additional indices provide a path and a process to fine-tune the indices in order to provide better and better context for both searching and data retrieval.

