John Fildes, Ph.D.
Spotlight Business Consulting LLC
www.SpotlightBC.com
jfildes@SpotlightBC.com
(630) 248-0836

Analytics is Important in Litigation-Related Technical Investigations

Analytics, which involves data mining and modeling, is likely not thought of as part of a litigation-related investigation, but it should be as indicated by Virginia Tech's College of Science having made analytics an overarching theme of its graduate degree programs. Accidents and product/structure failures do not happen in a lab under controlled conditions and the watchful eye of measurement instrumentation, and although there is often data, it is limited in its scope and amount, and limited in its applicability since it is often contaminated with uncertainty and unrelated factors. Analytics provides highly effective tools to overcome these data challenges and this approach is extremely powerful, but it requires specialized software, experience, and training in the applicable scientific principles, data mining, and modeling techniques.

How is Data Mining Used in Litigation Technical Investigations?

"Field data," as is available from accidents and failures, can come from sensors such as on automobiles and planes and from industrial processes. This data exists for purposes that are typically unrelated to creating a data repository for future failure analysis. For example, sensors may be providing data needed for a control system such as an auto cruise control or for a safety system such as to prevent fires and explosions in process equipment. Industrial processes and civil and construction processes typically contain sensors supplying data for process control and sometimes for regulatory reporting. The utility of this type data for failure analysis in a litigation-related technical investigation is limited by it not being uniquely determined by the accident or failure, not being all of the type of data that is needed to fully characterize the accident or failure, and either not being collected often enough causing gaps in the data or being collected too often causing an overload of data.

Data mining uses specialized, sophisticated software to visualize and identify patterns in the data that is available. Plotting data in different coordinate systems, or first transforming the data, can reveal systematic trends. Mathematically transforming data can make trends more obvious. The ways in which the data is plotted or transformed is sometimes suggested by the underlying scientific principles and sometimes by data analysis methodologies for discovering trends in the data that are not easily seen. The data can also be characterized statistically to establish the variability of the data, which can be used to identify what portion of the available data comes from the same population and is therefore similar and related to the same underlying scientific principles. Treating the data, especially data collected at different times, as vectors allows cluster analysis by statistical methods or by newer and potentially more effective neural network techniques, which provides another way to identify which samples of data are similar and which are different. Data mining may need to be an iterative process, especially when a large amount of data exists or when extensive mathematical transformations are used, in which case smaller subsets of data may be first analyzed to waste less time on analyses that produce little insight. Although data mining is powerful in the insight it can produce, it is common for significant gaps to still exist in the data.

Case Study

Virtually any type technical situation can provide a good case study of the value of analytics. The one described herein involves a claim of exposure to particulate matter from a nearby industrial facility.

Chemical analysis results are common in this and many other incidents. Although the chemical analysis is done in a well-controlled manner, the sample collection is typically not done in a controlled manner in litigation situations. This makes it challenging to determine the similarity between the chemical analysis and the claimed source of the sample, which often involves comparing several to tens of chemicals species. Statistical analyses where sufficient data exists to compute confidence or tolerance intervals can be effective. Treating the concentrations of the several to tens of chemical species in the sample as components of a vector and performing statistical or neural network cluster analysis can also be effective to establish similarity.

The industrial process from which the collected sample is claimed to have come can also be modeled from scientific principles and known materials properties to predict the probable chemical composition that the collected sample should have if it came from the industrial process. Modeling of this nature augments existing data from the industrial process and allows for estimating how a sample could vary due to the inherent variation in the industrial process. The chemical composition of other potential sources can also be included to test their contribution as can the transport of the material made in the industrial process to determine the conditions under which that materials could reach a remote location.

John Fildes, Ph.D.
Spotlight Business Consulting LLC
www.SpotlightBC.com
jfildes@SpotlightBC.com
(630) 248-0836

How is Modeling Used in an Investigation?

Modeling addresses the issue of gaps in the available data by using trends discovered by data mining, fundamental scientific principles, and known properties of materials to estimate the data needed to fill the gaps. A significant amount of data is available from academic research and industrial R&D, and although this data is often valuable, it is usually not generated specifically for the situation under investigation and needs to be adapted. The estimates produced by modeling are tested and validated by estimating data that already exists. The fundamental scientific principles that are used come (to a substantial degree) from: thermodynamics, which determines which processes can happen, chemical reactivity, which determines the rate at which a process happens, chemical bonding, which relates to the strength of materials and their susceptibility to degradation, solid state physics, which determines the properties of semiconductor materials and devices, and general physics, which underlies mechanical properties. These disciplines cover the strength and fracture of materials, corrosion, thermal degradation, chemical compatibility, vaporization and evaporation, adhesives, expansion/contraction of wood and laminates, and many others.

Some attorneys and even some experts may not appreciate the importance of analytics in litigation-related technical investigations and using analytics is challenging. Analytics requires a deep knowledge of scientific principles so that they can be applied in the unusual situations that accidents and failures often present. Analytics also requires extensive knowledge of statistics, data modeling, and increasingly of artificial intelligence methods, with which one has to be sufficiently experienced to apply in a practical way. Analytics also requires comprehensive searching for relevant data and the experience to select meaningful subsets of data during the exploratory screening phase so as to make the time required practical. This is a demanding combination of skills and experience that some experts may lack. Nonetheless, the importance and value of using analytics in some litigation investigations, especially ones that are more complex and multidisciplinary, cannot be overstated.

A message from the Dean of Virginia Tech's College of Science stated "At the Virginia Tech College of Science, we have reimagined scientific research…. We are focused not on data itself, but amplifying the relevance of that data with analysis, modeling, and interpretation." (Va. Tech Science, Fall 2019).

Combining empirical trends in the data uncovered by data mining in litigation-related technical investigations with estimates made from modeling based on fundamental scientific principles using data generated by academia and industry not only fills gaps in the limited data available from the accident or failure but is also uniquely able to provide insight as to what scientific principles were violated and why they were violated, resulting in the accident or product failure. Testing the hypotheses made in an investigation is inherent in this approach, which is extremely valuable since testing hypotheses is an essential aspect of the scientific method that guides the normal conduct of scientific investigations and that is also fundamental in litigation-related technical investigation to meeting the rules of evidence. The author's experience is that analytics is a substantial contributor to why the opposing side has sought to settle cases on the basis of his reports, sometimes without deposition.

Dr. John Fildes has a Ph.D. in physical chemistry, a B.S. in chemistry, and he was a post-doctoral research associate in a chemical engineering department. Physical chemistry provides the scientific basis for many engineering disciplines. Thermodynamics provides the basis for metallurgy, materials science, fire and explosion science, and others. Chemical bonding provides the basis for the strength of materials and electronic materials and devices. Electrochemistry provides the basis for corrosion science, and chemical kinetics provides the basis for chemical compatibility, reactivity, volatility, and chemical processes. Dr. Fildes has conducted over $27.5 million of funded R&D and/or litigation-related investigations in these areas because he is well experienced in the fundamental scientific principles as well as in analytics and chemical safety. He led large groups of scientists and engineers at Northwestern University and two scientific/engineering firms licensed to practice professional and structural engineering that conducted thousands of litigation-related technical investigations, which also makes him an expert in the conduct of these investigations.