Software engineering research: the need to strengthen and broaden the classical scientific method

Gonzalo Génova
Juan Llorens
Jorge Morato
Universidad Carlos III de Madrid, Spain

ABSTRACT

The classical scientific method has been settled through the last centuries as a cyclic, iterative process of observation, hypothesis formulation, and confirmation/refutation of hypothesis through experimentation. This “experimental scientific method” was mainly developed in the context of natural sciences dealing with the physical world, such as Mechanics, Thermodynamics, Electromagnetism, Chemistry and so on. But when we try to apply this classical view of the scientific method to the various branches of Computer Science and Computer Engineering, among which Software Engineering, we find two kinds of obstacles. First, Computer Science is rooted both in formal sciences such as Mathematics and experimental sciences such as Physics, therefore an excessive emphasis on the experimental side is not appropriate to give a full account of this kind of scientific activity. Second, the production of software systems has to deal not only with the behavior of complex physical systems such as computers, but also with the behavior of complex human systems (developers interacting with stakeholders, for instance, or users interacting with machines) where educational, cultural, sociological and economical factors are essential. Therefore, empirical methods in their narrow sense, even though valuable in some respects, are rather limited to understand a reality that exceeds the mere physical world. Moreover, neither formal nor empirical methods can provide a full account of scientific activity, which relies on something that is beyond any established method. Qualitative (i.e. meta-methodical) reasoning plays the directive role in scientific activity. In this chapter we claim that acknowledging a plurality of research methods in software engineering will benefit the advancement of this branch of science.

INTRODUCTION

In the final report of the ACM Task Force on the Core of Computer Science, Denning et al. presented an intellectual framework for the discipline of computing intended to serve as a basis for computing curricula (Denning et al., 1989). They characterized computing as a discipline that sits at the crossroads among mathematics, empirical science and engineering. Therefore, there is a plurality of research methods that can be applied for the advance of computing in all its branches, one of them being software engineering. The study of mathematics and empirical sciences is recommended in software engineering curricula to promote abstract thinking and the appreciation for the scientific method (IEEE, 2004).

Henri Poincaré wrote at the beginning of the 20th century about the experimental scientific method: “The man of science must work with method. Science is built up of facts, as a house is built of stones; but an accumulation of facts is no more a science than a heap of stones is a house” (Poincaré, 1952, p. 141). A good scientific work is not complete with a systematic recollection of data: it requires a rational explanation of their relationships. In other words, the essential ingredients of the scientific method are the description of phenomena (what happens, how it happens) and the explanation of their relationships (why it happens): “What and How describe; only Why explains” (Whetten, 1989).

Michael Polanyi, another scientist-philosopher, claimed that the scientific method is not a recipe that can
yield truths mechanically (Polanyi, 1958): explaining the relationships between observed phenomena requires intelligence, imagination, and creativity. Besides, a naïve but widespread empiricist account of science forgets that observed facts are not independent from theory (Chalmers, 1999): on the contrary, establishing the validity of observations requires human interpretation and a good deal of previous knowledge. The application of the scientific method in software engineering research has to take into account these warnings. Our main purpose is to show that the advance of science cannot be closed within the application of a “method”, however rigorous it may be. There is no universal method in science. In particular, empirical methods are rather limited to understand a reality that exceeds the mere physical world. Each branch of science needs its own method, and selecting the most adequate method is a task that lies beyond any formal or empirical method: it is a meta-methodical task.

In this Chapter we present a brief historical background of the development of the scientific method through the centuries. Then we discuss some philosophical issues that have arisen from the consideration of how scientists work, showing that the scientific method is not fully justified, and perhaps will never be. These issues are relevant when the scientific method is applied to software engineering, as we explain in the last part of the Chapter, where we emphasize the guiding role of qualitative, speculative or conceptual reasoning in research activities.

HISTORICAL BACKGROUND

The scientific method is rooted in the Greek culture, particularly in the laws of logic first defined by Aristotle (384-322 BC). However, Greek science was almost entirely lost for Latin West during the Early Middle Ages, after the fall of the Western Roman Empire. Science was meanwhile safeguarded and cultivated in the Arab culture, with influences from China and India. The main contribution of Indian culture that arrived to Europe through Al-Khwārizmī (c.780-c.850) was the place-value numeral system, which was fundamental for the development of algebraic and arithmetic operations (Mason, 1962). Another significant milestone in the recovery and development of scientific method is the Book of Optics by Muslim scientist Ibn al-Haytham (Alhazen, 965-1039), with his pioneering emphasis on the role of experimentation (Saliba, 2007). Aristotelian logic was also preserved in the Arab civilization, where it was studied by Islamic and Jewish scholars such as Ibn Rushd (Averroes, 1126-1198) and Moses Maimonides (1135-1204). The whole set of Aristotelian logical works (known as Organon, “instrument”) was not available in Western Christendom until translated to Latin in the 12th century. Then it became a major field of study and significant advancement for medieval Christian scholars, who regarded Aristotle as “The Philosopher”, in large part due to the influence his works had on Thomas Aquinas. The interest in logic as the basis of rational enquiry is manifested in the many systems of logic developed hereafter, but a real advancement did not arrive until the formulation of modern predicate logic in the 19th century (Łukasiewicz, 1957).

Robert Grosseteste (1175-1253) must be counted among the first scholastic thinkers in Europe to understand Aristotle’s vision of the dual nature of scientific reasoning (Crombie, 1971): a process that concludes from particular observations into universal laws, and then back again, from universal laws to prediction of particulars. Roger Bacon (1220-1292), inspired by the writings of Grosseteste, described a method consisting in a repeating cycle of observation, hypothesis, experimentation, and the need for independent verification (Hackett, 1997), surprisingly advancing what would be called much later the hypothetico-deductive method by William Whewell in his History of the Inductive Sciences (1837). However, the admiration for Aristotle was not universal. Francis Bacon and René Descartes were among the first thinkers to question the philosophical authority of the ancient Greeks (Cantor & Klein, 1969).

Bacon published in 1620 his Novum Organum (“The New Organon”) intending to replace traditional logic with a new system he believed to be superior. Bacon strongly criticized Aristotle, “who made his natural philosophy a mere slave to his logic” (Book I, Aphorism 54). According to Aristotle, scientific knowledge pursues universal truths and their causes, and this is achieved only by means of deductive reasoning in the form of syllogisms: it is deduction that allows scientists to infer new truths from those already established. On the contrary, while inductive reasoning is sufficient for discovering universal laws by generalization, it does not succeed in identifying the causes of observed phenomena. Therefore, although empirical observation has a place in the Aristotelian method, knowledge acquired by induction is not truly scientific and reliable.

In contrast, induction from the particular to the general occupies the first place in the Baconian method to investigate the cause of a certain phenomenon. In successive steps known as method of agreement, method of
difference, and method of concomitant variation, Francis Bacon compares different situations where the phenomenon occurs, does not occur, or occurs in different degrees, trying to find a factor that could be hypothesized as the cause of the investigated phenomenon. The proposed hypothesis must be scrutinized and compared to other hypotheses, so that the truth of natural philosophy (i.e. the laws of nature) is approached by a gradual ascent. The method was never described in full because Bacon’s work remained unfinished. In fact, this inductive method had been essentially described much earlier by Persian philosopher Avicenna (Ibn Sina) in his book *The Canon of Medicine* (1025) (Goodman, 2003). John Stuart Mill systematized and expanded the method in his book *A System of Logic* (1843).

A radically different approach to the practical and empirical method of Bacon is followed by Descartes: in his *Rules for the Direction of the Mind* (1619) and his *Discourse on the Method* (1637), he emphasizes the theoretical and rational aspects based on deduction, in order to avoid deception by the senses. Both Descartes and Bacon aim at discovering the laws of nature, either by deduction from first principles or by induction from observations. While Descartes doubts the accuracy of information provided by senses, Bacon stresses the many intellectual obfuscations that hamper the mind (his famous idols of the Tribe, the Cave, the Marketplace and the Theatre). The result is a certain fracture between two reasoning methods, deduction and induction, that should cooperate to reach the truth instead of competing as enemies to demonstrate which one is better.

Despising induction as a means to discover the laws of nature is more a doctrine of those who clung inflexibly to Aristotle’s teachings than of Aristotle himself. Certainly, according to Aristotle deduction is superior to induction, but both play a role in scientific inquiry. However, what Aristotle did reject was the use of mathematical reasoning in other sciences different from mathematics: his arguments to find the natural causes of phenomena are purely qualitative, what makes his physics much poorer than his logic. The merit to combine rational thinking, observation-experimentation, quantitative measurements and mathematical demonstrations corresponds mainly to Galileo Galilei, who is attributed with the saying: “Measure what can be measured, and make measurable what is not so” (Weyl, 1959). This is perhaps the most audacious, important and innovative step taken by Galileo in terms of scientific method, since the usefulness of mathematics in obtaining scientific results was far from obvious at that time, because mathematics did not lend itself to the discovery of causes (which was the primary goal of Aristotelian science) (Feldhay, 1998).

Moreover, being the founder of the experimental scientific method, it is worth noting that Galileo did not disregard theoretical thinking in favor of empirical proofs. In fact, one of his most famous demonstrations to disprove Aristotle’s theory of gravity, which states that objects fall at a speed relative to their mass, had the form of a “thought experiment”. The experiment involves two free falling stones of different weight that are tied together. Galileo explains that, according to Aristotelian physics, the lighter stone should retard the heavier stone in its fall, and vice versa, resulting in an intermediate speed. But the two stones together make a heavier object than either stone apart, so that they should fall faster than the heavier stone alone. From this contradiction Galileo concludes that objects fall at the same speed regardless of their weight. Incidentally, this purely mental proof contained in the First Day of his *Discourses and Mathematical Demonstrations Relating to Two New Sciences* (1638) demonstrates also that showing a logical contradiction in a theory makes it unnecessary to disprove the theory by experiments.

Isaac Newton consolidated the scientific method with an extraordinary development of applied mathematics, and laid the groundwork for most of classical mechanics, whose inductive-deductive approach other sciences sought to emulate. His *Mathematical Principles of Natural Philosophy* (1687), usually called the *Principia*, is probably the most important scientific book ever written. His “rules of reasoning” constitute a re-creation of Galileo’s method that has never been significantly changed, and in its substance is used by scientists today. The so-called hypothetico-deductive method is characterized by successive iterations of four essential elements: observation of phenomena, formulation of explanatory hypothesis, prediction of observable consequences from the hypothesis, and experimentation to confirm or refute the predictions (and thus, indirectly, the hypothesis). To reduce the risk of biased interpretations of results and achieve objectivity, data and methodology must be shared so that they can be carefully examined by other scientists who attempt to reproduce the experiments and verify the results.

**SOME PHILOSOPHICAL ISSUES**

The history of scientific method we have briefly sketched leads us to the consideration of some philosophical problems that remain largely unsolved. The purpose of the following exposition is not to solve them, but to show
that we must not be too naïve in explaining or using the scientific method.

The role of mathematics

The Baconian method intended to discover the cause of a certain phenomenon by finding regularities in observations that linked it to other phenomena, but it was still a fundamentally qualitative method. As we have said, it was Galileo who first took the significant step of giving those regularities the form of mathematical laws (in fact, other contemporary scientists share this merit, significantly Johannes Kepler with his formulation of the laws of planetary motion). Galileo was intimately convinced that the laws of universe are “written in the language of mathematics” (The Assayer, 1623, ch. 6). Galileo, Kepler and others believed that the universe has an underlying rational structure that is within reach of human understanding.

So, the beginning of the scientific revolution is marked by the creation of mathematical models and theories that formalize observed phenomena as measurable variables and link them to each other. Since that time the building of unifying models has been a constant effort in science. Establishing mathematical laws of behavior gives the possibility to make accurate predictions, paradigmatically demonstrated with the many discoveries in the field of astronomy. The confirmation of predictions proves the validity of the theory, and once it has been well established, the theory gives us the possibility to make new predictions we can trust. In this way, mathematical theories lay the ground for engineering all kind of devices with predictable behavior, designed for the well-being of humans.

But mathematics is not an experimental discipline in itself. Mathematics is grounded on axioms and pure reason, and its results do not require experimental verification to be valid. The fact that mathematical concepts are applicable to physical phenomena, even far beyond the context where they were originally developed, has been always a motive for perplexity. The Nobel Prize physicist Eugen Wigner invokes the example of the mathematical law of gravitation, originally used to model freely falling bodies on the surface of the earth, but then extended to describe the motion of the planets, proving accurate beyond all reasonable expectations (Wigner, 1960). He analyzes the obscure and “miraculous” connections between mathematics and physics, concluding that “fundamentally, we do not know why our theories work so well”. The philosopher Hillary Putnam explained this “miracle” as a necessary consequence of a realist view of the philosophy of mathematics (Putnam, 1975). Richard Hamming, one of the founders of computer science honored with the Turing Award, further reflected on Wigner’s ideas and tried to give partial explanations to this “unreasonable effectiveness” (Hamming, 1980). His tentative explanations ranged from the consideration of mathematics as a human creation to fit the observed reality, to the biological evolution as having primed mathematical thinking in humans. Hamming himself concluded that his own explanations were unsatisfactory as well. The debate is still very lively (Tegmark, 2008).

Regularity and causality

The emphasis on building mathematical models of predictable behavior leaves open the following question: do we reach the causes of phenomena with mathematics? The mere establishing of a regularity in the connection of phenomena is insufficient to state that one phenomenon is the cause of another phenomenon (Pearl, 2000). This is well known in the scientific method. Even though a mathematical law suffices to make practical predictions, a serious researcher will not be content with the formulation of a so-called “empirical law”: a law that lacks an underlying theoretical model, a mere distillation of the results of repeated observations. In this sense, the Baconian method (discovering causes by agreement, difference and concomitant variation) is an oversimplification of scientific inquiry.

Of course, science must distinguish between accidental conjunction and true statistical correlation of phenomena. An accidental conjunction occurs when two events happen at the same time, without having a direct relationship to each other besides the fact that they are coincident in time. For example, I open the front door of my house and it starts raining. Well designed experiments and proper use of statistical tools are essential to find true correlations between variables. The first step to find a true correlation, then, is to increase the sample size of events: I open the front door many times and in different situations, and find that there is no correlation between the kinds of events “opening the front door” and “starting to rain” (it would be very curious indeed, and deserving further investigation, to find a strong correlation in this case!).

But, even if a true correlation is discovered, inferring a cause-effect relationship would be a premature
conclusion (Simon, 1954; Holland, 1986; Aldrich, 1995). In other words, correlation does not imply causation. The opposite belief, correlation proves causation, is a logical fallacy known as “cum hoc ergo propter hoc” (Latin for “with this, therefore because of this”). A similar fallacy is “post hoc ergo propter hoc” (“after this, therefore because of this”), which is even more tempting because temporal sequence appears to be integral to causality. In general, a correlation between two variables A and B can be explained in different ways: A causes B, B causes A (reverse causation), some unknown factor C actually causes A and B (the so-called spurious relationship), or even a combination of the former (as in self-reinforced systems with bidirectional causal relationships).

Consider the crowing of the rooster (RC) and the rising of the sun (SR). A naïve conclusion, apparently supported by time sequence, is that RC causes SR. But a careful analysis of correlation and the use of counterfactual experiments lead easily to the conclusion that the independent variable is SR and the dependent variable is RC: the sun rises anyway, the rooster crows when dawn is approaching. So we can conclude with certainty that the sun’s rising causes the rooster’s crowing.

But, can we? Even a true correlation is not enough to establish a causal relationship. Some further explanation is required by the scientific method, a rational explanation rooted in a theoretical model that unifies concepts and observations. Something like “the rising of the sun increases the intensity of light, this awakens the rooster, that then crows”. This deeper understanding of the causal relationship is beyond the results of mathematics, beyond pure statistical correlation (in this sense we can say that Aristotle’s insight that mathematics does not lead to the discovery of causes was right). What the scientific method states is that correlation is a necessary but not sufficient condition for causation. Certainly, correlation suggests causation, and ignoring it would be unintelligent, but correlation in itself is not enough. A scientific explanation is a causal explanation that requires reasoning beyond pure mathematical regularity.

The philosopher David Hume went a step further when he criticized the idea of causality in An Enquiry Concerning Human Understanding (1748) and other works, according to his view that we know only what we perceive. He claimed that causality is purely a mental association between constantly conjoined events that cannot be inferred from experience. The causal connection is in our minds, and we cannot tell from experience whether it really exists. In a few words, causality is beyond experience, only correlation can actually be perceived (i.e. known). Hume strongly influenced Immanuel Kant, who further developed the critique of causality and metaphysics in general. The influence of both in modern philosophy of science is unquestionable. In a certain sense, it can be said that the metaphysical idea of causality (one that is beyond physical experience, as the word “metaphysical” denotes) has been abandoned in modern science. Even the theoretical models that “explain” the connections between observed phenomena seem to be satisfied with purely mechanical explanations, some kind of very elaborated systems of correlations between variables. In some fields of experimental science, especially in physics, causality is hardly mentioned. Nevertheless, most non-philosopher scientists will accept, against Hume and Kant, that knowing causality (at least in the weak non-metaphysical sense) is within reach of human reason.

The origin of hypotheses

The scientific search for cause-effect relationships, or at least for true correlations, begins with the formulation of a hypothesis. A hypothesis is a fact or theory that, if it were true, would explain the observed phenomenon. Following the hypothetico-deductive method, the hypothesis has to be confirmed or refuted by experiments. In fact, experiments specifically designed to refute a given hypothesis have a great value in the scientific method: if they succeed, then the hypothesis is discarded; if they fail to refute, then they provide a stronger confirmation of the hypothesis. But where does the hypothesis come from? Is there any logic in the origin of a hypothesis?

The common answer is: scientific hypotheses are the product of the process of induction. But this is obviously too simple. If induction is the process of formulating general rules on the basis of particular cases, then induction cannot provide a causal explanation that goes beyond simple generalization. The cause is not simply a generalization of the effects, therefore inferring the cause from the effects cannot be the product of induction (Génova, 1997). From the facts that the rooster has crowned today at dawn, and yesterday, and the day before yesterday… induction (generalization) can conclude that “the rooster crowed everyday at dawn”, but this is not a causal explanation. In fact, it is not an explanation at all, it is a mere “regularization”: making individual occurrences to be instances of a general rule. The explanation that “the rooster crows because it is awakened by
an increased intensity of light produced by the rising of the sun” is a very different kind of explanation that goes well further beyond generalization of the observed phenomena. It is an explanation that causally relates two different kinds of events. “By induction, we conclude that facts, similar to observed facts, are true in cases not examined. By hypothesis, we conclude the existence of a fact quite different from anything observed, from which, according to known laws, something observed would necessarily result. The former, is reasoning from particulars to the general law; the latter, from effect to cause. The former classifies, the latter explains” (Peirce, 1877).

Western philosophy, rooted in Aristotelian logic, has traditionally considered that there are two basic kinds of reasoning: 

- **deduction** is a kind of argument that shows that a conclusion necessarily follows from a set of premises; 
- **induction**, instead, draws a generalized but not necessary conclusion from a finite collection of specific observations. But Aristotle, in fact, acknowledged in his *Organon* a third kind of argument he called “backward reasoning”, different from induction. These two latter kinds of arguments, induction and backward reasoning, had been conflated through the centuries until the philosopher Charles S. Peirce recovered the distinction for modern logic (Peirce, 1867). Peirce called this backward reasoning abduction or retroduction: it is the logical process by which new ideas, explanatory hypotheses and scientific theories are engendered.

The main trend in modern philosophy of science, following Karl R. Popper, has ignored the logical problem of the origin of hypotheses (Hanson, 1958). The scientific method begins once a hypothesis is at hand to be tested by experiments, but the origin of the new ideas is not an issue that can be explained in logical terms (Popper, 1959). The act of conceiving or inventing a new theory is a kind of blind conjecture or guess, the fruit of chance or intuition. From this point of view, the discovery of new ideas can only be studied from a historical, psychological or sociological perspective, but it is not important for the rational description of scientific knowledge. New ideas are there, and that is all that matters.

By contrast, according to Peirce abductive reasoning provides a likely account of the facts that need explanation, therefore *it is a logical operation of the mind*, not a mere blind conjecture (Peirce, 1901). Of course, abduction (like induction) is not kind of argument that yields necessary conclusions: it is fallible, even extremely fallible, it is not a direct intuition of the laws of nature (as Cartesian rationalism would like); it is fallible, but rational. In its search for an explanatory hypothesis, abduction is deliberate and critical, which are elements of rational thinking (Ayim, 1982). Peirce is concerned with a notion of logic as a “theory of human reasoning” that is broader than pure formal logic. Therefore, Peirce considers that logic has to study not only formalizable kinds of arguments (deductive, necessary syllogisms), but also other kinds of arguments that are essential to human reason and progress in knowledge (fallible induction and abduction).

These various modes of reasoning are integrated in Peirce’s description of scientific method: abduction invents or proposes an explanatory hypothesis for observed facts; deduction predicts from the hypothesis testable consequences that should be observed; induction verifies the hypothesis by means of experiments, that is, the observation of particular cases that agree with the proposed hypothesis and thus confirm it (Peirce, 1901). In this sense, *induction does not provide any new ideas*, it merely corroborates or refutes the abductive conjecture. The role of generating new ideas via hypothesis or conjecture corresponds solely to abduction (Fann, 1970). Even an inductive generalization requires some kind of previous, perhaps unconscious, abduction: when the scientist concentrates on a certain set of facts in search for a general law, he or she has already made some kind of conjecture about the kind of phenomena that may be subject to a generalization. In trying to formulate the laws of movement, Galileo and Newton discarded from the start qualities such as the color, the smell or the source of moving bodies: in a not yet fully specified form, only their mass is considered relevant, which is already a kind of abduction. Perhaps the material composition of the bodies (wood, lead, stone, etc.) can be also abductively considered, to be later discarded by induction from experiments.

Therefore, the enumeration of phenomena reveals the crucial role of abduction as a preparatory step for induction. What do we enumerate? Why this enumeration, why not a different one? In order to enumerate something, we need to know already in some way what we want to enumerate. The general concept that should be derived by induction from particular cases has to be previously known, even though known only vaguely, to be able to enumerate those particular cases. And this is just what abduction does: it provides, via hypothesis or conjecture, the clue to the general concept the scientist has to follow to identify and enumerate singular data. Induction by enumeration is not enough to explain the formation of general concepts, since the enumeration of relevant phenomena requires a previous abduction to decide which are the relevant phenomena (Génova, 1997).
The strength and weakness of induction

Radical empiricism defends the idea that a work only deserves to be qualified as “scientific” if it is supported by “empirical evidence”. Indeed, it is very easy to criticize this thesis: the idea that “only those propositions that are obtained through experience are scientific, and thus acceptable as true”, is not supported itself by any kind of empirical evidence. Therefore, radical empiricism must be rejected as self-contradictory. Of course, we do not want to deny the extraordinarily important role that empirical evidence has in science. We only want to show the limits of obtaining knowledge through induction from experience. (Do not confuse with “mathematical induction”, which despite its name is a form of rigorous deductive reasoning.)

The big problem of induction is to determine whether it truly has a rational foundation, since the mere fact that particular experiences are repeated does not warrant the positing of a general law, as the critics of inducivism have constantly remarked since ancient times. In his Enquiry Concerning Human Understanding (1748), Hume argued that it is impossible to justify inductive reasoning: it certainly cannot be justified deductively, and it cannot be justified inductively (from the success of past uses of induction) since it would be a circular justification. Nonetheless, he continued, we perform it and improve from it. It has no rational justification, but it is rooted in instinctual habits; it is unreliable, but we have to rely on it.

During the 20th century two main philosophical stances have dealt with the problem of induction: Verificationism and Falsificationism. However the critics, Verificationism upholds an optimistic thesis: induction is possible. This optimism provides the ground for the most generalized attitude among scientists, which precisely leads them to seek the confirmation of their theories in experience. Verificationism admits a priori that regularities cannot be casual: there must be some kind of rationality in the universe, which human reason can discover. Bertrand Russell represents the modern sarcastic criticism to this view with his story of the “inductive chicken”, which after months of repeated experiences (most regular, indeed) came to the firm conclusion that the man who fed it every morning in the farmyard would continue to do so until the end of times, with all his affection... (Russell, 1997, p. 36): “It must be conceded, to begin with, that the fact that two things have been found often together and never apart does not, by itself, suffice to prove demonstratively that they will be found together in the next case we examine. The most we can hope is that the oftener things are found together, the more probable becomes that they will be found together another time, and that, if they have been found together often enough, the probability will amount almost to certainty. It can never quite reach certainty, because we know that in spite of frequent repetitions there sometimes is a failure at the last, as in the case of the chicken whose neck is wrung. Thus probability is all we ought to seek.”

Falsificationism, as set forth mainly in the writings of Karl Popper, considers also in a rather pessimistic way, together with the critics of verificationism, that induction is not possible (Popper, 1959; Popper, 1963). He argued that induction is not a part of scientific procedure, and that inference based on many observations is a myth. We cannot aspire to prove the truth of any scientific theory. Scientific hypotheses are no more than mere conjectures that are provisionally accepted until a new experience appears to refute them (what Popper calls “falsification”). This stance is informed by a commendable skepticism which has helped to give it credit among scientists, too. But the truth is that, if taken to its ultimate consequences (beyond the point Popper himself would have taken it), Falsificationism becomes absurd: scientists do not devote themselves to formulating and provisionally accepting whatever theory, and then to looking for counterexamples that refute it. On the contrary, scientists strive to verify hypotheses as much as to refute them, and they only accept hypotheses which are reasonable from the start and that have a huge explanatory power (Génova, 2010).

Our point here is that neither Verificationism nor Falsificationism can give a full account of scientific activity without referring to something that is beyond factual experience. We can say that both are right in what they deny, but they are wrong in what they affirm. Verificationism is right in saying that hypotheses must be experimentally verified, but it is wrong in claiming that induction from experience can reach scientific truth with absolute certainty. Falsificationism is right in saying that induction cannot be formally justified, but it is wrong in claiming that science essentially tries to refute theories. The reality of science is that induction is constantly used (and required by the scientific community) to validate theories, even if it lacks a formal justification and yields only “probable” results. The progress of science depends on principles that do not arise solely from formally verified experience. The limitation of reason to the empirically verifiable is more a hindrance than a help in the way of science.

Now, leaving apart the philosophical problem of induction, we still have the practical problem of determining what counts as a valid experiment that verifies a given theory. How many trials are enough? What percentage of
confirming results should be required to accept a theory? The absence of *a priori* answers to these questions clearly implies the impossibility to fully formalize the scientific method. In fact, the answer has to be found in the “public”, “social” character of science. This does not mean, far from it, that scientific truth is established by consensus, but that research results must be demonstrable to others: science is not a private affair. The sociological and subjective aspects of science have been highlighted after the failure of previous formalistic descriptions of scientific activity (Polanyi, 1958; Kuhn, 1962; Feyerabend, 1975). What the scientist looks for is to follow a way towards knowledge that can be followed by other researchers; the goal is to “convince” the scientific community of the validity of certain research results. This implies that, besides having empirical support, scientific works must be presented with adequate reasons and interpretation of results. So, how many trials? As many as the scientific community reasonably requires…

**THE SCIENTIFIC METHOD IN SOFTWARE ENGINEERING**

**Plurality of research methods**

The IEEE Computer Society defines software engineering as “the application of a systematic, disciplined, quantifiable approach to the development, operation, and maintenance of software; that is, the application of engineering to software” (IEEE, 2010). The application of the scientific method to research in this field has to consider both the products and the processes of engineering. The *products* of software engineering include all pieces of information that are created by developers in the course of building a software application, the “complete set of software and documentation” (IEEE, 2010): not only the running application and the source code, of course, but also other kinds of documents that are essential to develop, operate and maintain the application, such as requirements and design specifications, models, cost estimations, correctness tests, etc. The *processes* include all different kinds of procedures followed by software professionals to reach the final goal of building a software application, “a collection of steps taking place in a prescribed manner and leading to an objective” (IEEE, 2010).

The purpose of software engineering research methods is to improve both software products and processes. But these two aspects have fundamentally different properties that cannot be ignored when choosing an adequate research method. The products of software are information pieces that are essentially immaterial, certainly requiring some kind of material basis to exist (the hardware), but they are inanimate entities without free will in the end. On the contrary, the processes of software involve humans that can reflect on themselves and change their behavior accordingly. In this sense, the scientific method applied to software products shares principles with natural sciences such as physics, chemistry or biology, whilst the consideration of human factors in software processes demands a method closer to social sciences such as economics, psychology or pedagogy. The laws of behavior discovered for software products will be, like in the natural sciences, deterministic or non-deterministic. The source of indeterminism can be an inherent characteristic of the product, even intentionally wanted, or else the multiplicity of poorly known influencing variables in complex systems (improper indeterminism). It is indeed possible to discover probabilistic laws of behavior also for large human populations, but we cannot forget that non-deterministic causation means something completely different in social sciences: the subjects of research are human beings that can change their behavior through cognition and will.

We must also consider that software engineering is rooted both in *formal and empirical methods*. The origin of computer science is mathematics, and a large part of the research performed in the field of computers and their applications follows a more or less formal method, i.e. definitions, axioms and proof of theorems. A formal system is a human creation, a closed, idealized world with its own laws of behavior, like a game and its rules. In a formal system, be it about modeling languages, cryptographic algorithms or neural networks, the laws are deduced *a priori* from principles axiomatically established; the conclusions are drawn deductively and, properly speaking, there is no need for observation and experimentation of phenomena; the behavior principles are perfectly known because they have been created by humans. Whether a formal system has anything to do with reality is important for its usefulness and applicability, but does not affect its inner consistency. In contrast, the behavior of a real system, be it composed of machines, humans, or both, usually depends on factors that are *discovered* rather than established by human will. Therefore, observation and experimentation are essential to reach inductive knowledge of the laws of behavior of a real system.

Is there a place for experimentation with formal systems? Generally speaking, a formal system (sometimes
called a “model” in the context of software engineering) can be used to analyze an existing real system, or to design a real system to be built (Génova et al., 2009). Indeed, a formal system can be used to build a real system (for example, a computer running a certain algorithm) whose behavior can be observed so as to draw empirical conclusions. If the experiments contradict the predictions, it can be due to an incorrect application of the formal method, or to the fact that the formal system does not adequately represent the real system. Consider a set of dice, each one having equal probability for all faces, implemented as ivory cubes or as computer programs. One can mathematically demonstrate the probability of obtaining a particular result when throwing the dice. If repeated experiments contradict the calculated probability, then either the demonstration was faulty, or else the physical dice are bad implementations of the ideal dice. Obviously, the empirical conclusions are not so strong as the results of (correct) formal proofs, but when the formal method is difficult to follow strictly, the empirical method can greatly help to understand the formal system. A good example is artificial intelligence, where the final behavior of the system is difficult to derive only from its established principles (Cohen, 1995). Empirical conclusions can even encourage finding a formal proof of the obtained results.

In the last decades of the 20th century a growing conviction consolidated: the scientific method developed for studying and analyzing natural phenomena was not apt to understand the design and construction of human artifacts. The required method should not start with the observation of phenomena, but rather with the identification of a need, followed by artifact construction and evaluation (Hevner et al., 2004). This emerging field of construction-oriented research was called design science, the scientific study of design, and it was based on two assumptions: first, the design of artifacts can be a sophisticated task that contributes to the development of scientific knowledge; second, the scientific design of artifacts requires a specific research method (Frank, 2006). Note that the concept of “artifact” encompasses not only physical devices, but also conceptual and social systems: information structures, knowledge representations, methods, processes, organizations, etc. Probably the most prominent proponent of establishing research disciplines that do not follow the classical scientific model was Herbert Simon, who popularized the concepts of design sciences, as opposed to natural sciences, in his seminal work The sciences of the artificial (Simon, 1969). Natural sciences are aimed at “truth”, i.e. at the exploration and validation of generic cause-effect relationships, whilst design sciences are aimed at “utility”, i.e. at the construction and evaluation of generic means-ends relations (Winter, 2008). Design sciences are grounded on natural sciences: valid cause-effect relationships give the possibility to build useful means-ends relations. Besides, design sciences pose ethical issues about what ends are desirable and what means are legitimate and proportionate to achieve those ends, which demands a solid ethical education (Génova et al., 2007). Design science research is currently an expanding field that is on its way to achieve the same standards of rigor and relevance that is currently possessed by natural sciences.

Classification of research methods

In the previous reflections we have considered three different dimensions of research in software engineering: (i) the product-process dimension; (ii) the formal-empirical dimension; and (iii) the analysis-design dimension. Some authors have provided useful classifications of research methods that implicitly or explicitly consider these dimensions. We summarize here two of these classifications.

Hanenberg, who is strongly concerned by the missing consideration of human factors in the justification of new artifacts that supposedly improve the software development process (such as programming language constructs), classifies research methods in two main, partially overlapping categories: technical approaches and empirical approaches (Hanenberg, 2010). Technical approaches are adequate whenever the software developer or user does not play any role in the subject of research, that is, the subject of research is a formal system or a physical machine. They include:

- Pure formal methods that achieve deterministic results (the “classical approach”) or non-deterministic results (the “stochastic-mathematical approach”). Examples are research on inherent properties of deterministic algorithms and data structures, and research on parallel computing and randomized algorithms.

- Empirical methods that achieve statements using statistical methods applied on measurements resulting from experiments (the “stochastic-experimental approach” and the “benchmark-based approach”). An example is the research on software performance.

Empirical approaches include, besides the technical empirical methods, other approaches where the consideration of the role of the developer or user is essential (the “socio-technical approach”), since the research
is performed to study how *developers* can write better software (i.e. software that has fewer errors, is more maintainable or is more reusable), or how *users* can interact better with software systems. Examples are the invention of new programming paradigms such as object orientation or aspect orientation, the design of graphical user interfaces that bring an improvement from the cognitive point of view, etc. The formal-empirical dimension is predominant in this classification, even though the product-process dimension is also implicit: the roles of users and developers are crucial to respectively assess the quality of products and processes.

Hanenberg, following Tichy (Tichy, 1998), criticizes the lack of empirical validation in a large part of the research performed in the design of new development artifacts, especially when compared with the extensive and effective use of the socio-technical approach in the field of Human-Computer-Interaction (HCI) (Shneiderman & Plaisant, 2009). Too often only qualitative criteria are provided to justify the advantages and adequacy of a new language construct, without any empirical evidence that developers really become more productive by learning and using it. Therefore, Hanenberg claims, the socio-technical approach is currently burdened with an excess of speculative reasoning, while it should be enriched with the kind of empirical methods that are actually used with human subjects in HCI and other sciences like psychology.

Mora et al. are more favorable to conceptual methods, against a biased negative view of non-empirical methods (Mora et al., 2008). They perform a literature review of methodological taxonomies in information systems research and offer their own framework to classify research methods. This framework is structured by two dimensions: “conceptual vs. reality” and “natural/behavioral vs. purposeful design”. These two dimensions permit to classify research methods into four quadrants: the conceptual behavioral research (CB), the conceptual design research (CD), the empirical behavioral research (EB) and the empirical design research (ED). They further identify different activities that are typically performed in each quadrant, and provide examples, published in top scientific journals, of researches performed according to the four kinds of methods:

- **CB** – Conceptual analysis of behavior in existing knowledge management systems, in a tutorial and descriptive way.
- **CD** – Conceptual design of decision support systems, building predictive models of the effects of advanced information technologies in business intelligence capabilities.
- **EB** – Empirical analysis and validation of behavior and plausible causal links between different constructs in knowledge management systems.
- **ED** – Empirical design, construction and validation of a decision support system to select an investment portfolio.

The first dimension in this framework roughly corresponds to our previously identified “formal-empirical” dimension, except that, for Mora et al., conceptual methods encompass both formal and qualitative reasoning. The second dimension is also similar to our “analysis-design” dimension: the natural/behavioral method pursues a theoretical understanding (i.e. analysis) of the behavior of a given entity, be it natural, social or artificial, conceptual or real, without modifying it; the opposed method purposefully intends the design of a new artifact or the modification of an existing one. Mora et al. do not explicitly consider the “product-process” dimension.

Research methods can also be classified, from a different perspective, according to basic beliefs about scientific inquiry: ontology–what is reality, epistemology–how we know reality, and methodology–how we plan the research accordingly (Guba & Lincoln, 1994). The possible answers to these questions classify the methods into two main alternative inquiry paradigms, either positivist (reality is something given) or interpretivist (reality is something we construe), each one having several subtypes and including a variety of methods.

Summing up, the variety of research topics and objectives of information systems research cannot be covered in a satisfactory way by one method (Frank, 2006), and the purpose of a research project (exploratory, explanatory, descriptive, predictive) should be used to identify the most adequate research strategy in each case (Marshall & Rossman, 1995).

The role of qualitative reasoning

Formal methods provide strong deductive tools that greatly improve the power of human reason. A rigorous mathematical or logical demonstration offers certainty in the correctness of the result, and it can be checked by anyone who masters the technique. This publicity of the method is essential for science understood as a social enterprise. But the results of formal methods are valid only within the formal system, they depend in the end of the free choice of axioms. Formal methods suffer the risk of being correct, but irrelevant: formal correctness is not truth. How can we ultimately assure that the formal system is a good representation of a real system, that it is
useful or applicable? The adequacy to the real world is simply beyond the capability (and purpose) of the formal method, it has to be judged from outside the method. Yet it is good that it is so: a formal method would cease to be formal, losing its main strength, if it tried to answer this question.

Some will claim that empirical methods are the answer. Modern science is inconceivable without them. The repeatability of experiments fulfills the requirement of publicity and, indeed, the combination of formal and empirical methods is the distinctive mark of modern science: building theoretical models and contrasting their predictions against observation of reality. Accepting that empirical methods achieve a lower level of certainty than formal methods (remember the problem of induction), they assure instead the correspondence between theoretical models and reality. But, again, empirical methods are limited. We have already noted that justifying the reasonability of a hypothesis prior to designing experiments, judging on causality, and choosing a valid experiment to confirm a theory, are all activities that lie beyond the empirical method in and of itself. Empirical evidence must be adequately interpreted with good reasons to recognize what an experiment really demonstrates.

So, neither formal nor empirical methods can provide a full account of scientific activity, which relies on something that is beyond axioms and factual experience, beyond formal proofs and rigorous statistics. In other words, scientists have to reason also outside the scientific method, which is especially manifested when they are choosing the most adequate method for a certain research project, and when they are interpreting the results (Kitchenham, 1996). Even if empirical methods are useful in some fields such as HCI, qualitative methods involving heuristics and human judgment are also required (Nielsen & Mack, 1994). Logical evidence and empirical evidence require a work of clarification and development of concepts. This kind of meta-methodical reasoning, which we call qualitative reasoning, is not anecdotal; on the contrary, it is the very basis onto which the scientific method is developed, however weak be that basis. Reason uses the powerful tools of formal and empirical methods, but reason is not confined within the limits of any established method. Qualitative (i.e. meta-methodical) reasoning plays the directive role in scientific activity. Theorems and experiments without the guide of speculative thinking are worthless.

**Improving the methods**

We think software engineering needs to improve the application of the scientific method in two main directions: empirical and conceptual/qualitative methods (formal methods are strong and widely used, so we think they do not need such an encouragement at present). As Hanenberg, Tichy and others claim, better empirical methods have to be applied to verify the advantages of a new invention for the software development process, with a special consideration of human factors. Too often a new software artifact, embedded in a certain theory of the software process, is presented with a tool that implements it and shows its applicability (think of software modeling languages “demonstrated” with accompanying software CASE tools). But we should be careful to distinguish between experimentation of a theory and its practical application: the latter is particularly important in engineering, but developing a practical application does not properly constitute an experimental verification, according to inductive criteria, of the theory that supports it. The tool can demonstrate that the artifact and the theory supporting it are applicable, but it does not demonstrate that current practices in software development will be improved, i.e. that software developers will become more productive.

Tichy enumerates eight fallacies that prevent the widespread adoption of empirical methods in software engineering (Tichy, 1998): the current level of experimentation is already good enough, experiments cost too much, tool demonstrations are enough, etc. He reports that, in a random sample of papers published in ACM software-related journals before 1995, around 50% of those with claims that needed empirical support had none at all, contrasting the fraction of 15% found in other sciences. The “socio-technical” approach, then, should be fostered in software engineering research (Juristo & Moreno, 2001). However, this approach faces some obvious difficulties, as does any kind of empirical approach where the subject of research is a human population. This is not an excuse not to perform experiments, but a warning against a too naïve design and interpretation of experiments. Empirical software engineering has to learn its methods from social sciences (a good example of success is HCI). The design of controlled experiments is particularly difficult because every human person has a different culture, education and environment that clearly will influence his or her reaction to an experiment. People are not atoms or falling bodies that behave always the same: they learn from experience, and they can change their behavior.

The repeatability of experiments, then, is compromised. Can we really repeat an experiment with human subjects? Can we change the value of one single variable in the initial conditions of the experiment, everything
else being equal, so as to design a counterfactual experiment and draw valid empirical conclusions? We cannot rewind history and pretend the subjects are in the same state as they were before the first try. Another particularity of human subjects is that they behave in a different way as usual when they know that they are the subjects of an experiment (a phenomenon known as the “Hawthorne effect”) (Kitchenham, 1996; Katzer et al., 1998). To minimize this effect, the subjects should be unaware of the experiment. Yet this has ethical implications: we cannot experiment with humans without their consent and collaboration. The only way to escape from these limitations is to increase the population that is the subject of research, making it as heterogeneous as possible. Experiments performed on a group of students from a single university or even from a single class are of very limited value. The ideal would be extending the experiments to many software professionals from different countries and organizations. A recent and promising approach based on the Mechanical Turk platform (http://www.mturk.com) could help to overcome some of these difficulties with the potential to conduct experiments on large and varied populations (Alonso & Lease, 2011).

In any case, the researcher must be conscious that statistical conclusions drawn from a large population will have very limited predictive value in particular cases. Even if it were empirically demonstrated, say, that object orientation makes developers across the world more productive than structured programming, which is a controversial statement with defenders (Haythorn, 1994; Juristo & Moreno, 2001) and opponents (Hatton, 1998), we are not guaranteed that introducing object orientation in a particular organization would be beneficial. Developers (and so instructors and students) are intelligent beings that need to be re-educated to assume a new programming paradigm or whatever new software artifact that is presumed to improve the development process. Empirical arguments do not suffice, developers need to be rationally persuaded of the benefits, often with qualitative rather than statistical arguments. Therefore, improving conceptual methods, both formal and qualitative, is also essential for the advance of software engineering (in this respect, qualitative reasoning is “methodical”, in the sense that it has to be rigorous, but it is also “meta-methodical” in the sense that it cannot be subjected to the predefined set of rules of an established formal or empirical method).

There is a widespread negative view of non-empirical approaches that considers conceptual research is based on speculations rather than systematic data collecting procedures (Mora et al., 2008). The term “speculative thinking” is often referred to something that has no scientific value in itself, as if it were a game of free dancing ideas deprived of any rational significance. Surely, this negative view is supported by the frequent abuse of non-rigorous speculative thinking. Of course, conceptual, qualitative or speculative thinking cannot be reduced to such weak justifications as asserting subjective beliefs (“this approach is good because I like it”, “it seems intuitively obvious”, etc.). Taste and intuition are not objective enough to be shared by the scientific community: again, science is a public, not private, affair. On the contrary, qualitative reasoning must be grounded on a sound rational basis. Analytical rigor is difficult, but not impossible. The arguments must be presented in a clear, concise and objective way, pros and cons must be analyzed, opposing views must be carefully considered. But, in the end, qualitative thinking will remain a kind of non-empirical reasoning that requires a particular training to achieve evidence (non-empirical evidence) with the desirable degree of rigor in science.

CONCLUSION

Ultimately, we are seeking an appropriate balance between empirical methods and qualitative reasoning. They should not be considered exclusive ways, they are not enemies but cooperators on the road to scientific truth. Qualitative thinking is satisfactory to describe a radically new idea or a significant breakthrough in science and engineering, but it has to be complemented with validating experiments in subsequent works (Tichy, 1998). Qualitative or speculative thinking is also required to clarify concepts and to perform essential activities in science that cannot be closed within any established method, such as judging the adequacy between a theoretical model and the reality it represents, inferring causal relationships between observed phenomena, proposing reasonable explanatory hypotheses, deciding which experiments are necessary and interpreting their results. The practice of the scientific method requires to acknowledge (and overcome) its own limits; improving the scientific method is something that cannot be done from inside the method itself. Therefore, acknowledging the limits of formal and empirical methods, and opening the door to meta-methodical reasoning, is a must for software engineering. The new degree on Computer Science and Philosophy recently introduced in the University of Oxford demonstrates the interest to reach an understanding (University of Oxford, 2011).

Albert Einstein is often quoted as having said: “Not everything that can be counted counts, and not everything that counts can be counted”. In fact the words must be credited to sociologist William Bruce Cameron.
We do not intend to minimize the value and importance of empirical methods, when they are required. But we claim that empiricism is insufficient. There cannot be a complete scientific activity that consists solely of proving theories by means of experiments: first, theories must be formulated and developed, and their explanatory power must be demonstrated, so that the investment of human and material resources in the experiments, which may be very costly, can be justified; then, the experiments that will prove or refute the theories must be carried out. Moreover, experimental verification may say something about the truth of a theory, but it can say nothing about its relevance, i.e. its interest to the scientific community or society as a whole.

Not all branches of science are equal, not all kinds of research are equal. Experience and speculation must go hand in hand in the way of science. Some investigations will have a basically experimental character, while others will be primarily speculative, with a wide gradation between these two extremes. As long as all are demonstrable, we should not consider some to be more worthy of respect than others. If we are closed from the start into a radical empiricism that considers only empirical evidence is a valid support for a scientific work, then we will be unable to perceive the value of other kinds of thinking that historically have proved to be indispensable for the advancement of science (Génova, 2010). If this value is not perceived, rigorous qualitative thinking will not be taught in engineering schools. Since it is unavoidable that qualitative thinking plays the leading role in scientific activity, the result will be an impoverished application of the scientific method.

REFERENCES


**KEY TERMS AND DEFINITIONS**

**Scientific method:** A cyclic, iterative process of observation, hypothesis formulation, and confirmation/refutation of hypothesis through experimentation. Also called hypothetico-deductive method. In design sciences, as opposed to natural sciences, the steps are identification of needs, artifact construction and artifact evaluation.

**Philosophy of science:** The branch of Philosophy that deals with the metaphysical issues arising from the study of scientific method.

**Empiricism:** A philosophical stance that defends the idea that a work only deserves to be qualified as scientific if it is supported by empirical evidence. It is rooted in David Hume’s view that we know only what we perceive. Variants are Verificationism (induction is possible because regularities cannot be casual) and Falsificationism (induction is not possible, scientific theories are always provisional until they are refuted).

**Reasoning:** Inference of a conclusion from a set of premises. There are various modes of human reasoning, mainly deduction, induction and abduction (but also reasoning by analogy, and others).

**Formal method:** A deductive method of reasoning that achieves necessary conclusions from definitions, axioms and proof of theorems.

**Empirical method:** An inductive method of reasoning that achieves fallible conclusions using statistical methods applied on measurements resulting from experiments.
**Socio-technical method:** The empirical method applied in a research project where human beings play an essential role as subjects of experimentation.

**Hypothesis:** A fact or theory that, if it were true, would explain the observed phenomenon. Backward reasoning from effect to cause. Following the hypothetico-deductive method, the hypothesis has to be confirmed or refuted by experiments. Hypotheses are fallibly generated by abductive reasoning.

**Causality:** The relationship between an event (the cause) and a second event (the effect), where the second event is understood as a consequence of the first. Statistical correlation is a necessary but not sufficient condition to determine a causal relationship.

**Qualitative reasoning:** Reasoning outside any established formal or empirical method. It is essential to clarify concepts and to perform common scientific activities such as: justifying the reasonability of a hypothesis prior to designing experiments, determining causal relationships, judging the adequacy between a formal model and the reality it represents, deciding what statements need empirical validation, choosing a valid experiment, interpreting the experimental results, etc. Also called meta-methodical reasoning, speculative reasoning, or conceptual reasoning.

**ABOUT THE AUTHORS**

**Gonzalo Génova** has an MS degree in Telecommunication Engineering from the Polytechnic University of Madrid (1992), an MS degree in Philosophy from the University of Navarre (1996), and a PhD in Computer Engineering from the Carlos III University of Madrid. He is currently an Associate Professor of Software Engineering in the Department of Computer Science and Engineering at the Carlos III University of Madrid. His main research subjects, within the Knowledge Reuse Group and in close cooperation with The Reuse Company, are models and modeling languages in software engineering, requirements engineering, and philosophy of information systems.

**Juan Llorens** received his MS degree in Industrial Engineer from the ICAI Polytechnic School at the UPC University in Madrid in 1986, Spain, and his PhD in Industrial Engineering and Robotics at the Carlos III University of Madrid, Spain in 1996. He joined the Carlos III University in 1992 where he is currently a Full Professor of the Department of Computer Science and Engineering. Dr Llorens is the leader of the Knowledge Reuse Group within the University. His main research subject is information representation and processing for software reuse. Since 1998 he has split his educational activities between Carlos III University and the Högskolan på Åland (Mariehamn, Åland, Finland).

**Jorge Morato** is currently an Associate Professor of Information Science in the Department of Computer Science and Engineering at the Carlos III University of Madrid (Spain). He obtained a PhD in Library Science from the Carlos III University in 1999 on the subject of Knowledge Information Systems and its relationships with linguistics. From 1991-1999, he had grants or contracts from the Spanish National Research Council. His current research activity is centered on text mining, information extraction and pattern recognition, NLP, information retrieval, Web positioning, and Knowledge Organization Systems. He has published mainly on semi-automatic construction of thesauri and ontologies, topic maps, and conceptual and contextualized retrieval of semantic documents.

Address:
{ggenova, llorens, jmorato}@inf.uc3m.es
Universidad Carlos III de Madrid
Av. Universidad, 30
28911 Leganés (Madrid) – Spain