"Towards Reflection Competencies in Intelligent Systems for Science." Yolanda Gil. In "Artificial Intelligence for Science: A Deep Learning Revolution," Alok Choudhary, Geoffrey Fox, and Tony Hey (Eds.). World Scientific, London, UK, 2023.

Towards Reflection Competencies in Intelligent Systems for Science

Yolanda Gil
Information Sciences Institute and Department of Computer Science
University of Southern California
gil@isi.edu

Abstract

This chapter envisions a much more expanded role of AI systems that goes beyond their current use for simply learning from given data using a given metric. To tackle many challenges of our time, scientists will need to partner with AI systems that are capable of independent inquiry, proactive learning, and deliberative reasoning. The chapter describes thoughtful AI systems that can make good partners for scientists and will need to be more collaborative, more resourceful and independent, and more responsible. The chapter also describes six core competencies for AI scientists that will create powerful machines for discovery, and expands on the reflection competency which involves formulating scientific questions, devising general strategies to answer them, executing methods that implement those strategies, and placing new findings in the context of the original questions. A broader and more expanded role of AI systems in the scientific discovery process will in turn enable significant advances in AI, particularly in knowledge representation, reasoning, planning, meta-reasoning, and architectures for intelligence.

1. Introduction: Eighty Years of AI for Science

Scientific discovery has long been of interest to AI researchers. Herbert Simon (Nobel Laureate and Turing award winner) worked on cognitive modeling of scientific discovery as early as the 1940s [1]. Edward Feigenbaum (also a Turing award winner) worked with Joshua Lederberg (himself a Nobel Laureate) and other colleagues in Stanford University in the 1960s on using AI to automate the identification of organic molecules from their mass spectra [2]. Many more AI systems have been developed over the years to address major activities in scientific discovery such as problem formulation, experimentation and data collection, data analysis, machine learning, and model revision [3]. Indeed, a recent cover of *Science* states “artificial intelligence transforms science” [4].

Today, a major focus of AI for science is on machine learning. In the last few decades, advances in machine learning as well as data-intensive computing have pushed the envelope in the nature and scale of the scientific phenomena that can be addressed. Powerful learning paradigms and distributed computation work at unison to process data at scale, leading to spectacular discoveries in diverse areas such as high-energy physics, biomedicine, and geosciences. More recently, new deep learning approaches for machine learning and new intelligent techniques for data mining have given rise to modern data science, combining these powerful data-driven discovery capabilities with scalable computing and data systems that have completely changed how we look at data.

In addition to machine learning, many other areas of AI are making significant contributions to science. Examples include natural language processing to extract knowledge from publications [5], constraint reasoning to search for optimal solutions [6], experiment design and execution [7], and semantic data repositories to facilitate information integration [8].

Future AI systems can contribute much more to science. AI systems for scientific discovery today have very limited scope in the scientific research process, as they are given the learning goals, they are given the data, and they are given optimization metrics. The role of AI systems is limited to solving a well-defined task where the data and techniques are specified by a scientist. Confining intelligent machines to this narrow realm is severely limiting our ability to truly harness the potential of AI to tackle complex science questions. The limited incorporation of intelligent systems in science is also thwarting the pursuit of fundamentally new discoveries particularly at the fringe of current science practice. The increased complexity of the scientific questions that we face is challenging the abilities of human scientists. Imagine a new generation of AI systems that can formulate the learning problems needed to address a given science question, that can find or generate necessary data of appropriate quality, and that can incorporate background knowledge such as theories and scientific principles in order to discern what metrics would be appropriate to assess any new findings. Many scholars have shed light on the diverse and rich cognitive processes involved in scientific reasoning that AI still does not address, from discovering laws [1], to understanding causal mechanisms [9; 10], to collaboration [11], to prioritizing problems [12; 13; 14; 15], to producing paradigm shifts [16].

In this article, we argue that a much more expanded role for AI in scientific discovery will be necessary to tackle many of the challenges of our time. Scientists will need to partner with intelligent systems that are AI scientists capable of doing independent inquiry, proactive learning, and deliberative reasoning. A new generation of AI systems will enable a true partnership between scientists and machines. This partnership will be essential to tackle a new generation of science questions. And expanding the role of intelligent machines in the scientific discovery process will in turn enable significant advances in knowledge representation, reasoning, planning, meta-reasoning, and architectures for intelligence.

2. The Imperative of AI for Science

As scientific questions become more complex and multidisciplinary, the capabilities of scientists to do research will need to be augmented with intelligent machines. Compare the challenges of finding a cure for polio and finding a cure for cancer. Polio, a scourge that has affected humanity for millennia, was cured with a vaccine that was discovered by one scientist. Glioblastoma, a brain cancer that takes very few months to go to advanced stages and is very hard to detect and treat, is being studied by scores of scientists in multiple specialties. Different research groups have complementary information about the disease, with disparate data on genomics, proteomics, transcriptomics, MRIs, treatments, etc. It is a challenge to combine their independent partial findings and synthesize major discoveries. Similarly, compare the challenges of early river hydrology and physics modelling with the challenges of understanding the interacting hydro-bio-agro-human processes in our environment and ecosystems. Scientists with expertise in each of these areas develop complementary models that are very hard to integrate to study the intricate interactions that cut across them. Today's science processes require work that goes beyond what human scientists can do in the face of the complexity and multidisciplinarity of the research questions pursued. There are several key aspects in this matter.

First, keeping up with research innovation is challenging and costly for any scientist. Each discipline advances very quickly, with new sophisticated methods coming out continuously. It is challenging for any given research group to keep up with all the latest methods, so only a few are likely adopted. There is a high cost to understanding what new methods are becoming important and are crucial, learning their nuanced assumptions, and training younger researchers to use sophisticated methods properly. Scientists today do not have intelligent assistance to facilitate fast learning and adoption of new methods.

Second, collaborative research requires very significant effort. It is becoming harder for an individual researcher to tackle the more advanced science questions. Integrating all the information needed is very challenging, particularly across disciplines, and requires partnerships and collaborations. It can take a year of work by a dozen of scientists to integrate climate, hydrology, and agriculture models to understand and forecast food shortages. It takes two years of work by hundreds of scientists to assemble and analyze data to do a global climate study to see trends across the last few decades. It has taken thousands of scientists working for several years to discover the Higgs boson. These efforts require so much coordination (to secure funding, to organize the work, to coordinate responsibilities, to monitor progress, to assemble results, etc) that they are far from being the norm and in some cases can be described as heroic. Scientists today do not have intelligent assistance to support these kinds of collaborative research tasks.

Third, accounting for new data requires continuous updates of prior findings that the science enterprise is not well equipped to do. When cancer data for a new cohort of patients becomes available, previously published studies for similar cohorts should be reconsidered and their findings updated to incorporate the new data. When new environmental data is continuously captured through sensors, or with improved quality with a new type of sensor, previous findings should be reconsidered and extended to account for the new data. When a new and more powerful analysis method comes to light, it should be applied to existing data that was previously analyzed. Each paper in the published literature provides a static snapshot of how scientists would answer a question, but researchers seldom have the resources to revisit published results. Scientists today do not have intelligent tools that automatically reproduce the methods and update the results.

Fourth, significant innovation could result from reimagining research methodologies and processes. There are many aspects of scientific research that could be automated, improved, or redesigned to incorporate capable intelligent systems. This could open the door to advance the frontiers of science in fundamentally new directions by tackling new kinds of questions and creating unconventional approaches. Scientists today do not have intelligent tools that complement and augment their abilities to create innovative changes in research.

In summary, in order to pursue increasingly complex science questions effectively and efficiently we need:

1. AI systems that help scientists adopt new methods quickly so they can keep up with continuous advances in their field
2. Intelligent aids that reduce the effort to integrate knowledge across disciplines so scientists can tackle complex multi-faceted phenomena with reduced effort and therefore more frequently
3. Automated AI systems that can incorporate newly available data into previously published studies in order to continuously update findings so scientists can keep up with all the new data that is being continuously collected

4. AI approaches that can synthesize and innovate scientific research methodologies and processes so scientists have new avenues to address fundamentally new kinds of questions

In turn, these avenues of AI research will lead to a new generation of intelligent systems capable of understanding the scientific research process, knowing the different ways to carry out the steps involved, and able to learn the skills required to keep up with new methods and even create their own. These capabilities will be generally applicable beyond scientific domains, as they will become important tools for humanity to tackle complex problems.

What new avenues of AI research do we need to pursue in order to address these science needs?

3. General Intelligent Capabilities of AI Systems for Science: Thoughtful AI

Intelligent systems with some basic AI capabilities for science will soon become necessary for all scientists. At first, they will be used simply as tools that have no initiative or autonomy. AI systems will soon become assistants to scientists, carrying out tasks that a lab assistant or a research assistant would do. Intelligent systems will at some point become more independent, perhaps making contributions that deserve co-authorship and in some instances writing scientific papers on their own [17]. Eventually, they will become more than assistants and will act as partners to scientists. Some envision AI systems capable of making major scientific discoveries and even winning the Nobel prize [18].

In order to be assistants and partners for scientists, intelligent systems will need to be more collaborative, more resourceful and independent, and more responsible. These kinds of skills can be considered under the umbrella term of *thoughtful AI systems* which will act as is expected of human assistants and partners. Key design principles for such thoughtful AI systems are summarized in Table 1 and discussed in detail in [19].

Table 1. Principles for designing *Thoughtful Artificial Intelligence Systems* (from [19]).

	Principle	*Description*
1	Rationality	Behavior is governed by knowledge
2	Context	Seek to understand the purpose and significance of tasks
3	Initiative	Proactively learn new knowledge relevant to their task
4	Networking	Access external sources of knowledge and capabilities
5	Articulation	Respond with persuasive justifications and arguments
6	Systems	Facilitate integration and collaboration with other systems
7	Ethics	Behavior that conveys scope and limitations

Future AI research in these principles will enable successful approaches for developing intelligent systems that can be partners for scientists.

4. Specific Capabilities of AI Systems for Science: Core Competencies

What kinds of tasks and activities would intelligent systems carry out in pursuing scientific research? Defining core competencies required for scientific research that are shared across scientific disciplines is not an easy endeavor given the diversity and complexity of scientific work. As a starting point, we posit six core competencies needed in intelligent systems for science:

1. *Reflection.* This competency is focused on reasoning about scientific knowledge to formulate questions, identify strategies to pursue them, and situate new findings in the context of what is known. It will push the frontiers in AI in areas such as reasoning, meta-reasoning, and problem solving.

2. *Observation*. This competency focuses on data gathering through laboratory experiments, sensor management, remotely controlled robots and drones, and other science tasks focused on interactions with the physical world or the system under study. It will advance AI research in robotics and cyberphysical systems.

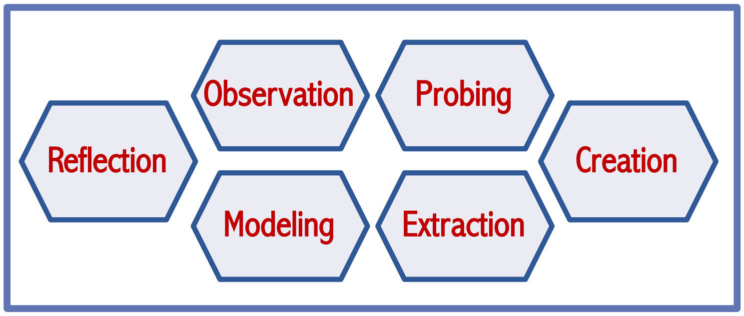
3. *Modeling*. This competency is about the analysis of complex scientific data to uncover patterns and create explanatory and predictive models. It will result in AI innovations in all aspects of machine learning, causality, and uncertainty reasoning.

4. *Probing*. This competency is focused on design and exploration of solutions through efficient search strategies. This competency is important in many science domains that are not so focused on modeling the world but in synthesizing new artifacts, such as drug design and materials discovery. This competency will emphasize AI search, constraint reasoning, and optimization.

5. *Extraction.* This competency is focused on pulling out and integrating information from the literature, online data repositories, and other Web sources. It will result in AI advances in natural language, vision, and information integration.

6. *Creation.* This competency is focused on generating new theories, designing new approaches, constructing new instruments, and other inventions that lead to significant inventions and paradigm changes that open fundamentally new directions in science. It will result in AI advances in creativity, design, and representation shift.

The six core competencies are summarized in Figure 1. Each will advance complementary areas of AI research and will need to be integrated together to create powerful machines for science. Each competency can be explored separately, which will enable the AI community to make significant progress. Most of the work to date on AI for science has focused on the Observation Competency, Modeling Competency, Probing Competency, and Extraction Competency. The Reflection Competency and Creation Competency have not received as much attention. We discuss next some of our prior work on the Reflection Competency.



Competency	Description	AI Research Targets
Reflection	Formulate questions, identify strategies, understand findings	Reasoning and problem solving
Observation	Gather data through laboratory experiments or sensors	Robotics, cyberphysical systems
Modeling	Analyze data to uncover patterns and create predictive models	Machine learning, causality, uncertainty reasoning
Probing	Design and explore solutions through efficient search & optimization	Search, constraint reasoning, optimization
Extraction	Pull out and integrate information from diverse online resources	Natural language, vision, information integration
Creation	Significant inventions and paradigm changes	Creativity, design, representation shift

Figure 1. Six core competencies of AI scientists.

5. Reflection Competency in Intelligent Systems for Science

Intelligent systems for science will need reflection capabilities in order to formulate scientific questions, devise general strategies to answer them, execute methods that implement those strategies, and to place new findings in the context of the original questions. These reflection capabilities are crucial to automatically generating new scientific findings, no matter the question or the domain.

Key research challenges for the Reflection Competency include:

- Representing scientific knowledge to capture questions, hypotheses, and methods, and relating those to one another
- Reasoning about hypotheses, the methods to test them, and the results obtained
- Implementing the scientific processes and steps involved in answering different types of questions, and that can be similar or differ significantly across science domains
- Integrating new findings into current theories and models, detecting inconsistencies, and resolving them with theory revisions or further questions
- Explaining findings and the supporting evidence appropriately, answering follow up questions about the findings in the context of what is already known

The development of a reflection competency will result in significant advances in many areas of AI, including cognitive architectures, knowledge representation, reasoning, planning, meta-reasoning, explanation, question answering, theory revision, and argumentation.

Figure 2 illustrates a proposed conceptual framework for reflection based on six major steps of the scientific research process, based on [3; 20; 21]. Scientific research often starts with an inquiry, which can be a hypothesis that can be tested or simply a question to explore. Next, a scientist will decide what approach to take to pursue the question, in terms of what kinds of data would be needed and how the data would be analyzed. Then there would be a step for gathering data, which may mean carrying out experiments or retrieving data from an existing data repository. Once the data is available, different analysis are carried out for different subsets of the data or with different assumptions. The results of these analyses are consolidated in relation to the original question or hypothesis. Finally, a scientist would reflect on the nature and significance of the findings and revise existing theories or models in their domain.

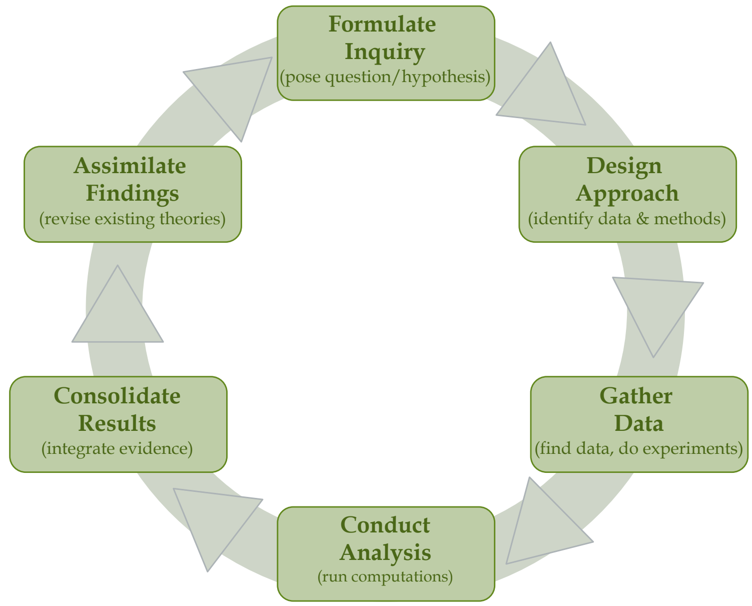


Figure 2. An initial conceptual framework for reflection in the scientific research

This reflection framework focuses on inquiry-driven research, where a question or hypothesis prompts a scientist to gather evidence necessary to answer the questions posed. This is by no means a universal account of all scientific research processes, which may be more exploratory (starting with some data rather than driven by questions), analytical (driven by representation change and redesign of some body of knowledge), instrument focused (design of a new instrument), synthesis (e.g. of new methods or algorithms), etc. This is a high-level process that captures commonalities in many scientific endeavors. This general reflection framework can be fleshed out based on the approaches and processes that are used in different science domains.

Figure 3 illustrates the use of this conceptual framework for reflection with real examples from multi-omics (from [22], neuroscience (from [23]), and flood prediction (from [21]). This highlights the generality of this framework, and its flexibility to adapt to different kinds of inquiries, data types, and analyses. In multi-omics, there is readily available data in shared repositories from many prior studies, and many specialized software tools that can be used for analysis of proteomic and genomic data. Combining the results from these different modalities is an open area of research. In neuroscience, there is data available but specific features of interest have to be extracted from brain image data.

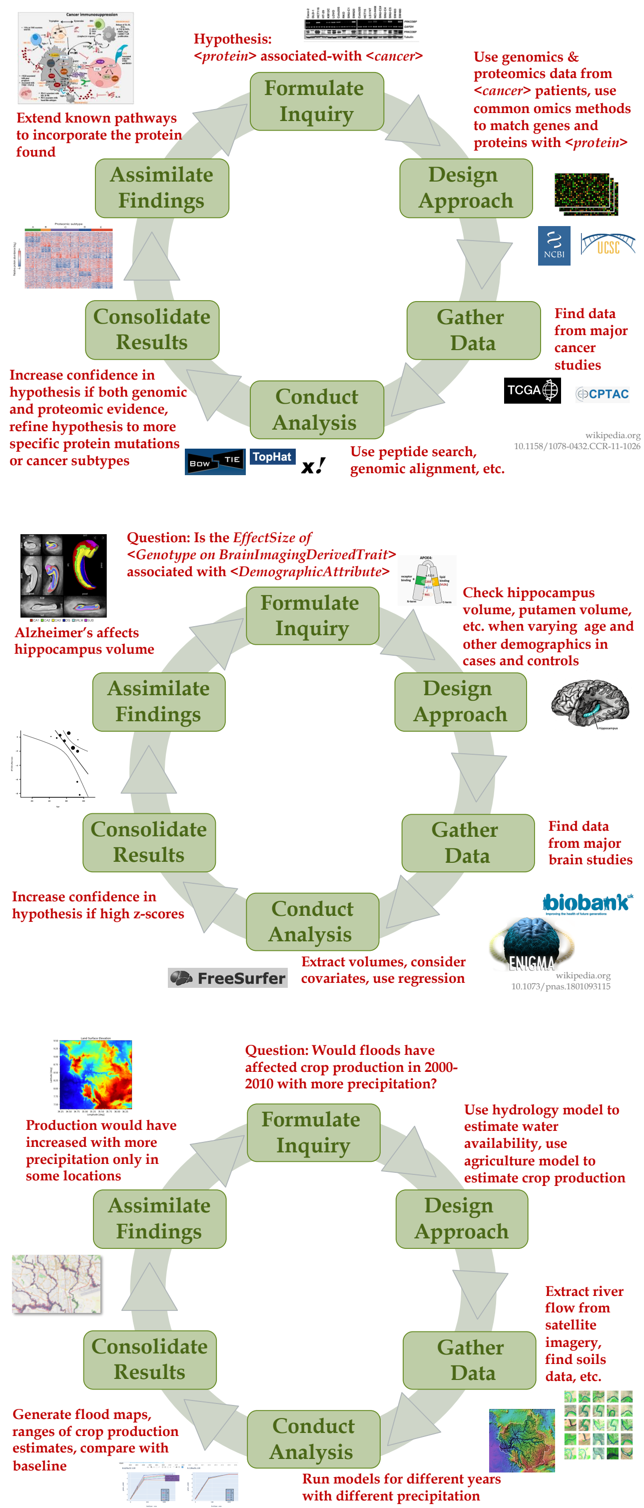


Figure 3. Illustrations of the reflection framework with real examples in three diverse areas of science: cancer omics (top), neuroscience (middle), and flood prediction (bottom).

Meta-analysis is needed to combine results from different studies. For example, in flood prediction the analysis consists of running many simulations and the meta-analysis combines them to generate prediction ranges and assess uncertainty.

The reflection framework illustrates how this core competency can drive the development of other competencies:

- The Experimentation Competency would be prompted by the needs of the Gather Data step. The reflection process can provide context for the Experimentation Competency in terms of the kind of data needed, the requirements for data collection, and the quality metrics that should drive experimentation.
- The Modeling Competency would be invoked by the needs of the Conduct Analysis step. The reflection process will determine what kind of modeling needs to be done, the data available, the performance criteria for the resulting model, and other important information for modeling tasks.
- The Probing Competency would be driven by the needs of the Conduct Analysis step. The reflection process can determine the search objectives, optimization criteria, and the domain knowledge and data that can make the search more efficient.
- The Extraction Competency would be triggered by the needs of all the six steps, as it can extract theories to be used in the Formulate Inquiries step, domain knowledge to guide the Gather Data and Conduct Analysis steps, and fusion and integration methods to synthesize Consolidate Results and Assimilate Findings.
- The Creation Competency would be needed when the Design Analysis step cannot generate appropriate strategies to find data or methods to answer key questions, or when no findings result after several iterations of the cycle, or when findings remain inconsistent despite the iterations.

6. Conclusions

Future AI systems for science will be capable of pursuing independently substantial aspects of the research and therefore make their own discoveries. They will be capable of taking on significant problems by formulating their own research goals, proposing and testing hypotheses, designing theories, debating alternative options, and synthesizing new knowledge. They will also be able to explain their reasoning, compare their lines of inference to other possible ones, and situate their findings. They will communicate with scientists who have different levels of expertise and understanding in any given research topic.

The required capabilities will only be possible through substantial research advances in a diversity of areas of AI, including cognitive systems, machine learning, knowledge representation, constraint reasoning, problem solving and planning, meta-reasoning, reasoning under uncertainty, multi-agent systems, natural language processing, collaboration, and robotics. AI research for science will also emphasize intelligent capabilities that have been received less attention in the past, such as representational change and creativity.

Acknowledgments

We would like to thank our collaborators over the years, particularly Daniel Garijo, Deborah Khider, Maximiliano Osorio, Varun Ratnakar, Hernan Vargas, Suzanne Pierce, Emmanuel Johnson, Parag Malik, Ravali Adusumilli, Neda Jahanshad, Alice Yang, Scott Peckham, Armen Kemanian, and Kelly Cobourn. We gratefully acknowledge support from the US Office of Naval Research through award N00014-21-1-2437, the National Institutes of Health through award 1R01AG059874-01, and the Defense Advanced Research Projects Agency through award W911NF-18-1-0027.

References

[1] Simon, Herbert A. “Models of Discovery and Other Topics in the Methods of Science.” Springer, 1977. ISBN 978-94-010-9521-1.

[2] Lindsay, Robert K.; Buchanan, Bruce G.; Feigenbaum, Edward A.; and Joshua Lederberg. “Applications of Artificial Intelligence for Organic Chemistry: The Dendral Project.” McGraw-Hill, 1980. ISBN 978-0070378957.

[3] Gil, Yolanda; Garijo, Daniel; Ratnakar, Varun; Mayani, Rajiv; Adusumilli, Ravali; Boyce, Hunter; Srivastava, Arunima; and Parag Mallick. Towards Continuous Scientific Data Analysis and Hypothesis Evolution. Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17), San Francisco, CA, 2017. https://doi.org/10.1609/aaai.v31i1.11157

[4] Science. “AI Transforming Science.” Special Issue, 357(6346), 7 July 2017. https://science.sciencemag.org/content/357/6346

[5] Callahan, Alison, Dumontier, Michel, and Nigam H. Shah. HyQue: evaluating hypotheses using Semantic Web technologies. Journal of Biomedical Semantics 2, S3. 2011, 2(Suppl 2):S3 http://www.jbiomedsem.com/content/2/S2/S3

[6] Gomes, Carla; Dietterich, Thomas; Barrett, Christopher; Conrad, Jon; Dilkina, Bistra; Ermon, Stefano; Fang, Fei; Farnsworth, Andrew; Fern, Alan; Fern, Xiaoli; Fink, Daniel; Fisher, Douglas; Flecker, Alexander; Freund, Daniel; Fuller, Angela; Gregoire, John; Hopcroft, John; Kelling, Steve; Kolter, Zico; Powell, Warren; Sintov, Nicole; Selker, John; Selman, Bart; Sheldon, Daniel; Shmoys, David; Tambe, Milind; Wong, Weng-Keen; Wood, Christopher; Wu, Xiaojian; Xue, Yexiang; Yadav, Amulya; Yakubu, Abdul-Aziz; and Mary Lou Zeeman. “Computational sustainability: computing for a better world and a sustainable future.” Communications of the ACM, 62(9), 2019. https://doi.org/10.1145/3339399

[7] Groth, Paul and Jessica Cox. (2017). Indicators for the use of robotic labs in basic biomedical research: a literature analysis. PeerJ. Nov 8;5:e3997. https://doi.org/10.7717/peerj.3997

[8] Tshitoyan, Vahe; Dagdelen, John; Weston, Leigh; Dunn, Alexander; Rong, Ziqin; Kononova, Olga; Persson, Kristin A.; Ceder Gerbran; and Anubhav Jain. “Unsupervised word embeddings capture latent knowledge from materials science literature.” Nature 571, 95–98 (2019). https://doi.org/10.1038/s41586-019-1335-8

[9] Craver, Carl F. and Lindlay Darden. (2013). In Search of Mechanisms: Discoveries across the Life Sciences. University of Chicago Press.

[10] Pearl, Judea. "The Book of Why: The New Science of Cause and Effect." Basic Books Publishers, 2018.

[11] Trickett, Susan B.; Schunn, Christian D.; and J. GregoryTrafton. (2005). Puzzles and peculiarities: How scientists attend to and process anomalies during data analysis. In Michael E. Gorman, Ryan D. Tweney, David C. Gooding, & Alexandra P. Kincannon (Eds.), Scientific and Technological Thinking (pp. 97-118). Mahwah, NJ: LEA.

[12] Thagard, Paul. (2012) The Cognitive Science of Science: Explanation, Discovery and Conceptual Change. Cambridge, MA: MIT Press.

[13] Wilkenfeld, Daniel A. and Richard Samuels (eds.) (2019). Advances in Experimental Philosophy of Science. London, UK: Bloomsbury.

[14] Addis, Mark; Sozou, Peter D.; Lane Peter C. and Fernand Gobet. (2016). Computational Scientific Discovery and Cognitive Science Theories. In: Müller, V.C. (eds) Computing and Philosophy. Synthese Library, vol 375. Springer, Cham. https://doi.org/10.1007/978-3-319-23291-1_6

[15] Chandrasekharan, Sanjay and Nancy J. Nersessian. (2015). Building Cognition: The Construction of Computational Representations for Scientific Discovery. Cognitive Science 39:1727-1763. https://doi.org/10.1111/cogs.12203

[16] Kuhn, Thomas S. (1962). The Structure of Scientific Revolutions. University of Chicago Press.

[17] Gil, Yolanda. "Will AI Write the Scientific Papers of the Future?". AI Magazine, 42(4). 2021. https://doi.org/10.1609/aimag.v42i4.18149

[18] Kitano, Hiroaki. "Artificial Intelligence to Win the Nobel Prize and Beyond: Creating the Engine for Scientific Discovery." AI Magazine. 37(1), 2016. https://doi.org/10.1609/aimag.v37i1.2642

[19] Gil, Yolanda. "Thoughtful Artificial Intelligence: Forging A New Partnership for Data Science and Scientific Discovery." Data Science, 1(1-2), pp. 119-129, 2017. http://doi.org/10.3233/DS-170011

[20] Gil, Yolanda; Khider, Deborah; Osorio, Maximiliano; Ratnakar, Varun; Vargas, Hernan; Garijo, Daniel and Suzanne Pierce. Towards Capturing Scientific Reasoning to Automate Data Analysis. Proceedings of the Annual Conference of the Cognitive Science Society, 44, 2022. Retrieved from https://escholarship.org/uc/item/85d2d1xf

[21] Gil, Yolanda; Garijo, Daniel; Khider, Deborah; Knoblock, Craig A.; Ratnakar, Varun; Osorio, Maximiliano; Vargas, Hernán; Pham, Minh; Pujara, Jay; Shbita, Basel; Vu, Binh; Chiang, Yao-Yi; Feldman, Dan; Lin, Yijun; Song, Hayley; Kumar, Vipin; Khandelwal, Ankush;

Steinbach, Michael; Tayal, Kshitij; Xu, Shaoming; Pierce, Suzanne A.; Pearson, Lissa; Hardesty-Lewis, Daniel; Deelman, Ewa; Ferreira Da Silva, Rafael; Mayani, Rajiv; Kemanian, Armen R.; Shi, Yuning; Lorne Leonard, Lorne; Peckham, Scott; Stoica, Maria; Cobourn, Kelly; Zhang, Zeya; Duffy, Christopher; and Lele Shu. Artificial Intelligence for Modeling Complex Systems: Taming the Complexity of Expert Models to Improve Decision Making. ACM Transactions on Interactive Intelligent Systems (TiiS), 11(2). 2021. https://doi.org/10.1145/3453172

[22] Srivastava, Arunima; Adusumilli, Ravali; Boyce, Hunter; Garijo, Daniel; Ratnakar, Varun; Mayani, Rajiv; Yu, Thomas; Machiraju, Raghu; Gil, Yolanda; and Parag Mallick. Semantic Workflows for Benchmark Challenges: Enhancing Comparability, Reusability and Reproducibility. Proceedings of the Pacific Symposium on Biocomputing (PSB), 24:208-219, 2019.

[23] Gil, Yolanda; Honaker, James; Gupta, Shikhar; Ma, Yibo; D'Orazio, Vito; Garijo, Daniel; Gadewar, Shruti; Yang, Qifan; and Neda Jahanshad. Towards Human-Guided Machine Learning. Proceedings of the 24th ACM International Conference on Intelligent User Interfaces (IUI'19). Association for Computing Machinery, New York, NY, USA, 614–624, 2019. https://doi.org/10.1145/3301275.3302324