

Department of Defense
RESEARCH & ENGINEERING

**TECHNICAL ASSESSMENT:
DATA-ENABLED
TECHNOLOGY WATCH & HORIZON SCANNING**



Office of Technical Intelligence
Office of the Assistant Secretary of Defense
for Research & Engineering
October 2015



Distribution A: Approved for public release; distribution is unlimited



About the Office of Technical Intelligence

The Office of Technical Intelligence (OTI) provides the U.S. Department of Defense research and engineering community and partners holistic, defense-relevant insights into emerging and potentially disruptive technology to enable U.S. and mitigate adversary technological surprise. To do so, OTI identifies emerging and potentially disruptive science and technology, recommends efficient research and development strategies, and coordinates intelligence collection, analysis and dissemination. OTI accomplishes these missions through three complimentary efforts: technology watch and horizon scanning, technical assessments, and tailored intelligence support and coordination.

OTI technology watch and horizon scanning efforts are developing methods to identify nascent and disruptive science, technology, and capabilities through the exploitation of tailored approaches and tools, including analysis of scientific literature, patents, and worldwide investment using both open source and internal data.

OTI technical assessments provide decision-relevant research and development strategy inputs on emerging and potentially disruptive technologies to the research and engineering community by exploring opportunities and threats the technologies could enable, conducting data-driven analyses of drivers to forecast future trends and identify unique DoD needs, recommending specific investment and policy approaches, and developing and seed funding projects to leverage those opportunities.

OTI intelligence support activities are focused on coordinating efforts across the research and engineering community, ensuring timely and valuable analysis reaches users, and providing mechanisms to enhance communication between policymakers, researchers, and analysts.

Table of Contents

Introduction	1
Potential Benefits from Data-Enabled TW/HS.....	2
Structuring Effective TW/HS Efforts.....	3
Characterizing Decisions	4
Selecting Data	4
Selecting Metrics.....	5
Conducting Analysis & Developing Decision-Support Products	6
Leveraging Knowledge Management	7
Beyond Workflows: TW/HS Infrastructure	8
Curated Data	8
Accessible Analytics	9
Conclusion.....	9

¹ Cover photo courtesy of Los Alamos National Laboratory.

Introduction

For more than five decades, the U.S. Department of Defense (DoD) has been a world leader in science and technology (S&T); however, it currently faces a range of challenges to maintaining that leadership. Cutting-edge research and development (R&D) is increasingly dispersed internationally. It has also expanded beyond the domain of established universities and large, longstanding corporations, and DoD does not have the same depth of relationships with newer technology companies, start-ups, and even community laboratories where exciting breakthroughs are occurring today. Meanwhile, the raw number of participants, amount of technical information, and sources of relevant data are all growing rapidly, creating major challenges to finding relevant information. Thus, DoD has multiple challenges to staying informed of cutting-edge work and guiding its investments appropriately, all while the importance of doing so is increasing. Competitors are challenging DoD's technical advantage, and budget pressures are limiting DoD's ability to expand what it funds.²

“Many, if not most, of the technologies that we seek to take advantage of today are no longer only the domain of DoD development pipelines or traditional defense contractors. DoD no longer has exclusive access to the most cutting-edge technology.”²

As these challenges have mounted, the data analytics field has grown rapidly, producing more sophisticated algorithms which run more quickly using more powerful, less expensive computing resources. Combined with the explosion in S&T data, this has created interest in the potential for data analytics to enable new and effective approaches to technology watch and horizon scanning (TW/HS). Technology watch is typically defined as the characterization of activity in a known field, and horizon scanning focuses on identifying new or emergent concepts. However, these concepts bleed together in many cases, so this assessment discusses them together, covering the identification, characterization, and forecasting of known and unknown science, technology, and applications.³ More specifically, this assessment focuses on data-enabled TW/HS approaches to benefit the Defense research and engineering community.⁴ While there are other approaches to TW/HS, data analytic approaches are both relatively new and especially promising. This assessment begins by reviewing the potential benefits of data-enabled approaches. Following this, the main body of the work discusses effective workflows to integrate TW/HS into decision processes, challenges to conducting effective TW/HS today, recommendations to enable further development.

² “Long Range Research and Development Plan (LRRDP) Request for Information,” (DoD, December 3, 2014). http://www.defenseinnovationmarketplace.mil/resources/LongRangeResearchandDevelopmentPlanRFI_Final.pdf

³ Typically, organizations draw a distinction between technology watch and horizon scanning based on whether analysis starts from a known topic or whether analysis is primarily descriptive or predictive. However, these are not clear divides in practice. More importantly, for the purposes of this analysis, we find no clear value in distinguishing between them.

⁴ This report does not focus on intelligence-specific applications of TW/HS, although many applications relevant to DoD may be of interest to intelligence organizations, and DoD applications may share technical needs with intelligence applications. It also does not directly consider analysis and forecasting of social, environmental, or other non-S&T areas, which some organizations include under the horizon scanning rubric, although these areas may play a role in informing S&T focused efforts.

Potential Benefits from Data-Enabled TW/HS

With the growth of worldwide S&T and constrained budgets, decision makers must make difficult choices as to how to allocate resources and develop appropriate policies. These choices affect decision makers across missions and levels. The basic research community is searching for promising and potentially disruptive new research, while DoD laboratories and other applied research organizations are seeking to enhance DoD capabilities. At the same time, senior leadership must make strategic choices that are partially based on developments in the S&T environment, which will shape the future of U.S. military forces. In addition, all levels of DoD must accurately plan for human capital needs and develop policies that stay current and can manage technology-enabled opportunities and challenges. All of these decisions can benefit substantially from a keen understanding of the current state of the art as well as acute insights into future developments. To develop these inputs, DoD regularly convenes expert groups to analyze and forecast S&T developments. However, data-enabled TW/HS has the potential to improve upon or augment current approaches by expanding the aperture of analyses and decreasing the influence of bias, while at the same time building institutional capacity.

...data-enabled TW/HS has the potential to improve upon or augment current approaches by expanding the aperture of analyses and decreasing the influence of bias, while at the same time building institutional capacity.

The S&T landscape is vast, characterized by both broad interdisciplinary study and deep fields of research. This poses a challenge to human analysis, as any realistically sized group will have limited expertise and insight across the range of potentially Defense-relevant S&T. Even for groups focused on a single field, it is increasingly difficult to monitor cross-disciplinary efforts and the potential for impacts from disparate fields. The diffusion of knowledge and globalization of S&T further exacerbates these limitations, as most of the experts DoD consults are from the U.S. and internal DoD experts are often hampered by limitations to travel, journal access, and, in some cases, even simply access to external websites. In contrast, data-enabled approaches have the potential to start from a broad base of knowledge, enabling the identification of important interactions and developments outside of the mainstream.

Compounding these challenges, there is little validation of the accuracy of expert judgments provided to DoD. While not focused on DoD groups in particular, a 2012 study looking at the accuracy of technology forecasting approaches found that many expert forecasts could not even be assessed for accuracy due to lack of clear and precise judgments, and for those that could, expert judgment fared poorly compared to other methodologies.⁵ Perhaps most interestingly, this study found that quantitative trend analysis proved most effective. This suggests that data-enabled TW/HS approaches have an advantage over traditional expert-led activities. One likely explanation for this is that data analytics has the potential to decrease the role of human biases in S&T analysis. Thus, while new approaches are not necessarily more effective, there is reason to believe that data-enabled TW/HS systems can be developed that prove more accurate, especially considering research that demonstrates how difficult it is for humans to overcome analytic biases, even when aware of them.

⁵ Carie Mullins, "Retrospective Analysis of Technology Forecasting: In Scope Extension Final Report" (Tauri Group, August 13, 2012).

Beyond analytic challenges, current expert-led approaches do not build institutional capacity. These efforts tend to be ad hoc processes without reproducibility and provide little direct benefit to the DoD experts who are tasked with providing data inputs. Data-enabled activities have the potential to lighten this task for DoD experts while also enabling reproducible analyses for more lasting value. For example, even simple functions such as saving searches and documenting initial analyses can enable much faster updates to analytic products and give analysts or experts the opportunity to review their thought processes. Thus, the initial investment in developing TW/HS efforts for a given topic can produce institutional capacity to repeat them. Data-enabled approaches still require experts to assist in interpreting results, but by doing so, data-enabled TW/HS efforts are likely to provide DoD experts a return on their time invested in the form of broader insights into their field, further creating institutional capacity. This capacity can be brought to bear on a variety of decisions, ranging from portfolio management to individual investments and human capital management.

Structuring Effective TW/HS Efforts

In order for TW/HS to be a valuable pursuit, it must provide valuable insights into S&T, *and* these insights must support decisions. Data-driven TW/HS will not replace human decision makers, so DoD will need to develop relevant technologies in concert with appropriate workflows to integrate these tools into decision processes. Otherwise, TW/HS activities will simply provide information that is “interesting,” but not impactful.

This section analyzes needs, challenges, and opportunities for integrated TW/HS workflows and technologies. OTI’s analysis in this area is based on a review of TW/HS efforts in the U.S. and allied governments, discussions with data analytics providers, interviews with R&D decision makers, insights from a recently organized TW/HS Community of Practice, and an OTI-Air Force Research Laboratory (AFRL) collaboration to test TW/HS approaches (see box at right). One of the key conclusions from these efforts is that approaching TW/HS efforts in terms of a workflow which integrates human analysts with data analytics throughout the process is critical in order to deliver valuable results. The following section discusses the key functions in a TW/HS workflow, which we divide into five phases:

- Characterizing Decisions
- Selecting Data
- Conducting Analysis
- Developing Decision Support Products
- Leveraging Knowledge Management.

OTI TW/HS TEST CASE

From April to September 2015, the Air Force Research Laboratory’s Materials and Manufacturing Directorate (AFRL/RX) and OTI collaborated to apply TW/HS methodologies to inform a future investment in structural materials. Because AFRL/RX sought to break new ground in the area, the program team proposed that data-enabled analysis might provide broader insights than in-house expertise alone. Based on a data-enabled analysis of the structural materials field, OTI provided 7 candidate focus areas. These results were still under review at the time this assessment was concluded. Conducting this study provided critical insights into the challenges of tying TW/HS efforts directly into decision processes and R&D needs.

Each of the following sections describes one of the workflow phases, identifies challenges to accomplishing it today, and provides recommendations for R&D and policies to enable future DoD TW/HS efforts.

Characterizing Decisions

In order to provide the most valuable information to decision makers, each analysis will have to take into account characteristics of the specific decision at hand. Organizations have their own particular goals and metrics for success, and without understanding these nuances, TW/HS analysis runs the risk of providing results that are interesting to the analyst, but irrelevant to the decision maker. In order to conduct effective TW/HS efforts, analysts must understand three critical factors: the decision itself, the program timeline, and the evaluation criteria.

In order to conduct effective TW/HS efforts, analysts must understand three critical factors: the decision itself, the program timeline, and the evaluation criteria.

Understanding these factors informs the scope, scale and context of the supporting analysis, which enables analysts to provide targeted, actionable inputs into the decision process in time for the information to be actionable. Defining the evaluation criteria is most critical. Evaluation criteria represent an organization's preferences with respect to the decision at hand. For example, with an investment decision, is an organization looking to invest in a novel area or a mature one to take advantage of an existing resource base? Characterizing evaluation criteria allows analysts to tailor the TW/HS program to the customer. To do so, analysts must work with decision makers to make evaluation criteria explicit and define them as clearly as possible in the organization's context. Examples of typical evaluation criteria include maturity, novelty, return on investment, and the degree to which technologies enable priority capabilities.

Challenges

At present, there is no broad analysis of the types of decisions DoD organizations undertake and the attendant evaluation criteria. While conducting specific TW/HS projects requires interaction with the customer, a broad understanding of decisions and evaluation criteria would enable analytic teams to better link similar efforts and would support the development of appropriate analytics to inform those decision criteria.

Recommendations

1. Conduct an analysis of decisions and decision criteria in the DoD research and engineering community to support future TW/HS development efforts.

Selecting Data

Based on the evaluation criteria, it is possible to identify the appropriate data to support the TW/HS analysis. Data selection requires careful balancing of relevance and breadth. It is critical to identify sources that are likely to provide signal relevant to the evaluation criteria and to maximize the signal to noise ratio. For example, patent data is less likely to serve an analysis in support of a basic research program, but it might be valuable for applied research efforts.

Managing the signal-to-noise challenge often requires selecting only a portion of a given data source. For example, in the OTI-AFRL collaboration, the OTI team found that while choosing only the materials science portion of S&T literature data source seemed an appropriate way to begin, the engineering and chemistry

sub-sets also contained substantial amounts of relevant information. At the same time, while data from biology-related subsets were potentially relevant to identify biomaterials, including the biology sub-set created an unmanageable level of noise for the software available. Thus, the selection of data requires a careful analysis of what is likely to be useful and, if possible, initial exploratory analyses to identify unexpected areas of signal and noise.

Challenges

There are three principal challenges to data selection in support of TW/HS activities. The first is that it is not yet clear which data sources contain the most relevant signal for various evaluation criteria. This challenge is discussed in further depth in the following section, as it is intertwined with the development of appropriate analytic tools to inform the evaluation criteria.

The second challenge is that query development is surprisingly time consuming and difficult. Even if an analyst knows the proper terminology for a field, developing queries can take days, involving search strings that can stretch on for pages and require the use of complicated 'languages.' To demonstrate the scale of this challenge, a modestly sized search the OTI team used during the AFRL collaboration contained 196 parentheses to satisfy the constraints of the query language. Not only is this process arduous for analysts, but determining when a query is "right" is a difficult process, often involving extensive trial and error.

The third challenge is developing methods for selecting sub-sets of data when analysts are not experts in all aspects of a field. Scientists, engineers, and analysts produce articles, patents, and other forms of data in the language of their own field which does not translate to another field. For example, the concepts used to outline a potential applied research program might not find relevant research conducted in a basic research context. As a result, even with subject matter expert inputs, analysts may still be challenged to develop queries or leverage other approaches to capturing relevant areas from disparate fields or emerging areas within a field because experts may not be aware of these.

Recommendations

2. Develop query tools which aid analysts in query generation and characterization.
3. Develop analytic tools that can use seed terms or exemplars to identify relevant terminology or data produced across fields, research contexts, and data types.

Selecting Metrics

In order to inform evaluation criteria, analysts must select appropriate **metrics**. Evaluation criteria are often complex human ideas which cannot be precisely calculated from data. For example, analytics cannot directly assess the maturity of a technology, but they could analyze the amount of activity which references the technology, growth rates of activity, or identify whether sources discuss prototyping or advanced testing to inform a technology readiness level estimation. We refer to these proxies or models for evaluation criteria as **metrics**. In each TW/HS effort, analysts must choose appropriate metrics and the attendant algorithms to calculate them based on the decision makers' evaluation criteria. This discussion separates the selection of data and metrics for clarity, but selection of each should inform the other in an iterative process.

Challenges

Across the TW/HS field, there is a lack of validated metrics. This means that, while analysts may make claims based on attributes of the data – for example, that the top publishers are leaders in a field – there is a limited basis on which to assert that these claims describe the ground truth. Continuing the publishing-leadership example, many analyses of patent data identify China as the clear leader in research fields due to the fact that Chinese sources publish far and away the most patents; however, the extent to which these document new and important work is not clear, so identifying China as a clear leader is not necessarily accurate. Because of the inseparable nature of metrics and the data that analysts use to calculate them, this is also a severe impediment to data selection. Thus, a lack of validated metrics is a critical weakness for the TW/HS field.

Across the TW/HS field, there is a lack of validated metrics. This means that, while analysts may make claims based on attributes of the data...there is a limited basis on which to assert that these claims describe the ground truth.

Beyond validating metrics, there is relatively little activity generating new metrics and algorithms to calculate them. This is a further challenge to TW/HS efforts, as it means that there are relatively few options to inform the broad range of evaluation criteria of potential interest across DoD. One of the major technical challenges to developing new metrics is that many current methods do not work effectively when analyzing across multiple data types. For example, while relationships between scientific publications and patents may be valuable to assess attributes of a technology, many approaches to analyzing S&T data – such as clustering to identify similar concepts for analysis – are not effective when using multiple data types.

Recommendations

4. Begin a program to validate existing metrics and their associated algorithms to ensure that they describe real S&T phenomena.
5. Invest in the development and validation of new metrics to inform the range of evaluation criteria of interest to DoD.
6. Develop approaches to analyze multi-source S&T data to enable the development of future metrics and their associated algorithms.

Conducting Analysis & Developing Decision-Support Products

With the selection of data and metrics, analysts can conduct their initial analysis. To enable more effective application of metrics, it is often valuable to develop a taxonomy of the field under consideration. Taxonomies allows for the identification of areas at the same level of abstraction. Breaking down research into categories and sub-categories enables the comparison of sub-fields to identify how they rank relative to various metrics. For example, calculating maturity-related metrics typically does not make sense for an entire corpus of data, but it may be valuable to calculate for individual technologies in order to prioritize them relative to an organization's decision framework.

Beyond calculation of metrics, analysts must integrate the disparate portions of their findings into a cohesive whole in order to make their efforts useful to decision makers. Creating a decision support product requires understanding what is useful to the decision maker, such as whether the individual metrics or a composite score would be most useful and how to communicate the findings so that they are both clear and most likely to be used effectively. For example, depending on the content and customer, a

beautiful graphic from data analysis software might provide deeper understanding, or it might distract from the focus of the analysis and confuse the audience.

Challenges

It is currently challenging to compare areas for decision makers because it is difficult or impossible to generate accurate, tailored taxonomies. Taxonomy generation is still a manual, expert-reliant process. For the OTI-AFRL collaboration, while materials science experts could provide insight into various areas of research in fields of interest, they were unable to provide a taxonomy of the structural materials field under review at the level of specific materials, which is the level at which AFRL sought to make an investment. Even where experts could provide valuable inputs, it was not clear that these provided a holistic view of the field, which would incorporate bias into the analysis if they did not include relevant technologies.

Partially due to the infancy of the TW/HS field, there is also little research on how best to present results to decision makers. Currently, analysts take a broad range of approaches – and get a broad range of results, from confusion to beneficial impact. The extent to which decision makers, especially senior decision makers, will gain value from direct access to TW/HS analytics is also unclear.

Recommendation

7. Develop semi-automated taxonomy generation approaches that allow for broad, accurate coverage of data sets, but that also enable analysts to tailor them to the specific TW/HS project.
8. Conduct research to identify how best to present S&T data to decision makers and conduct pilots to test the value of access to TW/HS analytics at various levels of leadership.

Leveraging Knowledge Management

In order to move from a successful TW/HS *project* to a TW/HS *program*, it is important to ensure that products can be kept up to date with manageable amounts of effort and to track the accuracy of analysis. While organizations save final products, the intermediate steps in TW/HS efforts are critical to reproducibility. For example, maintaining precise records of searches, data characteristics, and analytics versions allows analysts to update conclusions without repeating the entire project and to ensure that comparisons with prior work are appropriate. Tracking the accuracy of forecasts and other conclusions is also critical. This ensures that analytic methods are effective and allows for the prioritization of more effective approaches. While data-enabled TW/HS approaches appear promising, new technologies are not necessarily more effective than prior approaches, so measuring effectiveness is critical to demonstrating and increasing the value of TW/HS programs.

In order to move from a successful TW/HS project to a TW/HS program, it is important to ensure that products can be kept up to date with manageable amounts of effort and to track the accuracy of analysis.

Challenges

There are presently few resources or organizational incentives to track accuracy. Additional work to produce today's product just so it is easier to update in a year is often not done or captured in an enduring fashion.

Recommendations

9. Develop simple tools for making TW/HS processes more easily reproducible in order to maintain currency of analytic products and to enhance the return on time invested.
10. Fund retrospective studies to track the accuracy of TW/HS analysis to identify the most effective approaches and ensure TW/HS is providing valuable insights to decision makers.

Beyond Workflows: TW/HS Infrastructure

While the above steps are critical to conduct effective TW/HS efforts and build effective programs, the process is not possible without supporting infrastructure. In particular, TW/HS requires curated data and accessible analytics which are able to work together.

Curated Data

In order to generate the benefits promised by data-enabled analysis, TW/HS programs require access to full-corpus, curated, well-documented data. As one of the main promises of TW/HS is to deliver analysis from a broad base of data, analytic efforts require access to full databases – the ‘full corpus’ – as opposed to web-search or other access models that limit the amount of data users can analyze, which instantiate biases into the process from the outset. Having possession of the full corpus also allows for curation and cleaning. While quality varies substantially, all data sources require additional processing for the most effective use in TW/HS, especially disambiguation of authors and institutions. Because of the time and manpower required to effectively curate data, this almost always must be completed before starting a TW/HS project if the goal is to deliver results within a typical decision-support window. In addition, TW/HS programs should also provide strong documentation surrounding the source of the data, currency, and additional processing carried out on it in order for users to best tailor their analysis and document the strengths and weaknesses – such as potential blind spots – of their analysis.

Because the development of metrics is its infancy, it is not yet clear what data sources will be most useful to future TW/HS efforts. However, without data, it is not possible to experiment with workflows and metrics, so R&D efforts still require full-corpus, curated data.

Challenges

The principal challenge to delivering full corpus, curated data is cost. Providers of more highly curated data sets, such as the popular citation databases Scopus and Web of Science, charge substantial subscription fees varying depending on the number of users. Organizations typically purchase data access on a ‘per seat’ basis, so it is also difficult to negotiate contracts or agreements for data access without knowing future resource requirements. While free and unrestricted data sources may seem relatively appealing, they also come with substantial costs. For example, web and blog data can often be accessed for free, but curating this data to make it useful for analytics is challenging and resource intensive. An area that can carry further, special challenges is U.S. Government and partner internal data. Because of limitations on use – for example data marked For Official Use Only – these data sets may require special storage and network access limitations, which influence the accessibility of systems described in the following section.

Recommendations

11. A DoD organization should provide full-corpus, curated S&T data with minimal restrictions for access and which is available to use on commercial systems to support TW/HS development efforts. A central

coordinating office would decrease contracting challenges, minimize the duplication of curation efforts, and increase data availability, enabling metric and workflow development throughout DoD.

Accessible Analytics

While the development of metrics and the associated algorithms is still at an early stage, R&D activities and follow-on systems will require infrastructure to house those algorithms and the associated data. This infrastructure must be flexible enough to load new analytics into and powerful enough to return results in a useful timeframe for analysis. This will require non-trivial computational resources as we expect that many sophisticated metrics will require analytic engines to conduct complex processing on terabytes of data.

R&D activities and follow-on systems...must be flexible enough to load new analytics into and powerful enough to return results in a useful timeframe for analysis.

In order for TW/HS tools to be useful, they must be accessible to analysts at their desks, and they should enable collaboration. Most TW/HS projects will involve multiple analysts and experts consulting at varying times, and sophisticated analyses require many intermediate steps and refinements which benefit tremendously from collaboration. As with data, documentation is critical to enable analysts to understand what calculations analytic tools are performing and any updates that affect comparability with other analytic engines or earlier analyses.

Challenges

The major challenge to providing flexible, easy-to-use analytics and computational resources is accessibility. If systems are deployed on NIPRNet, they may have severe restrictions to software and connectivity to external resources. These potential restrictions are even more daunting considering that TW/HS is still developing, which will benefit from relatively rapid iteration cycles in algorithm development, which might require recertification for use on DoD systems each time updates are made. Depending on the source of data and algorithms, some activities may also need to take place on classified networks, which pose even further challenges.

Recommendations

12. Because efforts are still largely in the R&D phase, DoD should establish an unclassified development environment without the restrictions posed by NIPRNet and other systems. This should be a flexible system with built-in data that enables developers to integrate and test new analytics quickly and easily. Analysts should have concurrent access to these systems to enable a conversation between developers and future users to ensure the relevance of new approaches.

Conclusion

The data-enabled TW/HS field has the potential to revolutionize decision making in the DoD research and engineering community. However, the field is still in its infancy, and it will not achieve a high level of impact without a broad range of R&D efforts. Just as importantly, this field will only yield benefits to DoD if researchers and analysts develop it with the appreciation that humans are still making the decisions and that the data analytics is only there to support them. For these reasons, this assessment focuses on providing recommendations for process and research, with the goal of enabling DoD and partner TW/HS efforts to blossom.