Introduction

This special issue of Technological Forecasting and Social Change addresses the topic of “tech mining.” Tech mining strives to inform science, technology and innovation (“ST&I”) management by applying text mining to exploit searches of electronic science and technology databases [1]. It is richly informed by work in fields such as information retrieval, scientometrics and content analysis. Current tech mining efforts co-exist with related efforts, including technology futures analysis, research evaluation and competitive intelligence.

Contrast tech mining with traditional ST&I policy and management practices. Traditional processes rely mainly on expertise. This could be anything from asking the “right-hand man” to convening multi-tier expert panels. The former relies excessively on casual knowledge; the latter tends to be slow and expensive. Both neglect the rich intelligence to be gleaned from data resources. Tech mining takes information already compiled, for whatever purposes, and exploits it just when needed. It derives knowledge pertinent to the issue at hand. For instance, today we might search fuel cell patents to help assess how crowded the landscape is to help assess our new program initiative. Tomorrow, we might mine fuel cell patents to identify which other companies are the best prospects to license our intellectual property. In both cases, we get the intelligence immediately; we don’t have to wait for a third party to perform special studies taking months.

Tech mining is not restricted to mining abstract publication and patent records. It combines text and numerical data to best answer the questions confronting us. It draws on multiple content sources; for instance, checking websites after database search result analyses spotlight the key players. It particularly gains from incorporating expertise in formulating analyses and in reviewing preliminary results to spot gaps and offer interpretations. Tech mining is inherently interdisciplinary, drawing on skills of ST&I analysts, information professionals, technical specialists, and technology managers. Probably the best way to get a feel for this field is to read a sample of active research and practice – hence, this issue.
In this editorial overview we highlight the main dimensions of tech mining, relating these to the particular themes addressed by our contributors. We also introduce relevant research from disciplines outside of tech mining. The contributors in this special issue discuss the users of tech mining and their organizations. Like our contributors, this overview also discusses designing the tech mining system and the learning process. We consider the nature of ST&I systems – this is, after all, what tech mining is designed to appraise and facilitate. The variety of contributions, which we sample below, shows the health and vibrancy of this emerging field.

**Tech Mining: Our Users**

Tech mining practitioners are becoming increasingly aware of the needs of their users. In this special issue we see the variety of firms, organizations and institutions making use of tech mining to improve decision-making processes [2]. This is a healthy development, as is a growing awareness by analysts and researchers of the multi-faceted character of the user community.

Tech mining is used by private firms. Smaller firms play a critical role in societal processes of innovation. Smaller firms need to be alert to the research and development (R&D) efforts of large corporations. Yet, these firms have limited resources for technical intelligence; regardless, they must respond fast and flexibly to the emergence of new knowledge, changing technological platforms, and new products. Information on start-up companies and new knowledge initiatives may itself be mined to discover the correlates of success and failure. The funding of start-ups brings us to the roles of governments as sponsors of ST&I initiatives. Governments need to track emergence of capabilities and competition.

Tech mining thus serves government needs as well – i.e., ST&I policy-making, R&D program management, and research evaluation. Government agencies and large corporations pursuing mission-oriented research can and do use tech mining approaches to keep abreast of shifting opportunities [3]. Tech mining informs government policies and actions oriented towards the creation of new business and enhanced industry-university links. Government can additionally use these capabilities to evaluate effectiveness of their policies and actions.

Not surprisingly, tech mining users parallel those stakeholders most responsible for the development of new knowledge. The authors of this issue give a variety of perspectives on the use of tech mining by industry, government and academia. Additionally of note, Etzkowitz and Leydesdorff survey leading theories
of innovation [4]. The authors argue that institutional arrangements involving the
university, the firm and the state are all important in understanding the emergence
of knowledge-based societies. Our reflection on tech mining next follows a natural
progression from identification of users to consideration of the internal processes
and decision-making capabilities of target organizations.

Organizational Processes of Learning

The collected papers in this special issue make it clear that tech mining
contributes to decision-making in at least three distinct ways. Tech mining systems
enable organizations to identify experts both within and without the organization.
They allow organizations to speed learning over past efforts. And finally, tech
mining systems can help an enterprising organization synthesize new knowledge
from disparate sources of information.

Kostoff [5] provides a systematic way of exploiting science and technology
information to enable radical new innovations within an organization. This radical
discovery approach to tech mining calls upon externally identified experts via
literature-assisted discovery. This approach also uses literature-based discovery to
tease hitherto unexplored linkages and connections among existing bodies of
knowledge.

Tech mining necessitates a distinct set of social roles inside the organization.
Multiple authors in this special issue have examined the roles and responsibilities
created by a tech mining system in operation. Roles identified include the tech
mining role, the research management role, the development management role, and
the citizen. The tech mining role requires exceptional access of the analysts to
decision making within the organization. In addition, those analysts need skills at
communicating complex information. Other authors in this special issue have
examined the distinct responsibilities of those sponsoring new research, versus
those implementing existing and available knowledge. Tech mining promises to
assist in building deliberative, structured decision-making about new technologies
within organizations. And finally, as noted in these articles, we need to identify the
interests of concerned citizens and consumers of technologies in question.

For many organizations tech mining initiatives serve as a gateway to
accessing new knowledge about ST&I. A careful balance of internal competencies
and external access to information must be maintained within the organization.
Indeed, as our contributing authors have noted, even new knowledge about science
and technology is retrofit into existing patterns of practice within organizations.
Tech mining must be understood in terms of competence building, knowledge cumulation, and exploitation of new knowledge within the organization.

Trumbach and co-authors [6] highlight how existing social networks are complemented by the virtual network assets provided by tech mining activities. The authors fruitfully explore the idea of “weak ties” from social network theory. These ties, it is argued, provide the richest sources of new information for organizations. Tech mining can be seen therefore as a source of networked knowledge and assets that “allow a firm to have more strategic flexibility and to be more adaptive to changing environments.”

Tech mining practitioners are attempting and achieving integration between the social and technological architectures of their client organizations. As Brown and Duguid [7] argue, knowledge is the glue that binds the modern organization together. Attempting this integration requires focus on building new links between knowledge and practice, recognizing the role of translators and knowledge brokers within an organization, and correctly situating a firm in an often complex inter-organizational network.

Designing the Infrastructure for Tech Mining

A software infrastructure is undoubtedly a critical part of the extended tech mining system. ST&I databases today seem largely indifferent to their use by those looking for broad trends in knowledge. As a result, database interfaces are tailored to the “one user, one output” model. More fruitful for tech mining purposes – and we would suggest probably more valuable across the board – are “multi-user, multi-output” styles of use. Users increasingly need to find small and manageable collections of articles, and they need to disseminate research profiles (as well as selected articles) to colleagues in their organizations.

Some progress is being made along these lines in software. The database front-ends are beginning to include support for more elaborate analyses such as lists and “query by example.” Not surprisingly however, most tech mining work still requires dedicated software packages. Several practitioners in this issue are using VantagePoint software – more again are using basic spreadsheets to represent trends, graphs and other findings.

The challenges in designing better databases of ST&I – and for that matter, building better analysis software – may lie in a fuller appreciation of the task itself. Vakkari [8] reviews the state of the art in the field of information searching, noting how recent information system designs have moved from representations of the
documents, to a more cognitive approach that appreciates the variety of human tasks for which the database is built. Six categories of information seeking suit the activities of social and physical sciences (starting, chaining, browsing, differentiating, monitoring and extracting) [8]. The tech mining task may entail additional, unique operations.

The selection of the database itself is also an appropriate topic of interest. Authors in this issue note how awareness of database content and structuring are prerequisites for success in any tech mining effort. Intriguingly, multi-database studies draw increasing attention. Mining different databases in conjunction can reveal technology at different stages of research, development and societal diffusion.

The “inputs” to tech mining warrant attention – i.e., the query used for retrieving articles, patents, etc. Several authors in this issue use complex, multi-faceted queries in an effort to get precisely the data they need from the database. This approach emphasizes the precision of the match between the data and the topics of study. Other authors are exploring approaches to expand their queries, and to discover unexplored linkages across fields of knowledge. In contrast, this approach emphasizes strong data recall. Text mining software can take advantage of bolder recall (to capture related domains) by selectively removing items after the search. That is, high precision (low noise) becomes less critical.

Shapira and Youtie [9] highlight a design challenge that all tech mining practitioners face. Decision makers need to tap the “stocks and flows” of the modern knowledge economy. And it appears that the most valid measures of knowledge flows are the most complex ones. Yet despite this, our decision-makers mainly use simple indicators:

*We found that simple disaggregated methods are more apt to be taken on board by practitioners than cumulated, complex indicators. Data that could be “drilled down” to identify individual level information in listings was consistently valued by economic developers [9].*

This trade-off between the seemingly irreconcilable goals of validity and utility demonstrates that building a tech mining system is really a matter of design. In the next section we consider methodological issues associated with selecting an appropriate research design.

**Designing the Tech Mining System**

Multiple articles here consider how to combine tech mining with complementary methods. Contributors also discuss the importance of combining
method with process. Better process designs are needed, for instance to handle integrated ST&I decision-making and to address the needs of diverse stakeholders.

Becker and Sanders [10] examine the relationship between tech mining and other methods of investigation including meta-analysis, value transfer, and social impact assessment. Meta-analysis combines many smaller research studies into a single larger study, potentially increasing the statistical conclusion validity of the results. Value transfer is a formal technique for increasing the generality of model-building efforts. Those using value transfer take parameters estimated from one model, and apply these findings to new models and new policy domains. The authors find parallels between tech mining efforts to combine and summarize results, and the quantitative processes used in these two methods. In contrast, social impact assessment offers a more qualitative family of techniques which may elaborate social aspects of ST&I, as noted through tech mining.


There are multiple perspectives about, and within, technology. We need techniques to better assess multidisciplinary knowledge. Santos et al. [13] argue that new methods, and the inclusion of diverse stakeholder perspectives, can meet this assessment challenge. They also address the design and embedding of tech mining systems into larger organizational processes. The authors provide a structured decision-making process combining both qualitative and quantitative elements. The authors argue for inclusion of tech mining in the initial phase of technology decision-making, where there is particular need for broad and unbiased information about ST&I.

Tech mining techniques suffer shortcomings. Consider the following list, paraphrased from this issue:

- Not all R&D activities are patented or published, and thus cannot be measured by tech mining
- Database compilations of publications and patents are rarely timely, resulting in indicators that permanently lag the discoveries
- The counting of publications (or even citations) can not unequivocally determine quality
• Institutions have diverse intellectual property and knowledge management objectives, making tech mining efforts (at best) a complex, surrogate indicator of underlying strategy
• Classification and indexing is imperfect, resulting in unreliable innovation indicators.

Even this incomplete list is a sobering reminder of the challenges facing us. The intelligent recombination of techniques to address threats to validity is perhaps the most important reason to consider enriching tech mining systems with a broad awareness of techniques and processes.

The contributors to this special issue are not the only ones to make a call to action for integrating diverse methods and processes for forecasting new developments in technology. For instance, Porter and co-authors [14] write

_There are many overlapping forms of forecasting technology developments and their impacts, including technology intelligence, forecasting, roadmapping, assessment, and foresight. There has been little systematic attention to conceptual development of the field as a whole, isolated but uncoordinated research on improving methods, selection of methods, or integration of analysis and stakeholder management._

Such an integrated perspective can better forecast new technologies; assess multidisciplinary and interdisciplinary knowledge; and help apply tech mining studies for decision-makers and stakeholders. In the next section we consider the role of the innovation process, and the ecological validity of tech mining studies within this process.

**Innovation Systems**

An overarching tech mining theme concerns innovation systems. Our contributors reflect on structure and interpretation. We must know what aspects of innovation processes we are measuring to successfully perform tech mining. Innovation is complexly valued. In order to provide guidance to decision-makers we must understand the multiple ways that ST&I advances relate to system objectives.

Several of our authors rely on theories of network coordination. Contributors here depict “stocks” and “flows” of knowledge; these could be represented as “nodes” and “edges” in a graph. They discuss processes (such as communication and diffusion) across networks of innovators. Some contributors make explicit use of social network theory, relating the theory of structural holes to the valuation of research in a complex knowledge network. Network theories invite a specific view of research managers: the role of managers is to build bridges among disparate fields of knowledge, often by involving the appropriate disciplinary stakeholders.
Managers also uncover existing connections hidden within the wealth of available ST&I information.

Weinberg writes in his classic on “Criteria for Scientific Choice” [15] that decision-makers should consider both internal and external criteria when assessing science and technology. On the internal side, Weinberg asks: “How well is the science done?” On the external evaluation side, he asks: “Why pursue this particular science?” Elaborating on scientific merit, he writes

*The answer to the question: Does this broad field of research have scientific merit? Cannot be answered within the field . . . the scientific merit of a field can be judged better from the vantage point of the scientific fields in which it is embedded [15]*

This external criterion of scientific merit represents a challenge for tech mining.

The contributors address the embedding of knowledge among scientific fields. They consider the lack of enduring value of published knowledge. Our authors ponder the consequences of the “balkanization” of knowledge, whereby fields of ST&I are divided into ever smaller fields of specialization. They consider the challenges for tech mining in finding threads of related knowledge which threaten to be severed by disciplinary orientations of scientists and engineers. They also acknowledge the difficulty in finding a single measure of “social merit” for new science and technology. Processes and stakeholder management appear to be the best techniques for this form of assessment.

In contrast the contributors of this issue have been little interested in “internal criteria.” This point perhaps more than any other clearly distinguishes tech mining from the related field of “research evaluation.” Our peers in the field of research evaluation seek objective, partial indicators of scientific competence and excellence. Tech mining is rather more interested in growing the capability of organizations from within, and in uncovering the embedded nature and key markers of existing programs of science, technology and innovation.

Salo et al. [16] consider both internal and external criteria for science. Their task is novel: they mine a database of a national innovation program. They have measurable inputs and multi-valued outputs. Salo and co-authors acknowledge the challenges of measuring the preferences of their decision-makers. As a result these authors apply a robust form of multi-criteria analysis in their appraisal of an existing national innovation program.
Next Steps

We began this special issue with the desire to learn more about current research and activity in the field of tech mining. This overview strives to overview seven leading contributions to this new field. One thing that surprised us is the emphasis on tech mining organizations revealed by these articles. Tech mining practitioners are, rightfully, paying close attention to the flows of knowledge with their organizations. They are considering ways to enhance existing decision-making processes, and they are considering how new sources of information are going to impact their client organizations.

We’d feel remiss if we didn’t turn tech mining techniques on ourselves to ask “what are we reading?” and “where are the authoritative sources about tech mining?” There are six journals cited by multiple authors in this issue. Not surprisingly, one of the most cited is *Technological Forecasting & Social Change*.

Table 1: Journals Cited and Subject Categories

<table>
<thead>
<tr>
<th>Journal (alphabetical ordering)</th>
<th>Subject Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>California Management Review</td>
<td>Business Management</td>
</tr>
<tr>
<td>IEEE Transactions on Engineering Management</td>
<td>Engineering, Industrial Business Management</td>
</tr>
<tr>
<td>Management Science</td>
<td>Management Operations Research &amp; Management Science</td>
</tr>
<tr>
<td>Research Policy</td>
<td>Management Planning &amp; Development</td>
</tr>
<tr>
<td>Technological Forecasting &amp; Social Change</td>
<td>Business Planning &amp; Development</td>
</tr>
</tbody>
</table>

These journals range across several subject areas, with the categories Business and Management taking the lead. We note the presence of research policy and management of technology – suitable to ST&I foci. This set reinforces the emphasis on application rather than on information science.

In summary, we feel that the field of tech mining is well-spawned and beginning to mature. The topic is finding a home in journals like *Technological Forecasting & Social Change* (and other leading journals). Tech mining practitioners are extending and building their theoretical bases. Practitioners are looking into the organizational impacts of their efforts. Practitioners are also appraising, in a fundamental way, theories of innovation systems, and considering how these
theories can be gauged by analyzing search results in leading ST&I databases. The contributions that follow offer a valuable cross-section of ideas in this emerging field.

References


