The Roles and Economic Impacts of Technology Infrastructure

(version 3)

Gregory Tassey Senior Economist National Institute of Standards and Technology tassey@nist.gov

March 2008

Abstract

Trying to conduct R&D without accurate and consistent measurement techniques, or to control production processes to attain high yields and quality levels without sensors that instantly and accurately measure the performance attributes of a product as it is being manufactured, or to execute market transactions between buyer and seller without assurance of performance through accurate acceptance testing are unthinkable in today's highly competitive global marketplace. As a result, most high-tech industries are measurement-intensive. For example, semiconductor design and manufacturing requires a large and diversified set of measurement infratechnologies and associated standards that are applied directly to individual products and also to the highly automated production systems used in the modern semiconductor plant. Between 1996 and 2006, the semiconductor industry is estimated to have spent \$12 billion on measurement services, which generated \$51 billion in economic benefits.

Beyond measurement activities within an industry, increased specialization in response to the pressures of global competition has created larger numbers of market interfaces and hence the need for interoperability for a wide range of information flows. NIST economic studies have shown that inadequate interoperability standards are costing the American economy billions each year.

Regulation places another extremely large demand on the measurement infrastructure. In just one regulatory area—standard reference materials (SRMs) for measuring the content of sulfur in fossil fuels—the 29 sulfur-related SRMs provided by NIST are used by an entire supply chain to design and calibrate pollution monitoring and control equipment. A NIST impact study estimated that the net present value of economic benefits to these industries from more efficient equipment calibration and monitoring was \$441 million (2002 dollars). If the average benefits per sulfur SRM are assumed to accrue to all of the approximately 1300 NIST SRMs provided to industry, the total economic benefit (net present value in 2002 dollars) is roughly \$20 billion.

The Roles and Economic Impacts of Technology Infrastructure

Gregory Tassey

The Economic Roles of Technology Infrastructure. One of the reasons the Industrial Revolution succeeded to a greater degree in the United States than in other economies was the investment in "traditional" economic infrastructure. Roads, bridges, canals, and railroads unified regional markets and thereby greatly increased economic efficiency.

At the same time, standardized parts enabled the factory system to evolve and a growing increase in measurement standards increased equity in trade and improved the efficiency of market transactions for the plethora of new products emerging from many new technologies. These events were followed by the advent of communications networks, which, along with emergence of the "factory," were the first true technology-based systems. Both required new types of infrastructure support, including functional interfaces among components of these systems.

Today, winners and losers across countries from the rapidly evolving globalization of the technology-based economy will again be significantly affected by investments in new kinds of infrastructure. Two cross-cutting factors will determine competitive outcomes. One is the relative ability among economies to identify and accumulate the complementary private and public technology assets that address the increasing complexity of emerging technologies. The second is the need to manage technology infrastructure effectively. The latter requires determining (1) what is infrastructure, as distinct from private goods and services (i.e., wafer thickness measurements vs. semiconductor devices) (2) the sources and ownership of this infrastructure (i.e., company-specific protocols vs. industry standards), and (3) to whom and through what mechanisms should access be allowed (i.e., R&D cost sharing, IP ownership, user fees).

Infrastructure in the form of a wide variety of infratechnologies and associated standards are essential to conduct R&D, control production processes for quality and yield, and finally to execute market transactions for complex products.¹ The need for technology infrastructure at the product/component level has grown not only because technologies have become more complex, but also because this complexity has pushed companies to specialize in fewer technologies. The result is a greater dispersion of R&D

¹ Examples of measurement infratechnologies are measurement and test methods, process and quality control techniques, evaluated scientific and engineering data, and the technical basis for product interfaces.

and hence innovation within supply chains.² For goods, this phenomenon has created more market interfaces for each technology and therefore the need to communicate complex design, procurement, accounting, and logistics data across many more companies. Service industries, which are the last tier in many supply chains, now exchange huge amounts of information on a daily basis with the manufacturing tiers and their customers.

Without a comprehensive and efficient technical infrastructure, only a fraction of these information flows would take place at anything approaching reasonable cost and with mandatory security. For example, electronic funds could not be transferred without conformance to standards that ensure both accuracy and privacy, and information could not move across the Internet without standards to define and control the information flows. However, the explosion of data-driven business interactions constantly creates new needs for common formats to enable data to be exchanged amounts business partners and for algorithms to effective mine data and interpret results that unambiguously drive strategic and operational responses.

One of the key characteristics of measurement infratechnologies is their ubiquitous economic role. Every phase of the R&D stage and the subsequent manufacturing and commercialization (market development) stages of economic activity require measurement infratechnologies and associated standards. 19 microeconomic studies conducted by NIST have shown the economic efficiency gains realized from the availability of this infrastructure to be substantial with an average benefit-cost ratio of 44:1.³

Measurement and Testing Infratechnologies. Successful innovation and the attainment of "first-mover" advantage require a diversified and pervasive set of technical infrastructures that support all three stages of technology-based economic activity: R&D, production, and commercialization/market development. The number and variety has increased steadily due to the increasing complexity of modern technologies and the associated industry structures. Without such infrastructures, technological progress in an economic sense would come to a halt.

Imagine trying to conduct R&D without critically evaluated data on the properties of advanced materials, or accurate and consistent measurement techniques, controlling production processes to attain high yields and quality levels without sensors that instantly and accurately measure the performance attributes of a product as it is being manufactured. Because these needs are essential, most high-tech industries are measurement-intensive.

For example, semiconductor design and manufacturing requires a large and diversified set of measurement infratechnologies that are applied directly to individual

² A supply chain is a group of industries that provide (supply) each other with intermediate goods until a final product or service is produced. Thus, an electronics supply chain would consist of suppliers of raw materials (e.g., silicon), semiconductors, computers and other electronic equipment, systems of equipment and associated software (e.g., a communications network), and finally a service that uses the network (e.g., electronic funds transfers).

³ These and other NIST economic and policy studies along with staff papers are available at <u>http://www.nist.gov/director/planning/strategicplanning.htm</u>.

products and also to the highly automated production systems used in the modern semiconductor plant. A 2007 study of measurement investment and its economic impact in the U.S. semiconductor supply chain estimated that these industries invested \$12 billion in measurement over the 10-year period, 1996–2006.⁴ This investment generated \$51 billion in economic benefits for a net benefit of \$39 billion.⁵ Assuming that the remainder of what is commonly labeled as the "high-tech sector" invests in and benefits from measurement at approximately the same rate as the semiconductor supply chain, a rough estimate of the net economic benefits to this sector over the 10-year period is \$455 billion (extrapolated based on shares of GDP). Because the high-tech sector accounts for only about 7 percent of GDP and even allowing for the fact that the rest of the economy is considerably less measurement intensive, the net benefits to the entire economy are considerable.

However, many of the measurement infratechnologies needed to support this investment are not available, implying a shortfall in industry innovation and productivity growth. In the view of the Semiconductor Industry Association (SIA), "One of the biggest barriers to continued miniaturization of semiconductor circuits is the inability to measure many of the critical dimensions at the desired feature size. SIA, in its industry roadmap and other public statements, has emphasized the growing need for infrastructure support from government, in particular, from NIST.

Another compelling example is software. The high error rate in software makes testing for quality assurance critically important. Virtually every business in the United States now depends on software for development, production, distribution, and after-sales support of products and services. Innovations in fields such as robotics manufacturing, nanotechnologies, and human genetics research all have been enabled by low-cost computational and control capabilities supplied by computers and software.

Yet, a 2002 NIST study estimated that the direct costs to the software supply chain due to failure to identify (successfully test for) "bugs" when they were introduced into the development process and the subsequent failure to efficiently remove them when found are extremely high. The estimate of direct costs compiled from industry survey data for the U.S. economy was \$60 billion per year, and this estimate did not include costs to end users such as lost business (for example, the cost of shutting down the New York Mercantile Exchange in 1998 due to a software failure).⁶

Moreover, this cost estimate is only for one testing infrastructure. Many, many such infrastructures exist in the high-tech economy and are essential to the efficient functioning of the development and use of the product/process technologies they support. Thus, the aggregate dead weight economic loss to a technologically advanced economy from inadequate technical infrastructure is obviously huge.

Emerging technologies today have complex and deep scientific underpinnings. Consequently, the supporting measurement infratechnologies are equally complex and

⁴ The tiers (industries) in the semiconductor supply chain studied by RTI were chemical/material suppliers, equipment suppliers, front-end processing firms, and back-end processing firms.

⁵ All estimates are in constant 2006 dollars.

⁶ See <u>http://www.nist.gov/director/prog-ofc/report02-3.pdf</u>.

scientifically based. A 2007 economic study conducted by RTI International on behalf of NIST examined measurement and other technological infrastructure supporting the U.S. biopharmaceutical industry. The study found that improvements to infrastructural technologies in the area of bioimaging, biomarker development, informatics, and gene expression analysis could yield significant reductions in the time and cost associated with drug development.⁷

Specifically, the RTI study estimates that increased investment in technology infrastructure to support the biopharmaceutical industry could reduce the total cost of developing new drugs by between 35 and 48 percent and development time could be shortened by 20 percent. In addition, average annual manufacturing costs could be reduced by 23 percent. These substantial potential cost and time reductions indicate the critical importance of assuring adequate technical infrastructure to support this industry and to achieve the National goals of reduced cost and improve quality of medical care.

Systems Technologies and Interoperability. In addition to the growing demands on technical infrastructure to support development of increasingly complex technology-based products and services, a second major source of need has arisen. Today's technologies are increasingly being integrated into higher-level systems, such as combined voice, video and data services, which are delivered over increasingly complex networks. Medical services, traditionally provided by small independent suppliers (doctors) and hence largely a "black-box" technology, today are in desperate need of a much improved interoperable electronic information infrastructure to manage diagnoses, treatment options, and patient records across multiple computer-based platforms.

Beyond the impacts resulting from the systems nature of technologies on companies' internal operations, external market interactions among companies are increasing due to greater specialization and hence greater distribution of R&D and production among industries in high-tech supply chains. This "vertical disintegration" phenomenon increases the number and complexity of market exchanges, which, in turn, creates increased flows of technical and business data among the several tiers (industries) making up these supply chains. The result is a much greater need for an interoperability infrastructure to support these flows.⁸

However, achieving interoperability is technically demanding. Many incompatibilities exist with respect to database formats and the management of information flows over networks. Both hardware and software developed and marketed by various vendors for information management typically have incompatible architectures and formats. When interoperability standards and the supporting infratechnologies are inadequate, companies must expend considerable resources to avoid

⁷ Economic Analysis of the Technological Infrastructure Needs of the U.S. Biopharmaceuticals Industry, November 2007, prepared for NIST by RTI International (http://www.nist.gov/director/progofc/report07-1.pdf).

⁸ Many different definitions exist for the "high-tech" sector of an economy, which is unfortunate for policy analysis purposes. The definition of "the high-tech sector" used here includes both manufacturing and service industries grouped into four major categories: high-tech manufacturing (IT-related plus industrial electronics), communication services, software and computer-related services, and pharmaceuticals.

compatibility problems or mitigate problems after they occur. Moreover, such activities can result in delays with respect to product/service delivery, which adds additional costs.

A series of economic studies by NIST of the cost of inadequate interoperability demonstrate the magnitude of the problem. An initial study in 1999 of just one type of inter-company data flows (electronic product design data) in a single supply chain (automotive) yielded an annual estimate of \$1 billion in excess costs. Subsequent studies in 2004 estimated annual interoperability costs for all business data flows among companies in the transportation, electronic, and construction/building management supply chains to be \$5 billion, \$3.9 billion, and \$15.8 billion, respectively.⁹ Obviously, the total cost for all supply chains making up the U.S. economy is much larger.

These studies show that inadequate interoperability is a source of substantial economic loss. Virtually all firms in the U.S. economy are negatively affected, but for small firms the lack of interoperability can be a life or death matter. Access to targeted markets by these firms increasingly depends on providing not only an innovative product but also including a technically sophisticated interface to other components of the product system. Engineering complex, non-standard interfaces, which often vary significantly among customers, adds both excessive costs and time delays that can be fatal for small firms. The supply chain and the entire economy then lose the innovative capacity of this important segment of U.S. industry.

Even the act of buying a high-tech product requires a technically sophisticated infrastructure. For example, how would a communications services company determine if the thousands of kilometers of ultra-thin optical fiber it needs to buy meet its specifications for such performance attributes as band width, signal loss rates, core diameter consistency, insulation properties, etc. unless technologically advanced acceptance testing standards are available? Without such infrastructure, the high-tech sector of the economy could not grow and prosper. Market transaction costs are substantially higher without adequate supporting infrastructure and time delays can eliminate the "first mover" advantage of innovating firms.

Regulatory Compliance. Technology infrastructure is also essential for assuring compliance with a wide range of regulations implemented to achieve environmental, safety, and other social goals. The cost of complying with such regulations is high for both the regulator and the affected industries, so compliance efficiency is essential. This is especially the case in today's global markets, where such regulations are frequently more lax in competing economies.

An example of NIST's essential contributions to regulatory efficiency is its program of supplying standard reference materials (SRMs) for measuring the content of sulfur in fossil fuels. The 29 sulfur-related SRMs provided by NIST are used by an entire supply chain (instrument manufacturers, independent testing laboratories, industry suppliers of secondary standards, coal processors and petroleum refiners, electricity generating

⁹ See <u>http://www.nist.gov/director/prog-ofc/report04-2.pdf</u> for the transportation and electronics supply chain integration study and <u>http://www.bfrl.nist.gov/oae/publications/gcrs/04867.pdf</u> for the construction industry interoperability study.

utilities, and other major users of fossil fuels such as steel producers) to design and calibrate pollution monitoring and control equipment. A NIST impact study estimated that the net present value of economic benefits to these industries from more efficient equipment calibration and monitoring was \$441 million (2002 dollars). If the average benefits per sulfur SRM are assumed to accrue to all of the approximately 1300 NIST SRMs provided to industry, the total economic benefit (net present value in 2002 dollars) is roughly \$20 billion.¹⁰

The issue of regulatory compliance becomes magnified when international trade is considered. As trade has become more important to most of the world's economies, incentives to create barriers to imports that compete with domestic industries have increased.

For some time, tariffs were the main instrument for modifying relative competitive positions in favor of domestic industries. However, the importance of trade has created pressures to open domestic markets through tariff reductions. At the same time, the increasing complexity of traded goods has afforded opportunities to replace tariffs with a range of non-tariff barriers to trade (TBTs) that are based on technical specifications.

Some technical requirements placed on imports are blatant in their intent to restrain trade. Examples are requiring detailed product data, including proprietary information, to accompany every batch of the same product from the exporting country, or requiring testing of each unit imported (as opposed to sampling or self-certification according international standards).

However, in many cases, assessment of the possible use of technical specifications as trade barriers is complicated for two reasons. First, such specifications often have legitimate uses, such as ensuring efficacy, safety, environment quality, and compatibility with interfacing products. The separation of these uses from trade-restriction objectives within the same regulation is conceptually difficult. Second, although some qualitative information is collected on TBTs, estimation of the economic impacts of these barriers is limited by the lack of uniform and sufficiently precise quantitative data.

Because the social objectives leading to technical specifications being imposed on imported products are typically neither accurately nor uniformly defined across countries, excessive restrictions can easily result. For example, lack of uniform risk assessment procedures can lead to widely divergent performance specifications and testing requirements.

Therefore, TBTs need to be segregated according generic mechanisms and appropriate impact data collected. For example, some barriers impose a one-time cost, such as retooling a manufacturing plant to meet product design or performance requirements imposed by the importing country, while other barriers result in recurring costs associated with certifying each shipment's compliance with a particular regulation.

A NIST-sponsored study conducted assessments of TBTs in the pharmaceuticals and automotive products industries.¹¹ These case studies exemplify the wide range of

¹⁰ See http://www.nist.gov/director/prog-ofc/report00-1.pdf.

¹¹ See <u>http://www.nist.gov/director/prog-ofc/report04-3.pdf</u>.

possible TBTs and also the difficulties in identifying, characterizing and measuring them. The pharmaceutical industry is among the most highly regulated across all countries. Further, the variety of product offerings is great and the typical product, historically complex, is becoming even more so, as biotechnology becomes the driver for innovation. Not surprisingly, a number of regulations and conformity assessment procedures in this industry were found in the study to have TBT-like characteristics. A number of countries write regulations in such a manner that they appear to discriminate against innovative pharmaceuticals for which the United States is the leading supplier. Differential speed of approval intervals and inspection procedures, such as mandated testing for each lot of imported products or requiring certification documentation and other conformity assessment data, were found to have characteristics of TBTs.

Because of the emphasis on automotive imports, the fact is overlooked that U.S. exports of this product category are significant. In 2002, total U.S. automotive exports totaled \$76.5 billion, which was 11.2 percent of total exports. The study found that instances of possible product-related TBTs were largely occurring in emerging economies with young automotive industries. Process-related TBTs were primarily occurring in the repair and service industry. Most of these possible TBTs fall into the conformity assessment category.

The complexity of candidate TBTs makes accurate economic analysis for policy decision making quite difficult. Such analysis requires the development of metrics and measures that reflect both one-time costs and trends over time for barriers that impose repetitive costs. The qualitative analysis in the two case studies along with anecdotal information indicate that the economic costs of TBTs to domestic firms are high.

Effectively estimating these costs not only depends on the availability of an economic analysis capability within government, it also requires a body of technical expertise that can inform the economic impact assessments. The latter expertise is a technology infrastructure function—the same infrastructure that provides the technical expertise necessary for testing imported products to determine conformance with health, safety, environmental, and homeland security objectives.

Nature and Impacts of Infratechnologies and Standards. Infratechnologies contribute to technology-based growth in multiple ways and across the full range of economic activity (R&D, production, and market development). The following are examples of NIST-developed infratechnologies and standards and their primary economic stage of impact (SRM = standard reference material and SRD = standard reference data):

<u>R&D</u>:

Josephson volt standard (basic standard for product development) alternative refrigerants (SRD for product development) ceramics phase equilibria (SRD for process design) thermocouple (SRD for product development) radiopharmaceuticals (SRMs for product development)

Production:

sulfur in fossil fuels (SRMs for calibration of process and monitoring equipment)

laser and fiberoptic power and energy (calibration for process and quality control) cholesterol measurement (SRMs for quality control) alternative refrigerants (SRD for process control) ceramics phase equilibria (SRD for quality and process control) gas mixtures (SRMs for process control)

<u>Commercialization/Market Development:</u>

electronic data exchange formats (interface standards for supply chain efficiency) gas mixtures (SRMs for regulatory compliance/transaction cost reduction) sulfur in fossil fuels (SRMs for regulatory compliance/transaction cost reduction) laser and fiberoptic power and energy (calibrations for transaction cost reduction) thermocouple data (SRD for transaction cost reduction)

Many infratechnologies impact more than one stage of economic activity. For example, as indicated above, SRMs for calibrating pollution control or monitoring equipment are necessary for production efficiency (i.e., achieving regulatory compliance without resorting to operating in more resource-intensive modes to allow for measurement uncertainty), and they also reduce the market transaction cost associated with demonstrating compliance. NIST supplied 29 SRMs for measuring sulfur in fossil fuel to the entire energy producing and using supply chain.

Underinvestment in Technology-Based Infrastructure. Major social objectives such as healthcare, communications, energy independence, security and education all receive major government support. However, achieving these objectives requires several industries in each case to deliver the technology systems that achieve the objective. The efficiency of these technology delivery systems is determined to a large degree by the availability and quality of the supporting technology infrastructures.

Today's technologically complex infrastructures are not receiving adequate levels of resources due to a poor understanding of this infrastructure's roles in long-term economic growth. Several factors contribute to this situation:

- (1) <u>Segmented and ubiquitous nature.</u> Any one infrastructure element and the associated set of standards frequently have modest economic impact, but a high-tech industry depends on hundreds of such elements, so the collective economic impact is substantial. This fact is largely unrecognized and undocumented, except for the economic studies done by NIST.
- (2) <u>Invisibility.</u> The economic infrastructure (roads, bridges, canals, telephone networks) that enabled the Industrial Revolution was physical and large and hence highly visible, thereby facilitating consensus on optimal investment levels. However, as the above examples indicate, most technological infrastructure underlying today's high-tech industries is non-physical (i.e., knowledge-based) and largely invisible even, in many cases, to high-level corporate managers. Measurement and test protocols, scientific and engineering databases, interface specifications, quality assurance techniques, etc. are often delivered through systems of hardware and software and, as methods or procedures, are much more difficult to visualize and understand.

- (3) <u>Reliance on private markets</u>. More than its competitors, the U.S. economy relies on private institutions to provide both the infrastructure and the standards that facilitate deployment. These institutions, such as standard development organizations (SDOs), consortia, and various associations, are typically funded at meager levels. They are also staffed by industry personnel who rotate to these organizations for brief periods of time from their home companies, causing continuity, institutional memory, and other productivity problems.
- (4) <u>Management.</u> The number and variety of infratechnologies needed to support each stage of the technology-based economic process present a difficult portfolio management problem. Thus, even if public-private sector role issues were resolved, setting priorities, conducting research in a timely manner, and deploying the resulting technical information through an efficient standards process and other dissemination mechanisms present difficult challenges.
- (5) <u>Quasi-public good character</u>. Although some of the needed infrastructure is provided by industry, its public-good nature results in "free-riding" problems and consequently leads to substantial underinvestment. Figure 1 illustrates how the various elements of technology-based economic activity interact to produce value added (the industry's contribution to GDP). The arrows indicate the directions of impact of each element. The shaded areas (blue) denote the degree of public-good character and hence the relative need for government support. The non-shaded areas denote private-good content and therefore private investment. The multiple

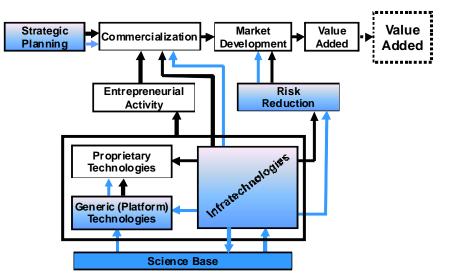


Figure 1 Technology-Element Model

G. Tassey, The Technology Imperative, Edward Elgar, 2007

arrows from the infratechnologies box indicate the ubiquitous role of measurement and other technical infrastructure in the technology-based economy. The fact that this technology element is partially shaded means measurement is what economists call a "quasi-public" good, which therefore requires both industry and government to supply portions of an industry's measurement needs. This clearly complicates assignment of investment responsibility between the two

sectors, which mandates considerable joint role assessment and strategic planning efforts.

Such barriers lead, in turn, to several negative economic impacts that compromise technology-based economic growth:

- (1) Underinvestment in infratechnologies;
- (2) Inadequate resources for the standards setting processes;
- (3) Inappropriate or ineffective standards, such as proprietary rather than open-source standards or simply low-quality standards;
- (4) Poor timing and ineffective deployment of standards relative to market needs, including poor management of transitions to new generations of the same standard.

The last of these impacts, market timing, is increasingly important as growing global competition shrinks technology and product life cycles. Private markets may eventually provide the optimal infrastructure. However, by that time either the technology life cycle being supported becomes obsolete and a new infrastructure is needed, or, worse, foreign competition with its own infrastructure in place takes over the domestic and global markets.

The bottom line is that in an increasingly flat world, the new technology-based growth paradigm resulting from globalization of corporate strategy means that governments must step up efforts to fund and disseminate the right amounts of technology infrastructure and do so in the time frames dictated by global competition. Technology infrastructure, which is part of the broader innovation infrastructure of an economy, takes considerable investment not only in the infratechnologies and standards needed by the domestic industry but also in the institutions that help develop and rapidly disseminate this infrastructure. Such assets are relatively immobile in the global economy sense, so that investment in domestic assets can convey considerable competitive advantage.