The Transportation Research Thesaurus: Capabilities and Enhancements

DETAILS

GET THIS BOOK

FIND RELATED TITLES

CONTRIBUTORS

Gail Hodge and Denise Bedford; National Cooperative Highway Research Program; Transportation Research Board; National Academies of Sciences, Engineering, and Medicine

NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM

NCHRP RESEARCH REPORT 874

The Transportation Research Thesaurus: Capabilities and Enhancements

Gail Hodge
INFORMATION INTERNATIONAL ASSOCIATES, INC.
Oak Ridge, TN

Denise Bedford
GEORGETOWN UNIVERSITY
Washington, DC

Subscriber Categories
Administration and Management • Data and Information Technology • Education and Training

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

TRB
TRANSPORTATION RESEARCH BOARD
2018

NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM

Systematic, well-designed research is the most effective way to solve many problems facing highway administrators and engineers. Often, highway problems are of local interest and can best be studied by highway departments individually or in cooperation with their state universities and others. However, the accelerating growth of highway transportation results in increasingly complex problems of wide interest to highway authorities. These problems are best studied through a coordinated program of cooperative research.

Recognizing this need, the leadership of the American Association of State Highway and Transportation Officials (AASHTO) in 1962 initiated an objective national highway research program using modern scientific techniques—the National Cooperative Highway Research Program (NCHRP). NCHRP is supported on a continuing basis by funds from participating member states of AASHTO and receives the full cooperation and support of the Federal Highway Administration, United States Department of Transportation.

The Transportation Research Board (TRB) of the National Academies of Sciences, Engineering, and Medicine was requested by AASHTO to administer the research program because of TRB's recognized objectivity and understanding of modern research practices. TRB is uniquely suited for this purpose for many reasons: TRB maintains an extensive committee structure from which authorities on any highway transportation subject may be drawn; TRB possesses avenues of communications and cooperation with federal, state, and local governmental agencies, universities, and industry; TRB's relationship to the National Academies is an insurance of objectivity; and TRB maintains a full-time staff of specialists in highway transportation matters to bring the findings of research directly to those in a position to use them.

The program is developed on the basis of research needs identified by chief administrators and other staff of the highway and transportation departments, by committees of AASHTO, and by the Federal Highway Administration. Topics of the highest merit are selected by the AASHTO Special Committee on Research and Innovation (R&I), and each year R&I's recommendations are proposed to the AASHTO Board of Directors and the National Academies. Research projects to address these topics are defined by NCHRP, and qualified research agencies are selected from submitted proposals. Administration and surveillance of research contracts are the responsibilities of the National Academies and TRB.

The needs for highway research are many, and NCHRP can make significant contributions to solving highway transportation problems of mutual concern to many responsible groups. The program, however, is intended to complement, rather than to substitute for or duplicate, other highway research programs.

NOTICE

The research report was reviewed by the technical panel and accepted for publication according to procedures established and overseen by the Transportation Research Board and approved by the National Academies of Sciences, Engineering, and Medicine.

The opinions and conclusions expressed or implied in this report are those of the researchers who performed the research and are not necessarily those of the Transportation Research Board; the National Academies of Sciences, Engineering, and Medicine; or the program sponsors.

The Transportation Research Board; the National Academies of Sciences, Engineering, and Medicine; and the sponsors of the National Cooperative Highway Research Program do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of the report.

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

The **National Academy of Sciences** was established in 1863 by an Act of Congress, signed by President Lincoln, as a private, non-governmental institution to advise the nation on issues related to science and technology. Members are elected by their peers for outstanding contributions to research. Dr. Marcia McNutt is president.

The **National Academy of Engineering** was established in 1964 under the charter of the National Academy of Sciences to bring the practices of engineering to advising the nation. Members are elected by their peers for extraordinary contributions to engineering. Dr. C. D. Mote, Jr., is president.

The **National Academy of Medicine** (formerly the Institute of Medicine) was established in 1970 under the charter of the National Academy of Sciences to advise the nation on medical and health issues. Members are elected by their peers for distinguished contributions to medicine and health. Dr. Victor J. Dzau is president.

The three Academies work together as the **National Academies of Sciences, Engineering, and Medicine** to provide independent, objective analysis and advice to the nation and conduct other activities to solve complex problems and inform public policy decisions. The National Academies also encourage education and research, recognize outstanding contributions to knowledge, and increase public understanding in matters of science, engineering, and medicine.

Learn more about the National Academies of Sciences, Engineering, and Medicine at **www.national-academies.org**.

The **Transportation Research Board** is one of seven major programs of the National Academies of Sciences, Engineering, and Medicine. The mission of the Transportation Research Board is to increase the benefits that transportation contributes to society by providing leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal. The Board's varied committees, task forces, and panels annually engage about 7,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation.

Learn more about the Transportation Research Board at **www.TRB.org**.

COOPERATIVE RESEARCH PROGRAMS

FOREWORD

By Andrew C. Lemer
Staff Officer
Transportation Research Board

The Transportation Research Thesaurus (TRT) is a structured, controlled vocabulary of terms in English, used by TRB and a variety of other organizations to support indexing, search, and retrieval of technical reports, research documents, and other transportation information. The TRT, covering all modes and aspects of transportation, has evolved over a number of years and is continuously being refined and expanded. *NCHRP Research Report 874: The Transportation Research Thesaurus: Capabilities and Enhancements* presents the results of a comprehensive assessment of the TRT's capabilities and strategies for the TRT's future development. The research and its results are applicable to storage, retrieval, and other aspects of managing transportation information generally.

The TRT was initially developed in the 1990s under NCHRP as a tool to improve the indexing and retrieval of transportation information. The thesaurus now covers all modes and aspects of transportation, providing a common and consistent language used by producers and users of information stored in the Transportation Research International Documentation (TRID) Database, the integrated records from TRB's Transportation Research Information Services (TRIS) Database, and the Organization for Economic Cooperation and Development's Joint Transport Research Centre's International Transport Research Documentation (ITRD) Database. The thesaurus is currently also used as an indexing tool for many federal, state, and university information repositories.

The TRT supports indexing, search, and retrieval of research documents, technical reports, and other information about transportation and thereby assists information users by encouraging consistency in language and cataloging of such information across repositories. Additionally, the TRT provides a framework for users who wish to tailor their own information storage and retrieval systems for their specific needs. While TRT refers to "research" in particular, the TRT's application is broader, covering all modes and aspects of transportation. The TRT is curated and maintained by TRB. The TRT web page (http://trt.trb.org/) allows for searching and browsing the thesaurus.

Since the TRT's initial development, the amount of available transportation information, demand for that information, information management technology, and standards of information management practice have grown and changed. There is little doubt the growth and change will continue. The TRT has expanded and evolved to cope with these trends, but TRB staff and other stakeholders have for some time recognized a need for strategic planning and substantial action to ensure that the TRT's effectiveness is maintained. The objective of NCHRP Project 20-109, "Enhancement of the Transportation Research Thesaurus," was to conduct a comprehensive analysis of the TRT's capabilities, explore strategic options for the TRT's future development, and identify early actions to

pursue that development. The research, conducted by a team led by Information International Associates, Inc., entailed interviews with TRT users and other stakeholders, in-depth analysis of the TRT architecture and governance tools and processes, and comparison with international standards of information management practice.

The product of this research, *NCHRP Research Report 874: The Transportation Research Thesaurus: Capabilities and Enhancements*, is intended to be a roadmap for the TRT's future development. The report is aimed primarily at TRB staff, TRT users, and other TRT stakeholders, but the research and its results are broadly applicable to storage, retrieval, and other aspects of managing transportation information.

CONTENTS

Note: Photographs, figures, and tables in this report may have been converted from color to grayscale for printing.
The electronic version of the report (posted on the web at www.trb.org) retains the color versions.

S U M M A R Y

The Transportation Research Thesaurus: Capabilities and Enhancements

Background

The Transportation Research Thesaurus (TRT) is a tool to support the indexing and retrieval of transportation information. The TRT is used by the Transportation Research Board (TRB) and a variety of other organizations to support indexing, search, and retrieval of research documents, technical reports, and other information about transportation. It assists information users by facilitating consistency in language and cataloging information across repositories. The TRT originated as a tool to support TRB's Transportation Research Information Services (TRIS) database. However, if the TRT is to maintain its value to a growing community of stakeholders, it must evolve to improve the user experience, make its services accessible to a broader range of users, accommodate changing transportation technology, adapt to changing international information management practices, and take advantage of advances in information technology.

Research is needed to develop a strategic plan for the TRT's future development and to identify initial tasks that will immediately enhance the TRT and initiate implementation of a strategic plan that will direct the TRT's future development to ensure that the thesaurus continues to serve well the needs of Transport Research International Documentation (TRID) users and the broader transportation community. The objective of this research was to enhance the TRT by determining a vision for the future of the TRT and the requirements for reaching it. The research includes an assessment of the ability of the current thesaurus content, procedures, and products to meet those requirements and the establishment of a strategic plan and short-term actions to achieve the requirements.

Research Approach and Methodology

The research team designed a framework for assessment consisting of five dimensions. The five framework dimensions are

- Dimension A: Thesaurus Content, Management, and Maintenance
- Dimension B: Thesaurus Access and Use
- Dimension C: Thesaurus Governance Processes
- Dimension D: Governance Tools—Guidelines, Standards, Principles, and Best Practices
- Dimension E: Thesaurus Architecture, Functionality, and Protocols.

134 assessment criteria were identified for the dimensions based on established practices determined from peer-reviewed literature and the research team's experience. In addition, as the data collection progressed, an additional 32 criteria were identified as pertinent to the

TRT's current context. Data to support the 166 criteria were collected from five sources, including the following:

- A transformed working copy of the TRT that enabled the research team to generate reports and statistics,
- Stakeholder interviews and feedback,
- Online data collection from transportation librarians (in lieu of interviews),
- Targeted input on information sources and use practices from subject matter experts (SMEs), and
- Focused searching of peer transportation information sources.

The data for criteria were first described and presented in a neutral and non-critical manner. The data for each criterion were detailed and were also supported with examples to facilitate interpretation and understanding. The research team then undertook an assessment of each criterion. The assessment was conducted against the current context. Each criterion was assessed as (1) fully supported in the current context, (2) partially supported in the current context, or (3) not supported in the current context. Each criterion that was not fully supported was then evaluated in terms of whether it could be enhanced in the current context or in a different future context.

Five use case scenarios were developed based on the peer-reviewed literature and gray literature reports of current thesaurus uses and applications. The scenarios take the reader from the improvements that can be made in the current context (Use Case 0) through a series of progressive improvements that end in a community-based designer thesaurus (Use Case 4). The five use case scenarios are the following:

- Use Case 0: Maintain and Enhance TRT Status Quo in TRID,
- Use Case 1: New ISO-Compliant Master Version of the TRT,
- Use Case 2: Enhanced TRT Local and Global Discovery Tools,
- Use Case 3: TRT Automated Classification and Indexing Tools and Services, and
- Use Case 4: Community-Based Designer Applications from the TRT.

With the exception of Use Case 0, the status quo, each step in the progression builds upon the previous scenario and lays out clear choices and decision points for TRB and the broader transportation community. The research team describes the use cases in terms of (1) the framework dimensions that are addressed; (2) the benefits and challenges they offer; (3) the resources and tasks required, including cost categories, human resources, training requirements, and community impacts; and (4) timelines.

As part of the research, the team also conducted a market analysis of thesaurus management software (TMS) options. The research team identified commercial and open source TMS products based on the expertise of team members and a search of the literature. Enhancement to the current TMS was also included as an option. Twelve TMS options were then evaluated against the requirements. In addition, preliminary information concerning the estimated cost, learning curve, staffing requirements, hardware/software environment, and licensing agreements was gathered by reviewing vendor websites and promotional literature and by contacting vendors. This information would be enhanced by information gained from a procurement process.

Findings

In the final assessment, none of the framework dimensions produced an aggregate "fully supported" result. Dimension A, which focuses on thesaurus content, had the highest number of criteria that were "fully supported" (15.55%). Dimensions A (48.88%)

and B (Access and Use) (57.69%) had the largest number of criteria that were "partially supported" or represented variant practices. Dimension C, which focuses on governance processes, had no criteria that were "fully supported" and a high rate of "not supported" criteria (66.66%) due largely to the lack of governance support functionality available in the current application. Dimensions D (Governance Tools) (68.75%) and E (Architecture) (75.00%) had high numbers of criteria that were "not supported" or had no current practices.

In general, the content of the thesaurus and its current use in TRID are areas of strength. However, Governance Tools (Dimension D) includes the scope and coverage of the field of transportation, which is an area of potential enhancement. The architecture of the current TRT is another clear area of weakness because of the constraints it places on improvements and enhancements. However, it may not be possible to enhance the architecture in the current context. The assessment of Dimension E focuses on the architecture of the thesaurus application and not the larger TRID system. The application supporting the TRT is designed to function as a search component rather than as a thesaurus management tool.

In this context, a key requirement for any TMS is International Organization for Standardization (ISO) compliance, which provides term-level records with explicitly labeled standard relationships. All commercial products evaluated as part of this project are ISO compliant with respect to the relationships that are supported; only the current TMS for TRB's TRT is not fully ISO compliant. ISO compliance is an overarching requirement for enhancing the TRT since it sets the stage for not only rectifying some of the current gaps in the TRT's capabilities but also is a requirement for any enhancements and advanced use cases, including Use Case 2 and beyond. ISO compliance brings with it a number of functional requirements regarding the standard relationships that are met by most products. It should be noted that the current TRT system can perform some of these functions even though it is not fully ISO compliant.

Beyond the compliance requirements, the differentiators among the commercial TMS products are non-functional requirements such as cost (including recurring costs), staffing, training, and licensing. There are relatively inexpensive single-user desktop products that are viable alternatives, particularly for Use Case 2 where there is limited integration with other systems or users. It would be feasible to start small and grow into higher-end products when needed. This decision requires an assessment of the degree to which the TMS will be integrated with external governance processes, such as reviews by the TRT Subcommittee or SMEs. The more distributed the involvement, the more likely the TRT is to need a network/web-based TMS or at least some way to publish updates or extracts of the TRT to the web for review. High-end products have many features that can take a significant amount of time to learn and implement. However, these products can also be used in very low-level approaches to perform the majority of the tasks that are needed, with the option for future growth by exploiting the more advanced features. Cloud-based TMS as a service option is becoming available. While these options eliminate some hosting concerns, they must be reviewed carefully for budget impacts, for their ability to support future growth and development, and licensing constraints.

Many of the decisions regarding the choice of a TMS product for the TRT will be subjective. Therefore, the research team encourages those who will be using the products under consideration to participate in demonstrations, take advantage of trial versions, and participate in webinars hosted by vendors.

Finally, the decisions and actions moving forward on the TRT are dependent on the strategic futures that are addressed in TRB's Strategic Plan. Key to enhancing the TRT and broadening its use is the importance of designing an innovative architecture that does not

adversely impact the TRT's use in TRID, but also supports the full functionality of a 21st century thesaurus management application.

Conclusions

As a result of the detailed criterion descriptions and assessments and the comparison to current and future use case scenarios, the research team offers the four general conclusions:

- **Conclusion 1.** The TRT's future value as an information management tool with application beyond indexing of TRID depends fundamentally on achieving Use Case 1. Unless this transformation is accomplished, there is no future scenario for the TRT beyond its current state. It may not be able to compete with other knowledge organization systems that may become available on the open source or commercial markets. Transformation to an ISO-compliant thesaurus and thesaurus management application is achievable as part of a potential Phase 2 of this research project (NCHRP Project 20-109).
- **Conclusion 2.** As a result of assessing TMSs against the requirements for the TRT identified by the research team and a comparison of additional improvements detailed in the use cases, the research team offers these observations regarding the future of the TMS to support the TRT:
  - Final decisions about the appropriate TMS are closely linked to strategic decisions regarding the future of the TRT, as represented by the use cases.
  - If Use Case 1, the ISO-compliant TRT, is the desired approach, the next step is to review the optional requirements identified during the TMS market assessment to determine which requirements the TRT Subcommittee and the TRT staff consider to be mandatory.
  - The initial focus for TMS procurement can be made from the products on the working list. The majority of these products satisfy the requirements for Use Case 1.
  - Non-functional factors such as initial and recurring costs, technical platform, staffing, licensing, usability, and training will be major factors in the final decision process.
  - If the decision is to move more in the strategic direction represented by Use Cases 2 through 4, it will be important to identify additional requirements to support the more advanced functionality. Each forward-looking use case brings additional requirements or modifications to current requirements.
- **Conclusion 3.** If an investment in Use Case 0 (the Status Quo) is the only possible option, the research team suggests that TRB and the transportation community limit the investment to tasks that will contribute to interoperability and correction of non-standard practices.
- **Conclusion 4.** Engage the transportation community, in particular, state departments of transportation, research institutes, university technology centers, and transportation libraries in drafting a long-term strategy based on a consensus across the field. This could be included as a Phase 2 task for this research project (NCHRP Project 20-109).

Background

The Transportation Research Thesaurus (TRT) is a tool to improve the indexing and retrieval of transportation information. The TRT is used by the Transportation Research Board (TRB) and a variety of other organizations to support indexing, search, and retrieval of research documents, technical reports, and other information about transportation. The TRT also assists information users by facilitating consistency in language and cataloging information across repositories. At the same time, the TRT provides a basis for users who wish to tailor their own information storage and retrieval systems for their specific needs. While the TRT applies to "research" in particular, the TRT's application is broader, covering all modes and aspects of transportation and thereby providing a common and consistent language for communication among producers and users of transportation information. The TRT is curated and maintained by TRB with input from the TRT Subcommittee. The TRT web page—http://trt.trb.org/—allows for searching and browsing the thesaurus in various formats.

The TRT originated as a tool to support TRB's Transportation Research Information Services (TRIS) database, as described in *NCHRP Report 450: Transportation Research Thesaurus and User's Guide* (Batty 2001), enabling indexers to describe documents in a consistent way and TRIS users to more easily retrieve relevant TRIS records in their areas of interest by searching the thesaurus terms. The TRIS database subsequently was consolidated with the Organization for Economic Cooperation and Development's Joint Transport Research Centre's International Transport Research Documentation (ITRD) Database into Transport Research International Documentation (TRID), the world's largest and most comprehensive bibliographic source on transportation information. TRID contains more than one million records of published information and ongoing research, covering all modes and disciplines of transportation. The TRID database is available on TRB's website at http://trid.trb.org.

If the TRT is to maintain its value to a growing community of stakeholders, it must evolve to improve the user experience, make its services accessible to a broader range of users, accommodate changing transportation technology, adapt to changing international information management practices, and take advantage of advances in information technology. For example, the TRT's functionality should ultimately facilitate (a) straightforward mapping between the TRT's terms and those employed in the thesauri and other information management tools of other organizations; (b) implementation and maintenance of scalable and robust standards-based thesaurus management software (TMS); (c) an export function to enable users to adapt the TRT to meet their specific needs; and (d) use of the TRT by a wider range of producers, users, and curators of transportation information.

Research is needed to develop a strategic plan for the TRT's future development and to identify initial tasks that will immediately enhance the TRT and initiate implementation of

5

the strategic plan. The strategic plan will direct the TRT's future development to ensure that the thesaurus continues to serve well the needs of TRID users and the broader transportation community.

Research Objectives

The objectives of this research were to (a) develop a strategic plan for the TRT's development and (b) identify a specific set of actions to enhance the TRT and initiate implementation of the strategic plan.

Research Strategy
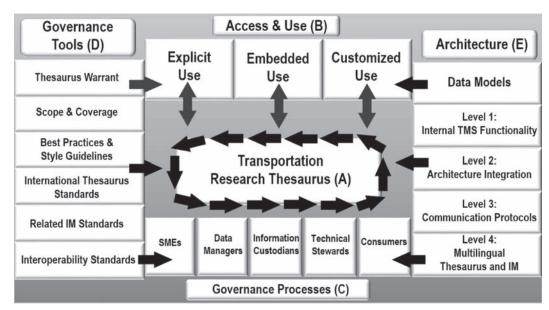
The TRT of the future must be founded on a well-executed assessment of the current TRT and the transportation community's vision of the role it plays in the future information environment. The assessment of current capabilities and potential enhancements included a critical review and assessment of the TRT, identification of opportunities for improvement, and proposed areas of enhancement.

Conceptual Framework

To organize the assessment, the research team developed a conceptual framework for a knowledge organization system that includes five dimensions, A through E. Figure 1-1 shows the dimensions, their detail, and interactions among them. Each dimension is further explained in the subsequent sections.

Dimension A: Thesaurus Content, Management, and Maintenance

The current version of the TRT includes terms, semantic relationships, ongoing day-to-day thesaurus management and maintenance tasks, and professional competencies and skills

IM = information management

Figure 1-1.   Dimensions of a knowledge organization system.

directly relevant to thesaurus management and information management. This dimension is the heart of the framework—without this component, there is no thesaurus. The TRT is a tool to consistently categorize research documents in TRB's TRID database and assist TRID users in retrieving relevant information as described in *NCHRP Report 450: TRT and User's Guide* (Batty 2001).

Dimension B: Thesaurus Access and Use

There are three categories of thesaurus access and use: explicit use (e.g., where the thesaurus is used directly by humans), use where it is embedded in another application (e.g., internal indexing, search system, recommender engine, dictionary or glossary term lookups from within a document, website, etc.), and custom use (e.g., based on specialized requests for use of parts or aspects of the thesaurus). Custom use designates products and services that may be generated to fulfill specialized requests. Each kind of access and use has implications for usability and for the architecture that supports the thesaurus. Failing to consider different kinds of access and use can lead to a suboptimal and constrained design, which, in the long run, may require fixes at additional cost or may prevent the thesaurus from supporting future opportunities.

Dimension C: Thesaurus Governance Processes

Standards in the field of information science are intended to provide guidance but will always be interpreted and adapted to the environment. Organizations can look to International Organization for Standardization (ISO) 25964 only for general guidance on governance, and the governance process for the TRT should be designed to support the transportation community. Governance is all about the community and how the community works. Good governance is a balancing act between getting something done and ensuring that the full community is engaged.

Dimension D: Governance Tools—Guidelines, Standards, Principles, and Best Practices

Governance guidelines include all of those standards, principles, best practices, and warrants that are used in the governance and assessment processes. This dimension is important because at its foundation it guides the day-to-day management and maintenance of the thesaurus, guides design tasks, and ensures the trust and surety of the thesaurus.

Dimension E: Thesaurus Architecture, Functionality, and Protocols

The architecture and functionality of the thesaurus are important because they define what capabilities a user has to work with, what a user can and cannot do, and what products and services a user can produce. There are four levels of architecture to consider: (1) the architecture of the TMS application, (2) the architecture in which the thesaurus must function, (3) the communication architecture to support interoperability, and (4) the architecture required to support multilingual applications.

Architecture Level 1.   It is important to have a clearly defined data model for terms, their attributes and semantic relationships, and any elements needed to manage and organize the terms and vocabularies. For example, the current TRT defines facets through the extensive use of top terms and hierarchical relationships among terms. The current TRT is also supported by an enumerated classification scheme, which provides a unique identifier for each

term and controls the hierarchical display. This is traditional thesaurus architecture. For example:

Water transportation (Aes)

> *Shipping* (Aesb)

> > *Ocean shipping* (Aesbb)

> > *Coastwise shipping* (Aesbc)

> > *Short sea shipping* (Aesbe)

In 2017, thesaurus architecture is usually grounded in a database structure where every term is represented by a full record, with administrative, access, and preservation fields. Architecture Level 1 also includes centralized versus federated support for thesaurus management functions, including permissions and version control. Additionally, Architecture Level 1 addresses issues of localization and collaboration. Some TMSs are designed to inherently support the creation and management of local vocabularies. These designs include all of the standard functionality for record locking and version control, as well as utilities for uploading, error checking, and reporting to manage exceptions. Note that the localization of vocabularies is different from micro-thesauri. Micro-thesauri are part of a larger macro-thesaurus structure and are created and maintained through the use of categorization or classification schemes (Cambridge Systematics, Inc. 2013).

Architecture Level 2.   The second kind of architecture depends on how the thesaurus is used within the organization and the level and nature of integration with other applications. This architecture may take the form of tight integration or loose coupling, depending largely on the type and extent of consuming applications. Generally, business rules for use and implementation of the thesaurus are addressed at this level.

Architecture Level 3.   This level is generally defined by communication protocols such as the Simple Knowledge Organization System (SKOS) or Resource Description Framework (RDF) that allow an organization to identify and define the structure and elements of their knowledge organization system for consumption by other organizations. The SKOS may also be used as an output format for consumption by other applications. The SKOS, though, should be defined to support the data model of the thesaurus. Architecture Level 1 of the thesaurus should not be defined by the SKOS, but Architecture Level 3 should have the capability to export the thesaurus or any part of it in SKOS format.

Architecture Level 4.   Architecture Level 4 pertains to the structures required to support multilingual content and applications. The requirements for displaying records in a multilingual thesaurus are different from those that are required to support multilingual search and machine translation.

This conceptual framework identifies at a high level the key challenges and opportunities that are important to delivering a successful result for the TRT community. Because the TRT has many stakeholders who play a variety of roles, the conceptual framework also provides a holistic and multifaceted view of the various dimensions of the thesaurus. A common understanding across the community is essential for making practical and wise decisions about the future of the TRT.

Assessment Steps

With this framework as background, the assessment of the current TRT was conducted in three steps:

- **Step 1: Formulate assessment criteria** in consultation with the project panel and the transportation community,
- **Step 2: Assess the current state of the TRT against assessment criteria**, and
- **Step 3: Identify gaps and opportunities for improvement and enhancement** in the context of the TRT vision.

Each step in this critical task had its own detailed methodology. Once completed, these steps were synthesized and structured into a draft report for review by the project panel. The three steps are discussed below.

Step 1: Formulate Assessment Criteria

In order to provide a foundation for understanding the current state of the TRT, the research team identified assessment criteria through a review of the peer-reviewed literature, the gray literature, and in consultation with the research panel members. The research team then aligned each criterion with a dimension of the conceptual framework. In addition, the criteria were grouped by impact categories within each dimension to facilitate assessment. The research team identified 166 criteria across the dimensions. The full list of criteria organized by dimension is provided in Appendix C.

Step 2: Assess the Current State of the TRT Against Assessment Criteria

The assessment against the criteria developed in Step 1 was completed through two activities. The first activity focused on a detailed description of the TRT based on the framework and criteria. The research team realized early in the process that a consistent, complete, and comprehensive understanding of the TRT was important as a foundation, prior to undertaking a critical assessment. This description represents a non-critical, fact-finding approach.

The second activity involved the critical assessment of the current state of the TRT. The assessment was grounded in the criteria-based description from Step 1. The critical assessment, presented in Chapter 3, discusses overall strengths and weaknesses of the TRT, based on each dimension of the conceptual framework.

Step 3: Identify Gaps and Opportunities for Improvement and Enhancement

Unlike the description and critical assessment of the current state of the TRT, which were straightforward tasks grounded in data and information, defining future opportunities presented challenges. The opportunities must be practical and affordable and be designed in such a way to avoid loss of functionality or disruption to existing applications. The opportunities must build value and capacity incrementally and be achievable over time. These opportunities must also be designed to accommodate future changes in the larger transportation field and in information and semantic technologies. The decision points and alternatives must be clearly specified for those opportunities. To meet this challenge, the research team identified five use case scenarios.

The five use cases were defined to progressively add value over time and to provide decision makers with options for defining the TRT's future. Use Case 0 represents the status quo. Use Case 1 focuses on the creation of a new master version of the TRT, which is standards based and will expand engagement with and use by the transportation community. Use Case 2 further develops the new master version of the TRT as an important information discovery tool for the field of transportation. Use Case 3 explains how the TRT could be developed to support both centralized and local automated categorization and indexing. Finally, Use Case 4 describes a designer version of the TRT that serves as the foundation for future tools and capabilities,

including those developed by the broader transportation community. The use case scenarios are presented in more detail as future strategies in Chapter 4.

Sources of Data and Information for the Assessment

The research team drew upon three major sources of data and information to complete the assessment: (1) a working copy of the TRT, (2) interviews and online data collection from representatives of TRT stakeholder groups, and (3) focused searching of comparable transportation information sources. In addition, the assessment, findings, and conclusions were informed by a detailed analysis of the TRT Facet X (Information Organization), from development of 50 definitions for terms already in the TRT, and from a market assessment of TMS. These "deep dive" tasks provided an opportunity for the research team to assess and evaluate aspects of the TRT in more detail.

Working Copy of the TRT

The research team began the project with access to the current web version of the TRT, all of the supporting documentation created by the TRT management team, and access to the administrative interface of the TRT. While these were rich sources of information for the assessment, the research team needed a standard thesaurus management and reporting tool in order to complete the assessment. For this reason, the research team set out to generate a working copy of the TRT in a thesaurus management tool. Having a full and complete copy of the thesaurus was necessary to conduct the assessment and provide all stakeholders with an accurate picture of both the TRT's current state and possible future strategies.

The TRT team provided several sources for the research team to work with to create the working copy. However, several challenges were encountered in the generation of this working copy, and it became clear that the only way to gain an accurate representation of all TRT relationships was to capture and reconstruct full records at the individual term level.

Figure 1-2 presents a snapshot of a full term TRT record reconstructed and transformed into a record compliant with American National Standards Institute (ANSI)/National Information Standards Organization (NISO) Z39.19 and ISO 25964. It includes the TRT components that serve as the source for the reconstructed and transformed record.

With a fully ISO-compliant thesaurus available in an application with extensive search and reporting capabilities, the research team was able to generate the data needed to address a large number of the assessment criteria.

Stakeholder Interviews and Feedback

To gain a better understanding of the current TRT environment, as well as challenges and opportunities, the research team conducted interviews with five stakeholder groups: (1) department of transportation (DOT) staff who have critical roles pertinent to the TRT; (2) transportation subject matter experts (SMEs) who are the primary intended users of the TRT; (3) the TRT administrative staff, TRB project panel and the TRT Subcommittee; (4) the TRT system architects and administrators; and (5) information science professionals in state transportation department libraries, academic transportation libraries, and transportation knowledge networks.

The research team used its network of contacts and input from the project panel, subcommittee and TRT staff to identify representatives for these five groups. The SMEs were selected based on expertise in each of the 21 high-level facets of the TRT.

Display the term, Shipping (Aesb), in TRT where all relationships are constructed based on the enumerated hierarchy (Aes**) and ISO compliant term record with full and explicit relationships to all other terms, and where the hierarchy is presented as a categorization value.
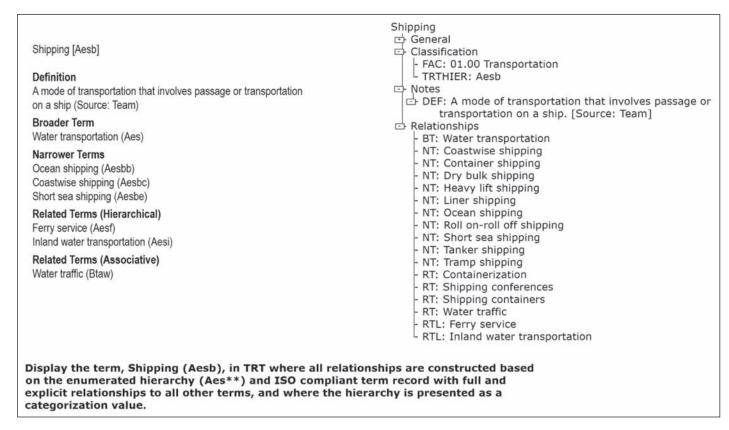
Figure 1-2.   TRT hierarchical structure compared to ISO 25964 structure.

An interview guide (see Appendix D) was developed and distributed prior to each interview. This guide included a general section about the purpose and protocol for conducting the interviews, followed by specific sets of questions for each interview group. In some cases, a particular individual represented multiple groups, in which case the question sets were combined. The questions served as a preliminary set of questions to guide the interviews. For each interview, the research team recorded information in a distinct document and, with permission, recorded the interview, if feasible. Each document was filed and archived as part of the research effort. In addition, reference materials shared by the interviewee were logged and filed in the research archive.

The research team conducted a total of 26 interviews, including five members of the NCHRP Project 20-109 Panel and key DOT stakeholders, nine state DOT information professionals and academic transportation librarians, eleven SMEs, and the TRT system administrator. The information gathered through interviews supplemented the quantitative assessment data generated by the research team working directly with the TRT.

The research team conducted hour-long, virtual interviews with several state and academic transportation librarians and managers. Several attempts were made to contact librarians in other states. When these states were unresponsive to the invitation, the research team shifted its strategy to an online data collection form (see Appendix D). The online data collection form was distributed to 81 state DOT, academic, and private organization transportation librarians. Fourteen responses were collected. These responses provided a level of feedback the research team hoped to achieve through the interactive interviews. This approach was also used to collect additional information from SMEs. Sections 6 and 8 of the

in-person interview guide for SMEs were transformed into an online data collection form (see Appendix D).

Focused Searching of Peer Transportation Information Sources

A significant number of the criteria in Dimension D (Governance Tools—Guidelines, Standards, Principles, and Best Practices) deal with the alignment of the TRT with the scope and coverage of the field of transportation. In addition to the transportation research found in the National Academies of Sciences, Engineering, and Medicine information repositories, transportation research is published in peer-reviewed journals, technical reports, and gray literature, accessible through open source search systems such as Google Scholar, as well as in commercial databases. The research team conducted two types of tests to gauge the alignment of the TRT with the field. The first type was a general search of terms in glossaries, manuals, and encyclopedias referenced by the research team's Advisory Panel of SMEs for use in the TRT assessment. The second type of search involved searching the TRT terms in other transportation databases and publications to determine their productivity and coverage.

Review of Previous NCHRP Reports

In addition to interviews and focused searching for terms, the team reviewed previous NCHRP reports that included assessments of or references to the TRT. All of these earlier projects and reports were leveraged in this research, which is the first comprehensive assessment of the TRT.

The TRT was initially developed under NCHRP Project 20-32 (2001) to provide a tool to improve the indexing and retrieval of transportation information. The project produced a machine application to replace the previous paper version of the TRT, updated the structure and content, and created a user's guide to support the implementation and use of the TRT. The initial project leveraged the Viewer software developed by CDB Enterprises in 1994/1995. While the original design of the TRT was intended to be ANSI/NISO Z39.19 compliant, the use of the Viewer software and its limited application as a component of search resulted in a suboptimal thesaurus application.

In 2007, NCHRP Project 20-70 was undertaken to develop a database at TRB that would allow TRB to maintain the authoritative version of the TRT in-house (i.e., to eliminate the external contractor role), to integrate the TRT with the then-new TRIS system for creating catalog records, and to publish the TRT on the Internet. Seven specific tasks were identified and completed. This project resulted in the current version of the TRT (http://trt.trb.org/trt.asp?). When NCHRP Project 20-70 was launched, the architectural design was guided by the TRIS database. Alignment with ANSI/NISO Z39.19 and/or ISO 25964 standards was not undertaken. These standards were not provided to the contractor as design specifications.

In 2008, NCHRP Project 20-79 was launched. This project focused on the addition of AASHTO definitions to the TRT and was completed in 2009. NCHRP Project 20-79 was implemented in the digital version of the TRT, a component of TRIS. While the project resulted in the integration of some existing definitions, a comprehensive strategy for the use and inclusion of definitions was not developed.

Since 2009, two additional NCHRP research projects have addressed or mentioned the TRT, NCHRP Project 20-90 and NCHRP Project 20-97. NCHRP Project 20-90 was launched in 2011 and completed in 2013 by Cambridge Systematics. The objective of this research was to prepare guidance describing the practices that state DOTs could use for capture, preservation, search,

retrieval, and governance of transportation data and information, as well as strategies and actions a DOT could follow to implement such practices. As part of this broader project, there was a small targeted review of TRT as a cataloging tool. While some suggestions were offered to enhance the TRT, the goal was not to provide a comprehensive review of the thesaurus. NCHRP 20-97 referenced the TRT as a tool that was designed to support findability of transportation information. The goal of this project, though, was to improve the findability of state DOT information by defining an information management framework, describing good practices for organizing and classifying information, and developing federated search procedures. The TRT was mentioned but was not treated or leveraged in depth for this research project. Since 2011, the state DOTs have also provided feedback and comments on the TRT, in particular feedback pertaining to the scope and coverage of the TRT and the general usability and ease of access to the resource.

CHAPTER 2

The Transportation Research Thesaurus—Historical and Current Context

The goal of this research was to describe and assess the current state and future potential of the TRT. However, the TRT has a long history, and it is important to understand when and why it was created and how it has evolved to its current form. This history is best understood in the context of the development of thesauri, TMS tools, and the use of thesauri since the 1950s. The research team identified four generations in the history and evolution of thesauri, thesaurus architectures, and technologies.

History of Thesauri and the TRT

In Generation 1, the thesaurus was in book form for manual perusing. In Generation 2, it became a component of information management systems. In Generation 3, stand-alone TMS applications were introduced. During these first three generations, a thesaurus was understood to be a tool for information professionals in subject analysis, indexing, and search. These early generations aligned with the early and progressive developments of technology and the increasing but limited access of the public to those technologies.

The fourth generation saw the impact of semantic technologies and artificial intelligence. In this generation, thesauri are seen as critical components of semantic architectures and as tools that stand on their own and are consumed by a larger community to navigate the new semantic world.

Where does the TRT fit into this historical context? The TRT was originally developed from 1993 to 1996, prior to the introduction of standards such as ISO 25964 (see Figure 2-1). The TRT was born in Generation 2, as the need for vocabulary management grew in the field of transportation.

As originally developed, the TRT was an elegant design, well suited to its purpose. The TRT was designed by one of the foremost experts in the field of thesaurus management and vocabulary control, Dr. David Batty. The TRT was designed to organize information into broad classes that could be searched-browsed in the TRID repository—similar to subject headings in library catalogs. Because TRID was developed in proprietary software, so too was the TRT component. In 2008, the TRT indexer interface and thesaurus management interface were embedded into the TRID application. In this context, they resemble the cataloger's interface in an integrated library system. This approach puts the TRT firmly in Generation 2 as a component of another information system.

Thesaurus Standards and Conceptual Model

How does the TRT align with the current semantic world? The conceptual reference model derived from the ISO 25964 standard for thesauri (see Figure 2-1) suggests that there are three essential components of a standards-compliant thesaurus: ThesaurusConcept, ConceptGroup,
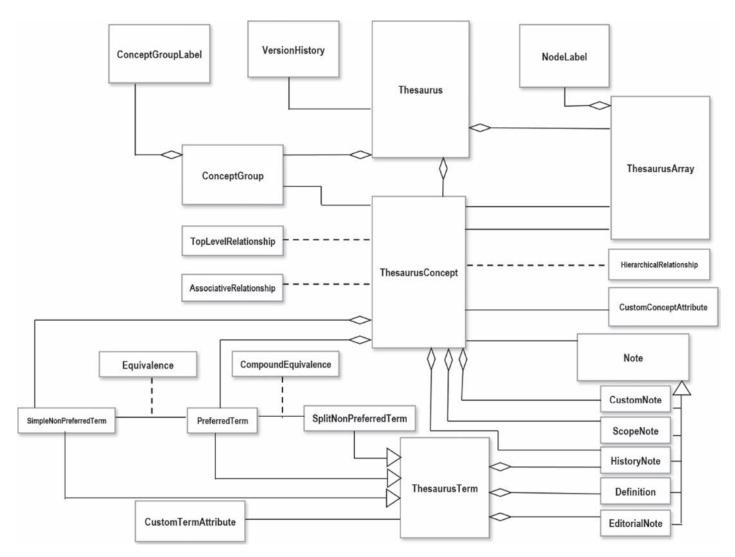
Figure 2-1.   Model derived from the ISO 25964 conceptual reference model.

and Relationships. These components translate to thesaurus term, facet, or category, and fully specified thesaurus relationships.

In the thesaurus standard, the central and essential element is the concept or term. Each term has a fully specified record that is used to manage that term. Designing around individual terms enables flexible definition of reports, search capabilities, multiple browse structures, the generation of Key Word in Context (KWIC) and Key Word Out of Context (KWOC) structures, and management of the thesaurus at multiple levels. This architecture also enables the assignment of terms to multiple "groups" or categories, which provides significant functionality, particularly for transdisciplinary domains such as transportation. Each term is linked to other terms through the specification of relationships. Relationships are structures that are attached to thesaurus terms. This model provides significant flexibility in representing the language of a domain.

Areas of Variance from Thesaurus Standards

How does the TRT vary from the ISO standard? The TRT varies from the standard conceptual model in that the hierarchical relationships represented in the enumeration scheme are neither a ConceptGroup nor are they fully specified hierarchical relationships linking terms. A

ConceptGroup may be created by extracting all of the terms that are assigned to a TRT alphabetically enumerated facet. However, this is not an existing feature of the TRT. Instead, the TRT does not have fully specified hierarchical relationships assigned to individual terms. Hierarchical relationships are interpreted on the fly for the TRT display from a translation of the alphabetical enumeration scheme.

The TRT as a Component of TRID Search

Why does the TRT vary from the ISO standard? The TRT varies from the standard because it was designed as a component of a search engine. Figure 2-2 describes the TRT as a schema that is part of the larger TRID search system. The conceptual model of a search system includes the following basic elements: bibliographic records, metadata repository, controlled reference structures, full-text index structure, hybrid index structure, attribute-based term index, query-processing algorithms, query-matching algorithms, search results display, and a search interface. Controlled reference structures help the searcher to select the "right" form of the term to search if they are searching by specific fields. In the TRID example, the TRT is a tightly integrated controlled reference source that supports attribute-based term searching. The TRT's primary function as a support to index-based searching explains why the TRT schema refers to index records and components rather than to thesaurus standard components. This context means that the focus is not on term records and their full functionality but on the specific use of those terms in metadata records to support TRID.
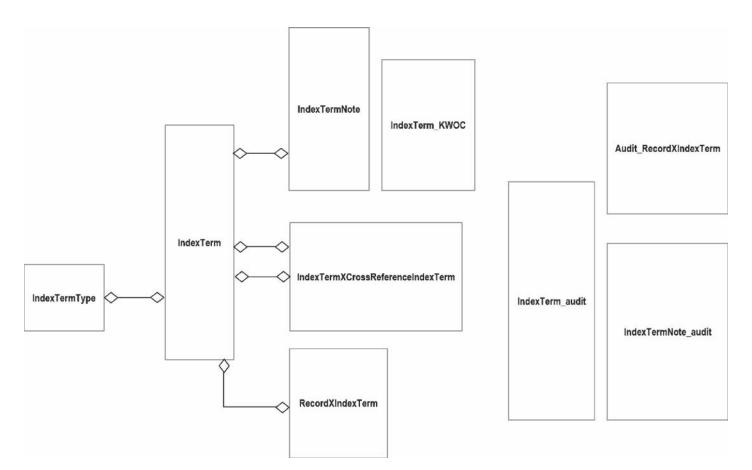
Figure 2-2. The TRT database schema.

Within the context of TRID, the TRT terms are linked through a well-developed and designed search index architecture. This is an important linkage point for thesaurus term records, including the linking of preferred (enumerated) and non-preferred (non-enumerated) terms. While this structure makes very good sense from a search perspective, it presents challenges for thesaurus managers and users. From a standards perspective, this means that the focus is not on term records and their full functionality but on the specific use of those terms in the search index. Although the TRT's database architecture may be efficiently designed from a search perspective (see Figure 2-2), this design does not support management of the TRT or its use in other applications. The design also presents a development environment that would be costly to reconfigure in order to extend the thesaurus management capabilities that are currently available in commercial thesaurus management applications.

The architecture and the functionality of the TRT are not consistent with Generation 3 commercial TMS applications. In addition, the TRT architecture imposes constraints on content that prevent the TRT from complying with current national and international thesaurus standards. These constraints present challenges for the TRT moving into Generations 3 and 4. Given the evolution of the external information and transportation environments, the TRT is now suboptimal in terms of scope and coverage, technology, governance, and architecture design.

Despite their differences, the ISO model and the current TRT schema are not mutually exclusive. The goal of the TRT as a controlled reference source in TRID can be accomplished by creating an ISO-compliant TRT. The elements needed to enable TRT-guided search in TRID would be effectively supported by an ISO-compliant thesaurus, and an ISO-compliant thesaurus would enable the TRT community to move content and use into Generations 3 and 4.

CHAPTER 3

Assessment of the Current TRT

This chapter describes and assesses the current state of TRT content. In a well-formed the-saurus, the expectation is to find terms that align with good practices and are, at a minimum, well suited to indexing and searching. Another expectation is to find terms that are at various stages of approval and of active or archived status and to find a full set of thesaurus relationships defined and applied in a way that reflects their use and relevance to the subject domain. Thesaurus content also includes the metadata that are required to manage thesaurus terms throughout their lifecycle, including source and origins (S&Os), definitions, scope notes (SNs), unique iden-tification numbers, notations of changes in status, and so on. In a multidisciplinary field like transportation, a well-formed, overarching structure is also expected in a well-formed thesaurus.

TRT Content, Management, and Maintenance

Using the reconstructed thesaurus as the main data source, the research team assessed the TRT content using 46 criteria. The criteria ranged from the number of terms per facet, to the numbers and types of relationships between terms, and the treatment of various types of terms such as homographs and adjectives. Particular attention was paid to the non-standard Related Term (Hierarchical) relationship, which is unique to the TRT and is intended to support the display of the terms. The occurrence of term-related metadata such as definitions and SNs was assessed, including metadata that support the management of the lifecycle of a term from rec-ommendation through deprecation. Through this detailed assessment, the research team also identified possible errors in the content. Finally, the research team analyzed the overall structure of the TRT, focusing on the balance of terms by level across the facets.

Thesaurus Terms

The total number of terms is identified as enumerated terms versus non-enumerated terms (see Table 3-1). Enumerated terms have codes that place them into the browse display hierarchy.

The non-enumerated terms are non-preferred terms. The number of non-preferred terms varies across facets. A common benchmark for non-preferred terms in a thesaurus used for information discovery, automated classification, and automated indexing is between 15% and 20%. This means that 15% to 20% of the non-enumerated terms are true synonyms, quasi-synonyms, lexical variants, language and spelling variants, common misspellings, initialisms, acronyms, or abbreviations. From a numbers perspective, many of the facets in the TRT appear to achieve this benchmark while others fall short (e.g., G [Testing], N [Organizations], T [Disciplines], U [Mathematics], V [Areas and Regions], and W [Time]). This is discussed in more detail in the next section, Thesaurus Relationships.

Total Terms All Facets	Enumerated Terms	Non-Enumerated Terms
12,125	9,591	2,534

The assessment also found a high incidence of qualified terms. Parenthetical qualifiers are used to provide context and distinction for terms with a meaning that may be ambiguous to users. This is particularly prevalent in the TRT to disambiguate homographs, terms with the same spelling but different meanings. An example would be Accelerators (Concrete), Accelerators (Materials), and Accelerators (Devices). The analysis found fairly heavy use of parenthetical qualifiers across the facets, although they are most heavily concentrated in Facets P [Facilities], Q [Vehicles and Equipment], R [Materials], S [Physical Phenomena], and T [Disciplines]. The preferred practice is to use a qualifying adjective and noun to represent the concept and to link the concept using thesaurus relationships rather than use parenthetical qualifiers. For example, the more standard thesaurus record would be the following:

Concrete accelerators

UF: Accelerators (Concrete)

BT: Materials accelerators

Another area of concern raised by this portion of the assessment is the use of container or structural terms. Structural terms are terms that represent empty classes in the structure of a thesaurus. They are designed to function as an entry point for a further breakdown by characteristic within a hierarchy. In a standard thesaurus structure, relationships are defined to represent such classes. For example, "xxx by type" would be defined as Narrower Term Type and applied as a formal relationship between the main term and the terms that represent types. The TRT design does not allow the management team to develop these types of relationships. During the conversion of the working copy, the research team found a significant number of structural terms used across the facets, with particular concentrations in Facets P [Facilities], Q [Vehicles and Equipment], R [Materials], S [Physical Phenomena], D [Communication and Control], and U [Mathematics].

The research team also found instances of stand-alone adjectives and adverbs (e.g., present, daily, weekly, etc.). These instances appear to be particularly prevalent in Facet V [Areas and Regions]. The use of stand-alone adjectives as search or indexing terms in repositories other than the TRB's will produce unintended and potentially irrelevant results and is not considered best thesaurus practice.

The specificity of terms varied by facet, with the largest number of specific terms in the facets with the most terms. This is not surprising, but does result in an uneven, unbalanced thesaurus.

Thesaurus Relationships

Hierarchical relationships define links between broader and narrower terms. In a standard thesaurus, these relationships are represented as BT/NT. They may also be defined at a more granular level to represent Whole/Part, Concept/Type, or Concept/Instance relationships.

No formal hierarchical relationships are specified in the TRT. Rather, a hierarchical structure and display is automatically constructed using the alphabetical enumeration, as previously shown in Figure 1-2. In the construction of the working thesaurus, the research team reconstructed the enumeration structure into formal BT/NT relationships. Explicit hierarchical relationships are important if the TRT is to be used in any context beyond TRID and other related TRB resources that now utilize the TRT.

In thesaurus standards, associative or related term relationships link terms that are related semantically or conceptually, although not in an equivalent or hierarchical way. Associative

links may include, but are not limited to: (1) sibling terms that are loosely associated and sometimes interchangeable; (2) terms that are linked by familial or derivational relationships; (3) disciplines or fields of study and the objects studied or its practitioners; (4) operations, processes, and their agents and instruments; (5) processes and their counteragents; (6) actions and their results or processes and their products; (7) actions and their targets; (8) objects or substances and their properties; (9) concepts and their causes; and (10) concepts and their units of measure. In the master version of the TRT, instances of associative relationships were found. However, these relationships are only discovered by viewing and capturing individual fully built out term records.

In most semantically elaborated thesauri, one would expect to find that between 30% and 40% of the relationships defined for an individual term are associative relationships. Analysis of the working copy of the TRT showed that the number of associative relationships varied widely by facet, with some facets having no associative relationships and others having as high as 49% associative relationships. However, the analysis suggests that those terms that have a related term have only one related term. This does not achieve the benchmark of 30% to 40% of all the relationships defined, suggesting that the TRT has a low level of semantic relationing, a feature that is important for effective use in other information contexts.

In constructing the working copy of the TRT, the research team discovered that a second type of associative relationship is defined. It was characterized as Related Term (Hierarchical). This relationship is undefined in thesaurus standards. Through further investigation, the research team determined that this relationship type represents sibling terms to the term referenced. They are the narrower terms of the referenced term's parent. Figure 3-1 is an example of a term with Related Term (Hierarchical) relationships.

Ocean currents
  Broader Term
  Hydrologic phenomena (Jbh)
  Narrower Terms
  Sinking (Oceanography) (Jbhms)
  Related Terms (Hierarchical)
  Aggressive waters (Jbha)
  Degradation (Hydrology) (Jbhb)
  Desiccation (Jbhc)
  Drainage (Jbhd)
  Floods (Jbhf)
  Hydrologic cycle (Jbhh)
  Streamflow (Jbhi)
  River currents (Jbhl)
  Peak discharge (Jbhn)
  Ponding (Jbho)
  Runoff (Jbhp)
  Sediment discharge (Jbhr)
  Seepage (Jbhs)
  Storm surges (Jbht)
  Upwelling (Jbhu)
  Water table (Jbhv)
  Water waves (Jbhw)
  Sea level (Jbhx)
  Phenomena of frozen water (Jbhy)

Figure 3-1.   Example of term with Related Term (Hierarchical) relationships.

Related Term (Hierarchical) relationships are not formally constructed between terms in the TRT, but are automatically generated and displayed for individual terms. The relationships are neither deliberately nor semantically defined for those terms that are siblings to the parent term. For example, "Aggressive waters (Jbha)" is a child term of "Hydrologic phenomena (Jbh)," and thus a sibling term to "Ocean currents (Jbhm)." Because the relationships are not deliberately created, the nature of the relationship between the terms is undefined, presenting design challenges for using the TRT outside of TRB information repositories.

For the term in Figure 3-1, "ocean currents," the research team notes that no true Related Term (Associative) relationships have been created. Semantic richness would be achieved if there were true Related Term (Associative) relationships built for the term. For example, semantic richness would be achieved if the relationships for the term "ocean currents" were expanded to include NTs for "surface currents" and "deep water currents." In addition, semantic richness would be improved if Related Term (Associative) relationships were added from "ocean currents" to terms such as "measuring currents," "ocean current impacts," "weather," "ocean going vessels," "ocean tides," and "ocean waves."

Finally, the team found some possible errors in the construction of relationships in the current TRT. Examples are documented in the Final Interim Report 1 and suggest that as part of any enhancement to the TRT, a thorough review of the relationships should be performed. This should be based on clear guidance and requires generation of a report for management that makes it easier to identify these issues.

Metadata to Manage Terms

Metadata for managing terms include definitions, SNs, and S&O fields. While definitions are included in the TRT, the practice is not consistent. Some of the smaller facets have a higher percentage of definitions. Those facets that have a higher concentration of terms and whose terms are more granular and specific to transportation objects tend to have fewer definitions. SNs are intended to guide indexers in the use of thesaurus terms. They generally explain the ways in which the term might be interpreted and how it should be applied. The occurrence of SNs is very sparse throughout the TRT and consistently sparse across facets. In the detailed task to develop definitions for 50 current TRT terms, it became clear to the research team that the definitions are often used as SNs and, in fact, may be duplicated in the SN field for some terms. This means that the practice of SNs is inconsistently applied as well. No S&O fields are included in the current TRT, which makes it difficult to trace the provenance of a term in the future.

Another key element in managing terms is the assignment of unique identifiers. While the files provided by the TRT management team each included a number that appeared to be a unique identifier for each term, through discussions the research team understands that the values do not serve that purpose. In most thesaurus management applications, a unique identifier is automatically generated by the system and assigned to the term. The unique identifier is a database-level control point that cannot be changed by an end user and is never reused at the database level. The unique identifier should be understood as a database-level control point rather than as a definable or changeable attribute of the thesaurus term. The research team's understanding is that unique identifiers are not assigned to individual terms in the TRT database. Neither are unique URLs or digital identifiers assigned to individual thesaurus terms as digital object identifiers so that they can be globally persistent outside the TRT system.

Treatment of Terms Throughout Their Lifecycle

Lifecycle management of terms is supported by the inclusion of status fields and workflows to support candidate, provisional, and deprecated terms and a history/comment log that details

what has occurred throughout a term's lifecycle. These functions are not supported in the current TRT. This, and the fact that all dates were changed to the same date when the last conversion was performed, made it impossible to determine from the TRT itself the number of additions or changes to the TRT during a given period of time.

Regarding deprecated terms, from the information provided by the TRT management team, the research team was able to observe a few examples of changes of term status. Where a term changed from a preferred to a non-preferred term (e.g., when a term shifted from enumerated to non-enumerated), the non-preferred term was still linked to old TRID records.

The research team learned that deprecated terms had their status change to uncontrolled and were removed from the enumerated hierarchy. Their previous enumeration was assigned to the new preferred term. In interviews with indexers and transportation librarians, the research team learned that uncontrolled terms retain their links to bibliographic records in TRID. Uncontrolled terms may be reported out as a list available to the TRT management team, but they are not explicitly manageable as non-preferred terms.

Overall Thesaurus Structure

A thesaurus that is composed of multiple facets should strive to achieve balanced coverage of the language of the domains it supports and serves. This issue is addressed as scope and coverage in thesaurus standards. In practice, robust scope and coverage results from close working relationships with experts in the domain. In a complex field like transportation, where there are multiple modes with specific vocabularies, interdisciplinary factors, and emerging topics, a thesaurus structure is most effectively managed by an overarching facet or category structure. Within each facet, one would expect there to be a robust second level of subfacets within which are built out specific vocabularies of terms and their relationships. Balanced structures enable thesaurus managers to better assess whether a facet's vocabulary is supporting the domain or area of practice.

On average across all facets of the TRT, the ratio of Top Terms to Level 2 Terms is 1:8. This would be a good ratio if it were implemented consistently across facets. However, the second level of the current version of the TRT does not follow a predictable or consistent pattern. The smallest number of Level 2 divisions is found in Facet R [Materials], and the greatest number is found in Facet M [Persons and Personal Characterizations]. As noted earlier, the variant practices observed are attributable to the strict hierarchical foundation of the TRT in contrast to the extended semantic relationing one would expect to find in a standards-compliant thesaurus. Each Level 2 represents a different browsing structure for end users. This is one of the reasons why stakeholders generally reported that they do not use the hierarchical browse structure of the TRT.

Below Level 2, the research team found additional variance in the depth of the structures across facets (see Figure 3-2).

These variances do not show good thesaurus practice and result in an unbalanced structure across the TRT.

Findings

In general, the descriptive information suggests that current TRT content is strong in some areas, but there are clear weaknesses in others. The weaknesses appear to be an un-intended consequence of the rigid hierarchical enumeration that is used as a backbone structure. The research team understands that there are constraints to changing this structure that are associated with
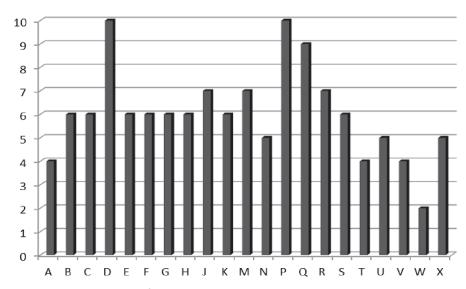
Figure 3-2.   Depth of terms below Level 2.

its design as part of TRID. Where these constraints could be mitigated there is an opportunity to enhance the TRT. There are important opportunities for improvement in the TRT relationships used in the TRT, but again there are constraints imposed by the integrated TRID architecture. Similar opportunities for improvement were found for the metadata for managing terms, the management of terms through their lifecycles, and for the overall structure of the thesaurus.

Access and Use

Dimension B addresses the various means of accessing and using the TRT content. Failing to consider different kinds of uses can lead to suboptimal and constrained design, which limits the products, services, growth, and potential long-term value of the TRT.

The research team identified eight criteria to assess access and use. Access and use includes explicit, direct use by people. The criteria to assess explicit use include using the TRT for TRID indexing or searching, to understand the transportation research domain, and to locate definitions for transportation-related terms. Embedded use is where the TRT is used within systems. The criteria to assess embedded use include external systems using the TRT content and a number of criteria related to the use of the TRT terms for searching and for TRID indexing. The latter are described at various levels and by facet. Customized use, the partial extraction of TRT content to create customized vocabularies, is assessed anecdotally. Finally, general TRT usability was assessed using criteria based on Usability.gov.

Explicit Use

Explicit use is defined as the direct access of the TRT by people—interactive use. It includes direct manual use whether through a web interface or another access point made available to the community. Explicit use can include manual consultation of the TRT and searching the TRT to identify index terms or to select terms for searching in TRB repositories or other commercial information sources. This type of use also includes direct use of the thesaurus to identify definitions or to extract terms or term records from the TRT.

The TRT is often used by librarians at universities, state DOTs, transportation research centers, and private-sector organizations to identify terms to use in indexing. Catalogers and

indexers most often reported searching for transportation terms as a way of validating the terms they had identified for a particular document or publication. Explicit use of the hierarchical browsing structure was noted by a few catalogers, particularly when a pre-identified term that they were searching for was not found in the TRT search results. Catalogers and indexers also reported selecting a closely related term and reviewing the list of terms associated with that term. In some cases, this provided guidance. Most stakeholders interviewed, though, did not report using the Related Term (Hierarchical) terms in the display.

The TRT is rarely used to identify terms for searching in TRB information sources. The strategy most often described for searching in TRB information sources was a trial-error approach—an initial term or query is searched, and depending on the results, the query is refined. This anecdotal information was validated in Google Analytics reports (discussed below). Reference librarians were aware of the TRT, but generally reported that they did not consult the TRT in preparing their search strategies in TRID or other transportation information sources. SMEs were generally unaware of the existence of the TRT and, therefore, did not report explicit use.

While SMEs were generally unaware of the TRT, when alerted to its existence, they noted its potential value and use for other information discovery tasks. One SME, who had worked in both private-sector transportation organizations and an academic environment, noted the TRT's value as a resource for gaining a quick understanding of the state of research in different aspects of the field. Another SME, who has many high-profile review and editorial responsibilities, noted the potential value of the TRT as a visual navigation tool for transportation resources within and beyond TRB. All SMEs noted the potential value of the TRT as an embedded tool supporting current awareness and new resource recommendations.

In general, the research team found that while the TRT is available for explicit use, it is rarely accessed. While the research team did not have sufficient user log data to determine exactly how the TRT was being accessed directly, the interviews with librarians and researchers suggested that they do not use the TRT as a resource for identifying search terms for other information resources. In fact, there was little to no awareness on the part of researchers (as represented by our SMEs) of the existence of the TRT.

Embedded Use

Embedded use is defined as the integration of the TRT into information- or business-related applications, such as a search system, an information or content management system, an automated classification or indexing system, or a workflow or decision support system. Embedded use implies that the consuming application uses and interprets the TRT without any human intervention other than the initial setup of business rules. Therefore, this type of use does not include manual addition of TRT terms to cataloging applications or authority control systems. The use of the TRT in TRB information repositories is an example of embedded use. For some criteria in this section, the research team relied on input from stakeholder interviews. For other criteria, such as the use of the TRT embedded in TRB information sources, Google Analytics reports provided reliable data for 4 months of activity.

The two primary examples of embedded use are (1) the TRT's integration into TRB repositories for indexing and (2) the TRT's integration into TRB repositories for searching. Data and use patterns for these examples are presented and discussed below. Beyond these two primary examples, the stakeholder interviews and data collection efforts surfaced only two other embedded examples. The first was a conversion of the TRT rectangular tables into an Oracle Thesaurus component to support searching in a business application at a transportation research center. The second was a manually converted integration into a metadata management application. By

and large, all other examples described in the stakeholder interviews or data collection efforts fell under the customized use category.

Since the TRT's main purpose is as an embedded tool for both TRID indexing and search, the team performed extensive analysis of the TRT terms for both searching and indexing. The use of the TRT terms alone in search queries is rare; the vast majority of the searches are keyword searches. The ratio of the TRT terms to their use in indexing was overall considered to be good, but it varied widely when calculated at the facet level. With the exception of the transportation mode terms in Facet A, all facets showed an extensive use of the top term. This may suggest underdevelopment or overdevelopment of some facets. Facet R [Materials] has the most terms and is also the most heavily used for indexing. Suggestions were made for rebalancing and improving some of the facets based on the assessments performed and discussed with the project panel as part of the TRT assessment and the detailed Facet X (Information Organization) analysis.

In terms of its use as an embedded tool for other systems, the TRT's tight integration with TRID indexing and searching makes it difficult for the TRT to support embedded use in any other system that is not closely aligned with TRID. The limitations of structure and functionality and the lack of a fully ISO-compliant thesaurus also severely limit the embedded uses. Those uses that were identified were anecdotal and reported significant cost and difficulties in doing customization.

Customized Use

Customized use is defined as a partial extraction of the TRT based on a particular facet or selection of a particular level of terms to create a classification scheme or a customized vocabulary management system focused on a particular mode or emerging topic. Customized use in specialized applications may include local uses requiring customization of the content, uses that are focused on a single transportation mode, or uses that extract terms from across all facets of the TRT to generate a new thesaurus product. In addition, specialized applications may include the extraction of thesaurus terms, with the construction of new relationships to describe conditions in a particular country or region. Customized use is important for supporting future semantic applications and the generation of new products, and it enhances the ability of the transportation community to use the TRT as a basis for new products and services.

For these criteria, the research team relied on the input from stakeholder interviews. Just as with embedded use outside of TRID, limited custom use was found. This is primarily because of the lack of extraction capabilities in the current TMS. Cases where there have been attempts at custom use have been small scale and have required significant manual effort.

Usability

Usability criteria focus on the explicit use of the thesaurus and the functionality that is available to stakeholders. The research team used the criteria available at Usability.gov, an authoritative source used across the industry, for guidance in assessing the TRT's usability.

While the TRT meets the usability criteria for the functionality provided by the interface, there are significant difficulties, particularly in the use of the navigation structure to browse. The navigation structure is dictated by the enumeration codes, and the facets that one would anticipate as the most heavily browsed are also the most difficult to navigate because of the structure. Some functionality, such as the generation of KWIC and KWOC, which are basically oriented toward print indexes, are of questionable value to stakeholders.

Findings

Improving the TRT in aspects related to access and use is difficult because these criteria are so heavily dependent on, and therefore constrained by, the TRT application. While some enhancements to the TRT interface to improve explicit use are possible, they may not be cost-effective given the lack of usage. The majority of contemporary applications that involve embedded or custom uses are impossible or very difficult in the current TRT context.

Governance Processes

Standards in the field of information science are intended to provide guidance but will always be interpreted and adapted to a specific environment. While one cannot look to the standards for specific assessment criteria for governance, the governance process for the TRT should be designed to support the transportation community. This dimension was assessed using the following categories: (1) community engagement in TRT governance, (2) level of focus of governance, (3) use of the TRT system to support decision making, and (4) TRT change management.

The research team found that engagement in TRT governance is dependent on the role played by the individual; involvement is highest among members of the TRT Subcommittee, who are primarily librarians and information professionals who oversee and guide changes to the TRT. The information science community provides an important kind of understanding of the thesaurus and its value to information management and discovery. However, even though some information professionals are also SMEs in certain areas of transportation, there is a general lack of direct engagement on the part of transportation SMEs, and a subsequent lack of transportation expertise in the process. The research team found some references to consultation with transportation experts, but they appeared to occur when there was a need for clarification or guidance in the use of a particular term.

On a practical level, however, the research team understands that the current structure of the TRT constrains the use of expert guidance. Let's consider that an SME in intelligent transportation systems or intelligent cities advises that these are fast-growing areas of focus for transportation, and they warrant full vocabulary development. The TRT Subcommittee's effort would be substantial. The current architecture of the TRT would constrain the TRT Subcommittee from adding a new facet—a challenge not imposed by a standard thesaurus. Instead they would have to find an appropriate placement within the existing structure or they would have to redefine and refocus an existing facet. Depending on how deeply the new "top term" is placed in the hierarchy, there may be limitations to how broadly the new field may be represented. Before it is suggested that the governance process be expanded, there have to be equivalent opportunities for the TRT Subcommittee to easily and practically act upon expanded input. The governance process is focused on the selection, addition, and placement of terms, which may limit the development of related terms around the term's broader concept and, therefore, the growth of the thesaurus. The TRT system provides limited support to the governance process.

While there is a process for recommending terms, the feedback to recommenders is limited and may discourage further suggestions. The system lacks the functionality to track suggested terms and the history of discussion regarding the decisions that are made or to produce reports that would provide the comprehensive views of the content and structure that are needed to make strategic decisions about the content or to perform periodic evaluation/audits of the content.

Overall guidance of transportation experts on the direction of the TRT was noted as missing from the governance process. In addition, SMEs interviewed—with one exception—were unaware of the

TRT. The research team could not find evidence of annual reviews or audits of facets with SMEs. Nonetheless, SMEs expressed an interest in being involved in the governance process. The challenge will be how to structure that engagement so that it has a low impact on SME time and still provides critical information to guide the TRT's development.

This is the area where significant improvements can be made both to the governance process and to the TRT system's support for that process. The research team acknowledges the difficulties of engaging transportation researchers in the governance process, but there are also significant benefits in terms of the content and promotion of the TRT. Another area of opportunity is the addition of reports that would support the governance process and provide more complete feedback to those involved.

Governance Tools

This analysis focused on the TRT's alignment with thesaurus standards, the alignment of the TRT with the scope and coverage of the field of transportation, and the productivity of the TRT terms when used in TRB and non-TRB information sources. The research suggests that the TRT partially supports thesaurus standards. The primary criterion for improvement is the strength and interpretation of literary warrant. The variation may be attributable to different understandings of the structure and use of thesauri versus library subject headings. In addition, variation may be attributable to different views on whether the TRT is an extended classification scheme or a thesaurus. Regardless of the reasons for different understandings of the TRT's warrant, these variations are reflected in the scope and coverage of the field.

TRT Alignment with International and National Thesaurus Standards

National and international thesaurus standards provide guidance on the formation of terms that may lead to more effective search queries and to indexing terms that represent the way that people in a domain think and speak. The readability and understandability of terms outside of their context in the thesaurus is an important guideline. By and large, the TRT content supports this criterion with the exception of the structural terms, the use of parenthetical qualifiers, and stand-alone adjectives and adverbs.

Where a thesaurus has developed and published style guidelines it is important that terms align with those guidelines. The TRT governance process ensures that terms are aligned with stated guidelines. However, while the guidelines refer explicitly or implicitly to some of the style issues highlighted in thesaurus standards, there are a few minor gaps (e.g., qualitative and quantitative nouns). In addition, there was feedback from indexers and transportation librarians on the need for either clarifications or expanded guidance.

Finally, perhaps the most important characteristic of a thesaurus is its warrant—the rationale that is at the foundation of all decisions to include, exclude, or link terms. The two most prominent forms of warrant are literary and user warrant. The TRT guidelines suggest that literary warrant guides the development of the TRT. Through the research team's interviews with the TRT Subcommittee members, literary warrant was confirmed. However, there was some uncertainty as to how literary warrant was being interpreted. In particular, does literary warrant mean use of a candidate term in TRB publications only, or is it use in the general transportation literature? Does it mean heavy use or is emerging use sufficient? The literary warrant was also developed in the 1990s. Transportation literature and the use of transportation language in that literature have changed significantly since the 1990s. Because the transportation literature is extensive, there is a need to clarify what is meant by literary warrant.

TRT Alignment with Scope and Coverage of Transportation Domain

For this assessment, the research team identified 20 resources that represent the transportation domain. They ranged from TRB Annual Meeting program areas to handbooks and peer-reviewed journals. A set of TRT terms was selected and searched against these resources—recording "exact matches," "no matches," and "potential relationships" (see Figure 3-3). Potential relationships are defined as cases where a term identified in a non-TRT resource could be linked to an existing TRT term. For example, "economic development" exists in the TRT, but "urban economic development" is found in a non-TRT resource. It is considered a potential relationship match since it could be added to the TRT by adding a BT/NT relationship with "economic development." There is a place to semantically "hook" or "anchor" the new term in the existing TRT structure. No matches may be viable terms for the TRT, but would require a new structure in order to be added.

Handling of Transportation Modes

Additional analysis focused on how well the TRT deals with particular modes of transportation. Transportation modes are not currently represented by dedicated facets. Instead, the TRT aims to cover all modes of transportation in a modal-agnostic structure. Vocabularies for all modes, specialized topics, and community perspectives have been integrated into this
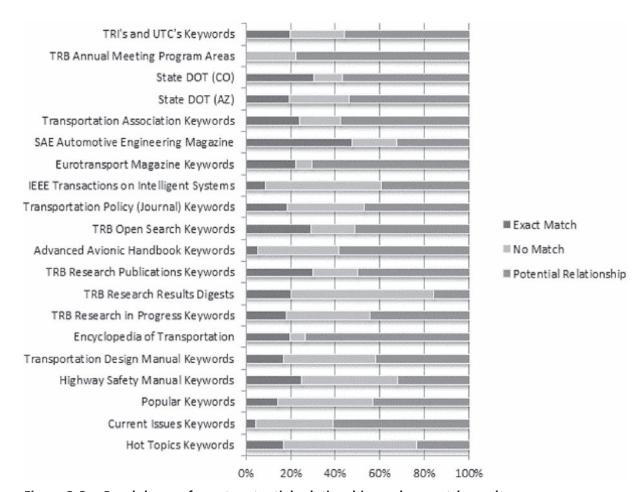
Figure 3-3.   Breakdown of exact, potential relationship, and no match results across sources.

modal-agnostic structure. This was a point of assessment raised by the SMEs supporting the project during the interviews.

By far, the greatest number of terms pertains to highway and road transport, but even that mode is not represented by a dedicated facet for roads and highways or a way to bring the terms together. In fact, this facet is lightly treated as a subdivision of land transport. The other modes—air, rail and water transport—are scattered across the TRT. An illustrative analysis is provided in the following paragraphs and Figures 3-4, 3-5, and 3-6; there may be other modal views that need to be addressed.

Air Transport Terminology Scatter

Terms related to air transport are scattered across the thesaurus without term-level relationships that would allow one to see the full scope and coverage of that mode. This structure makes it difficult to identify and assess the coverage of a specific mode. Figure 3-4 demonstrates the distribution of air transport terms across the current thesaurus structure.

Rail Transport Terminology Scatter

Terms related to rail transport also are scattered across the thesaurus without term-level relationships that would allow one to see the full scope and coverage of that mode. Figure 3-5 demonstrates the distribution of rail transport terms across 11 of the 21 facets of the current thesaurus structure. The greatest concentration of rail terms is in Facet Q, which represents a vehicles and equipment perspective.

Water Transport Terminology Scatter

Terms related to water (marine) transport are also scattered across 13 of the 21 facets (Figure 3-6). About half as many terms related to water transport are included in the TRT as are provided for the other two modes. Terms for water transport are concentrated in Facets H [Safety and Security], P [Facilities], and R [Materials].

TRT Term Alignment with Other Transportation Vocabularies

Finally, the research team explored the degree to which the terminology in the TRT matches that of other vocabularies with significant transportation coverage (see Figure 3-7). 218 terms
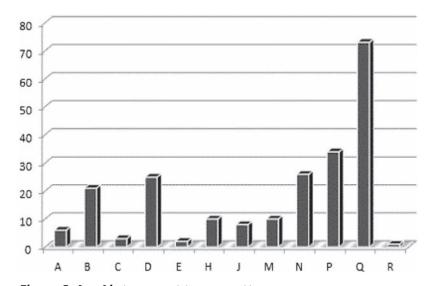
Figure 3-4.   Air transport terms scatter.
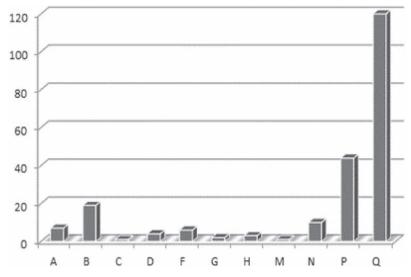
Figure 3-5. Rail transport terms scatter.
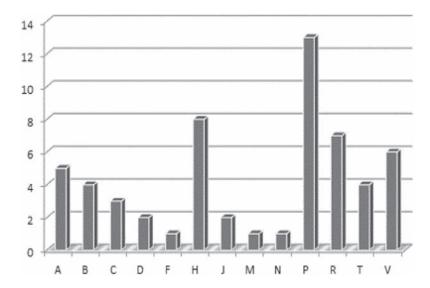
Figure 3-6. Water transport terms scatter.
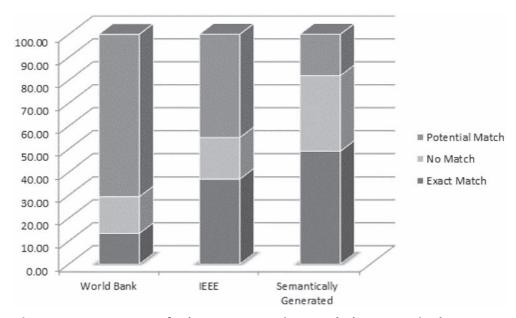
Figure 3-7. Coverage of other transportation vocabulary terms in the TRT.

from the TRT were selected and searched against The World Bank Thesaurus (2010 edition). There were 37% exact matches, 18% no matches, and 45% potential relationships. A similar matching was performed against the IEEE Thesaurus, using 51 selected terms. The results were 13.30% exact matches, 16.06% no matches, and 70.64% potential relationships.

The final source of other controlled vocabularies was derived from an exercise in 2012 to generate terminology from a corpus of transportation documents using semantic technologies. The list, called "Semantic Technologies" for purposes of this assessment, was based on noun phrase extraction and a manual review and trimming. The list was also categorized into areas of transportation. The original source was in the tens of thousands of concepts. From that list, the research team selected 92 keywords for testing. The results were fairly evenly distributed across the three categories, with 23.91% exact matches, 33.69% no matches, and 42.39% potential relationships. The rate of exact matches is higher than for other sources.

The results show that while there is some overlap with these other vocabularies, there are some gaps that can be filled. The use of this technique to identify such gaps could help to support further semantic enrichment of the TRT, utilizing these resources to identify both new terms and additional relationships. As noted earlier, there are opportunities within the no match and potential match results to identify terms that have logical links to the TRT terms and could be integrated into the TRT.

Findings

The TRT has a rich foundation of terms. However, the foundation only partially aligns with the scope and coverage of the field of transportation. This is understandable, since the current TRT application presents constraints and challenges to thesaurus managers, consumers, and stakeholders, and the structure encourages a focus on term-level management rather than on facets or subdomains.

The results of the search tests against other resources are consistent with the gaps surfaced in the Google Analytics search logs discussed earlier in this chapter under Embedded Use. The terms found in the Open Keywords search logs were not unlike those keywords selected from authoritative transportation research products.

The search tests against other domain resources also highlighted the generic nature of many terms in the TRT and gaps in transportation-specific terms. Of particular note was the lack of alignment of the TRT's overall structure with TRB subject areas as represented on the TRB website and in TRB Annual Meeting programs. The research team also notes that a simple review of the 21 facets of the TRT surfaced only two that would be recognized as transportation-focused from their names. Nineteen of the facets are generic and might be understood to be part of any other controlled vocabulary. A simple example would be Transportation Management and Organization. While the transportation context may be implicitly understood when TRT is used in TRID, this context is lost if the TRT is used outside of TRID.

The research team also found gaps in coverage when comparing the TRT terms to other transportation thesauri and controlled vocabularies. There was a low exact match rate of transportation-focused terms drawn from three sources (e.g., World Bank, IEEE Thesaurus, and Semantically Generated Keywords) with the terms in the TRT. While the overall no match rate was high, there is opportunity in the potential matches between terms from other thesauri and terms in the TRT.

This portion of the assessment also addressed the productivity of the TRT terms in TRB and non-TRB information sources that are frequently used by transportation professionals and practitioners. The research team notes that the assessment criteria for the productivity of the TRT terms are partially supported because there are wide variations in results where one would expect more consistent behavior. In addition, the structural terms, parenthetical qualifiers, and other

grammatically non-standard terms are not productive in searching other sources. For the TRT to have value to the full transportation community, it should be applicable to other environments and used in contexts beyond TRB.

The research team addressed the question of pre- and post-coordinate indexing practices. Pre-coordinated terms involve the combination or two or more concepts (i.e., this should not be confused with simple multiple word concepts). An example of a pre-coordinated concept is "economic development in urban areas." In contrast, an example of a multiple word term is "urban economic development." The use of pre-coordinated terms assumes that an indexer must predefine or pre-coordinate the terms that are likely to be searched and assign them as "whole" indexing terms. Pre-coordinate indexing or subject description practices were first used in physical card catalogs where there was a single physical entry point that a user was likely to find. This practice assumes that the search system or the searcher cannot efficiently handle searching for two distinct concepts. This earlier context was also limited by what search system architects refer to as "left edge matching." If the index design did not allow for matching second, third, or fourth words in the index entry, no match would be found. In the traditional context, pre-coordinate indexing could lead to missed matches to search queries. This risk is reduced in today's full-text search index architectures.

With the advent of online catalogs and fielded search systems such as BRS and DIALOG, it was possible to build search system indexes that could "slice and dice" or combine indexing terms. Post-coordination of terms—enabling the search system to find specific index entries for multiple terms and retrieve all relevant entries in a combined result set—was an effective advancement for that point in time and for the context. Using the example from above, the search system might find terms for "economic development" and for "urban areas." The main challenge with a post-coordinated approach to indexing is the large and often irrelevant recall that results. Where post-coordinated terms are generic and not context specific, this can result in low-performance searching. The post-coordinated approach also was dependent upon (1) the searcher specifying how to combine the two different sets of results, or (2) the search system architects setting the correct default matching algorithms for results. Because there is no one correct default for all contexts, search system architects typically set the default to include all results. Depending on the matching algorithms, the most relevant results may not appear in the top results, which are most likely to be viewed by the searcher.

With the increased computing power of the 1990s and the development and continued architecture advancements of full-text search indexes, indexers are no longer constrained to either pre- or post-coordinated indexing strategies. Most full-text search systems and most bibliographic search systems now build their internal index architectures based on what are called "rolling parsed entries," which are the equivalent of KWIC structures. This means that an index entry will be constructed for every word in an indexing term field. It is possible and efficient to try to match a multiple word search query against such a full-text index without the searcher having to specify how to query and match every search term. The question has shifted to "how are the query-processing and query-matching strategies defined?" It is now possible to manually and machine index content to reflect how people talk and think about what they are looking for in a search. Thesauri are now important tools in developing query-processing and query-matching strategies, whereas in the past they were primarily used for guiding manual indexing.

For TRB, the future development of the TRT highlights the underlying question—what type of search systems will the TRT be used with in the future? In the past 20 years, the question has moved beyond pre- and post-coordinated indexing practices. If the TRT is to have a future, it must more closely reflect the vocabulary and language of those involved in transportation, without concern for pre- or post-coordination of terms.

A positive observation from the tests of the TRT is the opportunity to expand the scope and coverage of the TRT by leveraging "literary warrant" more extensively—through the integration

of transportation-focused terms from search logs, commercial sources, and other transportation vocabularies.

The overall results of the assessment of scope and coverage suggest a shift to a proactive literary warrant strategy to support the TRT. There are important gaps in scope and coverage. Such a shift would involve a continuous and proactive review of the literature to ensure that current coverage supports the full domain. This does not have to be a labor-intensive manual process; it can be automated to increase the pipeline of terms flowing into the decision/review process. While a shift to a proactive literary warrant strategy is suggested, it is not clear that this approach could be implemented within the current TRT architecture.

Architecture

Dimension E focuses on the architecture, functionality, and protocols of the application that supports the TRT. The architecture and functionality of the application are important because they define what capabilities are available to work with, what can and cannot be done, and what products and services can be produced. The research team identified four levels of architecture that impact Dimension E. These levels are (1) the architecture of the TMS application, (2) the architecture in which the thesaurus functions, (3) the communication architecture that supports interoperability, and (4) the architecture required to support multilingual versions and uses of the thesaurus. Forty-eight of the original 134 assessment criteria pertained to Dimension E. Of these, 39 pertain to the architecture of the TMS, 5 to the context in which the TRT functions, 3 to interoperability, and 1 to multilingual use.

Sources of data for these assessment criteria were derived from the research team's construction of the working copy of the TRT, comparisons to other thesaurus management tools and to the ISO 25964 conceptual architecture, and from extensive interviews with the TRT management team and the TRT system administrator/designer.

Dimension E is the least amenable to enhancements in the current TRT context. The current architecture only partially meets the assessment criteria. This is not surprising considering that the thesaurus architecture is also constrained by the tight coupling of the TRT application to TRID indexing and searching. The database schema revolves around this primary purpose rather than a more general thesaurus standard-oriented approach that would support greater flexibility in terms of thesaurus management functionality and easier interoperability with other applications. There are many functions that are not currently provided within this context. The opportunity exists to review these functions, especially those that support interoperability, export, and governance reporting to determine how the architecture might be supplemented to address more of the requirements in this dimension.

Because of the importance of Dimension E to the capabilities of the TRT, a separate task was performed to assess the TMS market. The findings for this task are reported in Chapter 5.

Overall Assessment of the TRT

Based on the detailed criteria for each of the five dimensions defined for the assessment of the TRT, the team generated an overall assessment of the TRT against the criteria. Aggregate assessments of criteria in each dimension are presented in Table 3-2. The data are presented as percentages of the total number of criteria in the dimension.

None of the dimensions produced an aggregate "fully supported/fulfilled" result. Dimension A, which focuses on thesaurus content, had the highest percentage of criteria that were fully supported. Dimension C, which focuses on governance processes, had no criteria that were fully

Table 3-2.   Overall assessment of the five dimensions defined for the assessment of the TRT.

Overall Assessment	Dimension A	Dimension B	Dimension C	Dimension D	Dimension E
Fully Supported/Fulfilled	15.55%	7.69%	0.0%	12.50%	8.3%
Partially Supported/Fulfilled	48.88%	57.69%	33.4%	20.83%	16.6%
Not Supported/No Practice	35.55%	34.61%	66.6%	68.75%	75.0%
Total Criteria	*45*	*26*	*12*	*33*	*48*

supported. Dimension A and Dimension B (Access and Use) had the highest percentages of criteria that were partially supported or represented variant practices. In contrast, Dimension C, Dimension D (Governance Tools), and Dimension E (Architecture) had high percentages of criteria that were not supported or had no current practices.

In general, the content of the thesaurus and its current use in TRID appear to be areas of strength. Dimension D (Governance Tools) has areas of strength and weakness. One area of potential enhancement, however, is the scope and coverage of the field of transportation. Dimension E (Architecture) is another clear area for improvement of the current TRT.

The team performed further analysis across the assessment criteria and determined that the majority of the criteria could not be readily improved from partially or not supported to fully supported within the context of the current TRT system and architecture. The cost, time, and risk involved, based on the constraints of the integration with TRID, suggest that this approach would provide limited benefits for the cost.

Future Strategies

In order to look at the TRT's future, the research team identified five potential use cases or scenarios that describe a range of options for the future of the TRT. These five use cases take the reader from the improvements that can be made in the current context (Use Case 0) through a series of progressive improvements that end in a community-based designer thesaurus (Use Case 4). With the exception of Use Case 0, which maintains the status quo, each step in the progression builds upon the previous scenario and lays out clear choices and decision points for the TRB and the broader transportation community (see Figure 4-1).

Use Case 0 focuses on enhancements that can be made to the current master version of the TRT as it is designed within TRID. Use Case 1 shifts the master version of the thesaurus from its current application in TRID to an ISO-compliant application. In this scenario, the TRID TRT becomes a consuming system that draws content from an extract-transform-load or a web services application tied to the new master version. (A web service is a standardized way of communicating between electronic devices or system to system, over the Internet. It allows the integration of data between applications.) Use Case 2 focuses on enhancements to the new master version of the thesaurus to support both local and global information discovery needs throughout the transportation community. Use Case 3 enhances the new master version of the thesaurus by providing services for automated categorization and indexing across the transportation community. Use Case 4 draws from enhancements made in Use Cases 1, 2, and 3 to offer a community-based designer TRT for the transportation community to use to develop new information and knowledge products and services.

Together, the five use cases paint a possible future vision for the TRT. The community may decide that it wishes to target Use Case 1 and go no further. The community may wish to set a 3-year tactical plan for achieving Use Case 2. The community may also decide that it does not wish to move beyond Use Case 0, or it may decide to set a stretch goal for Use Case 4. The research team understands that there are challenges with finding the resources to support and advance the TRT. Given potential uncertainty about future public-sector financial support, the use cases provide options for achieving long-term goals through community support.

Each use case is discussed in detail below. Benefits, challenges, resources, and actions are also outlined. These are addressed in more detail in the TRT strategic plan and the immediate action plan (see Appendixes A and B).

Use Case 0: Maintain and Enhance TRT Status Quo in TRID

Use Case 0 focuses on the enhancements that can be made to the thesaurus in its current application as a component of TRID (see Figure 4-2). The enhancements cannot negatively impact the functionality of search and indexing in TRID. All current embedded applications must be respected.
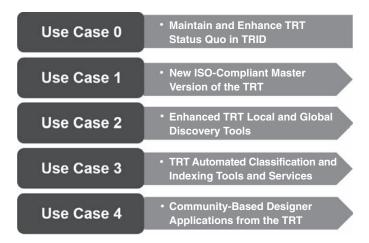
Figure 4-1.   Five use case scenarios.

The aim of Use Case 0 is to expand the management capabilities to support the TRT team and to make adjustments to terms to improve search and usability.

Some enhancements could be made to TRT content without disturbing TRID. However, these enhancements would have to be carefully reviewed. These enhancements include (1) expanding the inclusion of SNs for terms in selected facets; (2) expanding the use of synonyms to include regional variations, lexical variants, abbreviations, and initialisms; (3) reviewing all unapproved terms for consideration as "Use/Use For" (UF/USE) relationships; (4) reducing the number of parenthetical qualifiers in favor of "real speak" terms; (5) rebalancing the structures of those facets that were highlighted in the assessment and discussed with the panel as part of the TRT assessment; (6) increasing the number of transportation-specific terms in facets that address lifecycle stages; (7) and potentially expanding the definition, scope, and coverage of Facet X from information organization to transportation information and technology.

The research team suggests developing an extract of the thesaurus content that is in alignment with ISO standards. This means translating and fully representing the thesaurus relationships to support use in other standards-based thesaurus management applications. This
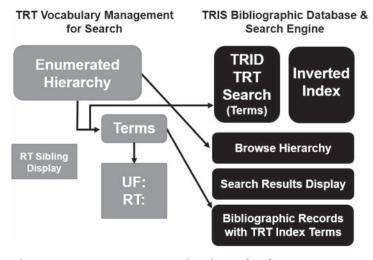
Figure 4-2.   Use Case 0: Maintain and Enhance TRT Status Quo in TRID.

enhancement would help to alleviate the manual labor that local organizations are using to work with all or parts of the TRT. A translation map from the current TRT database scheme to an ISO-compliant scheme must be developed to support this enhancement.

A thesaurus term suggestion tracking tool is also suggested for Use Case 0 and all other use cases to follow. Such a workflow would assist the TRT Subcommittee in managing change requests provided by any stakeholder. The current turnaround time for considering terms is too long, and there is limited follow-up communication with the recommender. This application would allow thesaurus managers to specify the input and explanations needed to better understand why a term is being suggested; to track the research inside and outside the thesaurus at the term level; and to track, record, and publish decisions. The workflow would also support expanded communication with the broader community and provide support for an improved workflow.

Another task that could be addressed in Use Case 0 is the development of an online or an on-site thesaurus training workshop for any interested stakeholders. The interviews and data collection tasks highlighted the need for a common understanding of current thesaurus standards, functionality, and applications.

Use Case 0 might also include a greater effort to promote awareness of the TRT. However, given the current limitations and constraints for enhancement, the awareness-raising effort might be better aligned with Use Case 1.

Benefits and Challenges

Use Case 0 offers important enhancements. It provides some enhanced capabilities to the thesaurus management team, but not to the extent that would be achieved with a standard, commercial-off-the-shelf thesaurus management application. Use Case 0 addresses some of the more serious constraints, such as a lack of synonyms, and reduces the use of artificial structural terms. Development of a mapping from the current database scheme to an ISO-compliant structure can improve interoperability with other thesaurus applications. Use Case 0 maintains the integrity of the existing structure in TRID.

Use Case 0 is perhaps the only use case where the challenges outweigh the benefits. The challenges outweigh the benefits because the investment needed to make the enhancements may not have a future payback. Use Case 0 does not build additional value into the thesaurus, and it presents a hard stop for future development. Given the low levels of use found in end user search and the extensive effort devoted to manual indexing to support those low levels of search (as reported in Chapter 3), the research team suggests that TRB seriously consider whether Use Case 0 is worth the investment.

Resources and Tasks

The suggested enhancements could be accomplished as part of Phase 2 of NCHRP Project 20-109. Unlike the other use cases, however, Use Case 0 would require dedicated time from the TRT system administrator/designer because the work would occur within the current application, and it would be necessary to guard against impacts to TRID.

Timelines

Use Case 0 could be accomplished within the timeline for Phase 2 of NCHRP Project 20-109. However, no further enhancement for the thesaurus would be expected.

Use Case 1: New ISO-Compliant Master Version of the TRT

In Use Case 1, the TRT would be transformed into an ISO-compliant thesaurus (see Figure 4-3). The new master version of the thesaurus would reside in a new thesaurus management application that adheres to the data model and relationships and provides a full set of thesaurus management capabilities for the TRT management team. In this use case, the TRID version of the TRT would not change except to incorporate any enhancements to the content that are manageable by TRID. A new service point would be added to support TRID's consumption of the TRT content—an extract-transform-load or a web services application. However, the utility of the web services would extend to Use Cases 2, 3, and 4.

Currently, the TRT is an embedded component of TRID. TRID searches involve an internal query of the search index, which is tightly integrated with the hierarchical structure of the TRT. This is only one type of TRID search, but it would be important to avoid any loss of functionality in moving to a standards-based thesaurus management application. If the TRT were managed by a formal and full-function TMS, there would need to be a connection between the new TMS and the component of TRID that supports TRT term searches. One way to accomplish this would be to periodically extract, transfer, and load terms from the TMS into the TRID component. The preferred approach, however, is to enable TRID to query the TMS interactively and on demand.

A new master version of the TRT in an ISO-compliant TMS would enable TRB to semantically enrich the thesaurus to improve scope and coverage of the field of transportation. Use Case 1 would also involve a review and semantic linking of all currently uncontrolled terms in the TRT. Use Case 1 might also include a comprehensive review of key transportation sources selected by the TRT Subcommittee and key stakeholders. The review would expand the work begun by the research team (described in Chapter 3 of this report). The goal of this review would be to build out a proof-of-concept semantically enhanced vocabulary. The review might focus on two new facets, perhaps focused on modes, as a proof of concept.

In addition to the semantic expansion of terms, Use Case 1 would enable the TRT management team to shift from the rigid hierarchical structure to a category approach to organizing terms.
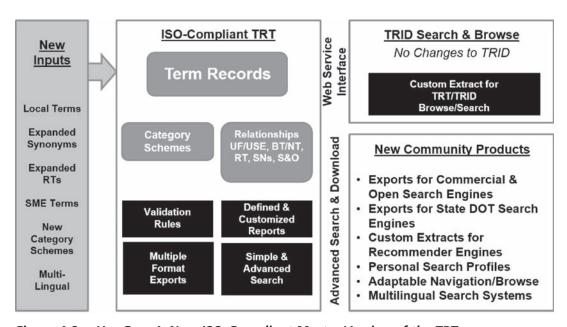
Figure 4-3.   Use Case 1: New ISO-Compliant Master Version of the TRT.

The hierarchical enumeration scheme and codes would be retained as controlled fields in the new master version to ensure continued support for TRID browse and navigation structures. The current facets would be redefined using a numerical encoding scheme that allows for expansion. The expanded structure would also allow for the TRT to define new facets that are focused on modes of transportation. It also would enable the TRT team and the transportation community to consider rebalancing some of the second-level classes described in Chapter 3 of this report.

The new ISO-compliant and more flexible approach to developing and defining relationships would allow the TRT management team to address (1) parenthetical qualifiers, (2) structural terms, and (3) the expanded use and definition of equivalence relationships (e.g., abbreviations, initialisms, true synonyms, lexical variations, etc.). Additionally, upon batch load of the transformed thesaurus into a new application, unique identifiers could be assigned to manage all terms (controlled and uncontrolled). In a commercial application, it would be possible to define multiple definitions and SNs, where those may be of value.

Use Case 1 would also provide an opportunity to define a set of standard reports to communicate the status of facets of the thesaurus and to support routine management of the TRT and changes to the current content. Most thesaurus management applications provide a rich set of standard reporting options, as well as options for customized reports.

A semantic expansion would have impacts on the governance processes and provide opportunities to involve others in the suggestion and review of terms. The workflow application described for Use Case 0 would be valuable in Use Case 1. The workflow application and the expanded management capabilities provided by a commercial TMS would support more extensive communication with the transportation community. The workflow would enable stakeholders to provide more information about their term suggestions, as well as enable the TRT team to track research related to, and the status of, recommendations and to more easily publish decisions about recommendations. In addition, a new TMS would enable the TRT team to include not-approved terms in the database for future tracking and reference.

Benefits and Challenges

The benefits of Use Case 1 are significant. In effect, it provides a more effective management environment for the thesaurus team and a foundation from which to develop new uses of the thesaurus. The challenges are not insubstantial, however. The primary challenge is a shift in mental model from a TRID search-based thesaurus and application to a more fully functional ISO-compliant one. However, Use Case 1 is a foundational step; if TRB does not move to Use Case 1, no further development or non-traditional use of the thesaurus is possible. The other benefit of Use Case 1 is that it gives TRB the option to continue to support TRID in its current form and to explore other options for TRID where there are built-in thesaurus-consumption modules.

Resources and Tasks

Sample ISO-compliant vocabularies may be available to serve as a basis for Use Case 1. Phase 2 might then focus on the additional enhancements to the content (Dimension A) and to the scope and coverage (Dimension D). Use Case 1 requires an ISO-compliant TMS. By definition, a new TMS will address standard criteria described in Dimension E. The development of a workflow application could also be addressed as part of Phase 2.

Timelines

It would be possible to accomplish Use Case 1 enhancements and tasks within the Phase 2 timelines of this project (NCHRP Project 20-109).

Use Case 2: Enhanced TRT Local and Global Discovery Tools

Use Case 2 presumes that the developments highlighted in Use Case 1 have been completed. With a new ISO-compliant master version of the thesaurus, it would be possible to generate standard and customized exports of the content to support applications throughout the transportation community. Use Case 2 would support local use of the master version for search and information discovery (see Figure 4-4). In the assessment of the TRT, the research team saw that there were opportunities for state DOTs, research institutes, and libraries to leverage the thesaurus if it could be adapted to their needs. The ability to generate custom-defined reports and to export these reports in a variety of formats, without the cost of development resources and time, would expand the use of the TRT significantly. Local terms would continue to be managed at the local level but could also be integrated into the master version. Local terms could be tracked in both the workflow application and in the use of term workflows within the TMS. Another option that arises from Use Case 2 is the potential use of the TRT by commercial publishers. In this use case, the TRT could be used as a front-end and navigation tool for web search engines. It could also be used as an integrated component for search engines to support automated query expansion and as a tool to support the clustering and filtering of search results.

In addition, with the capability to define the source and ownership of terms as a traceable field in a new application and to define multiple status conditions for terms, it would be possible to support local terms. The ability to assign unique identifiers to terms and to associate unique digital object identifiers with terms would make it possible for organizations to leverage TRT terms to support local linked data projects.

With an expanded community use of the thesaurus, it is possible that there would be an expansion of involvement of the community in governance (Dimension B). Governance would continue to focus on terms, but the expanded reporting capabilities would enable governance to easily focus on other aspects of the TRT as well. A commercial thesaurus management application would support expanded permissions and securities. The enhancement of content and the increased use by local organizations in local business applications data repositories and for data curation could increase the need for more comprehensive style and best practice guidelines.
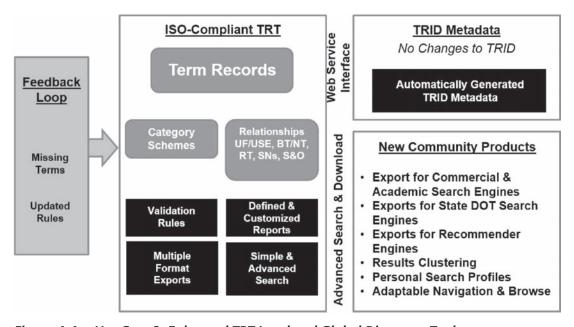
Figure 4-4.   Use Case 2: Enhanced TRT Local and Global Discovery Tools.

Benefits and Challenges

The benefits of Use Case 2 extend beyond the immediate TRT management team to any organization, agency, or publisher who manages and supports access to transportation resources. Use Case 2 also presents a major opportunity to increase awareness of the importance of information management and knowledge organization systems across the field of transportation. This use case provides the broader transportation community with a foundation for exploring common and adapted solutions to information discovery.

Resources and Tasks

Use Case 2 requires engagement of the transportation community. The time required to build engagement should not be underestimated. The major factor in the timeline to achieve Use Case 2 is the information management community's success in engaging the larger transportation community (Dimension B). Most of the work in Use Case 2 involves supporting local applications and learning from those applications. As this engagement grows, the scope and coverage (Dimension D) would also increase. To succeed in Use Case 2, SMEs must be drawn into the governance process.

Timelines

The timeline for Use Case 2 is best aligned with the completion of NCHRP Project 20-109 and Use Case 1 because completion of the activities involved in these endeavors will provide a solid base upon which to support local and extended uses.

Use Case 3: TRT Automated Classification and Indexing Tools and Services

Use Case 1 sets a new foundation for thesaurus services, and Use Case 2 promotes the value and use of thesaurus tools across the transportation community. Use Case 3 addresses services that support the large-scale indexing and categorization of transportation information (see Figure 4-5). Achievement of Use Cases 1 and 2 would create a higher level of thesaurus use. That new demand would increase the pressure on manual processes to manage and organize information resources. Experience demonstrates that manual approaches would probably not be able to meet the new demand. Automated solutions—whether fully automated or human-assisted—would help to meet the demand.

Use Case 3 considers how the TRT—as it would be transformed in Use Case 1—could be used as a feeder system for automated categorization and automated indexing applications (Dimension B). While there are significant questions about what services would be provided and how, Use Case 3, which represents a longer term vision, would provide support for state DOTs, which are under pressure to manage and provide access to information assets without the funding needed. Use Case 3 creates an opportunity for TRB to consider providing centralized indexing services to the states and other transportation organizations, including the private sector and non-profit organizations. Use Case 3 would place either the National Transportation Library or TRB as a centralized service point for metadata generation—functioning similarly to a centralized bibliographic utility such as the Online Computer Library Center.

Benefits and Challenges

The value of this scenario is a more rigorously and effectively managed transportation information landscape. The benefits accrue not only at the national level, but at the local organization
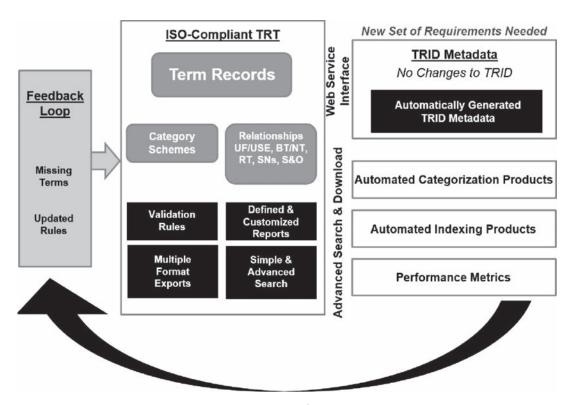
Figure 4-5.   Use Case 3: TRT Automated Classification and Indexing Tools and Services.

level. In most cases, the development costs and resources associated with developing semantic applications are significant. Only the most well-funded organizations can consider these options. Leveraging the TRT as a feeder application would significantly reduce the costs. The challenge would be making sure that there is sufficiently deep understanding about these technologies to effectively develop the semantic profiles and to support the services. A centralized service would require a new set of competencies and a dedicated staff. This would be an opportunity to reposition manual indexers to support such a service, once additional training is provided.

Resources and Tasks

If Use Case 3 is an option for further consideration, we will work with the TRB to develop a detailed cost model. Much of the work accomplished in Use Case 1 will be essential to developing the profiles. The expanded scope and coverage built into Use Case 2 further improves the profiles. Both Use Cases 1 and 2 are predecessor tasks, though Use Case 2 may be implemented concomitant with Use Case 3.

Timelines

Depending on the level of interest of local organizations and state DOTs in Use Case 3, it might be launched at the conclusion of NCHRP Project 20-109.

Use Case 4: Community-Based Designer Applications from the TRT

Use Cases 1 through 3 describe how the TRT could be enhanced to support the more traditional information management and discovery tasks. In Use Case 4, the TRT would be the source of many new designer- and custom-built applications (see Figure 4-6). The TRT itself
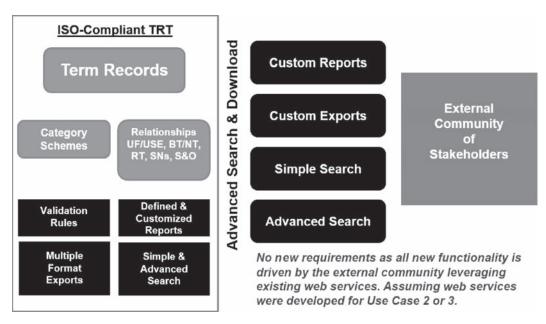
Figure 4-6.   Use Case 4: Community-Based Designer Applications from the TRT.

would be fully enabled as a web service tool, and the full functionality of thesaurus manage-ment that is available to the TRT management team would now be available to stakeholders and consumers.

Some possible new products and services that would be generated through Use Case 4 might include:

- Custom transportation dictionaries that are constructed from thesaurus definitions and modeled around the EPA's MyGlossaries product;
- Custom-developed "Term Lookup" and "Term Mouse Overs," which enable a reader to iden-tify key terms in a document and automatically look those terms up in the TRT or a custom dictionary constructed from the TRT;
- Use by web page developers to select terms, select a definition, and provide a hidden link to a page to improve the understandability of the pages for the non-expert public;
- A "Did You Mean?" disambiguation service that supports lookup of acronyms and initialisms and enables the searcher or reader to select the correct, fully spelled out term;
- Visual thesaurus applications to navigate all resources available from a single source (e.g., TRB) or embedded in a web search engine;
- The use of terms to construct user profiles for recommender engines;
- Use in business rule repositories, workflow applications, and business vocabulary applications;
- Extended uses in smart cities and enterprise information architecture applications; and
- Extended uses in intelligent transportation system applications.

Benefits and Challenges

The benefits of Use Case 4 are difficult to estimate. Essentially, the TRT would become a tool for new and innovative services and products. The challenges are equally difficult to discern because Use Case 4 is dependent upon engaging with the development and technology commu-nity and making a case for the value of the TRT in these new applications. As new applications are developed, community engagement would expand significantly. Each step forward in a use case has implications for TRT governance and management.

Resources and Tasks

In Use Case 4, the main task for the thesaurus team would be to deploy the thesaurus management capabilities as an open data source. This use case would not only provide the content of the thesaurus, but the functionality for stakeholders and end users to do development from start to finish. Beyond this task, the burden of resources and support tasks would reside with the consuming community. The TRT community would provide the resources needed for further development by others. It might be the case, however, that new applications will introduce new TRT requirements for consideration.

Timelines

The timeline for Use Case 4 is several years down the road. Use Cases 2 and 3 are not absolute requirements for Use Case 4, since custom development can begin following completion of Use Case 2 (but only with support from the thesaurus team). Nonetheless, Use Case 4 would be more effectively and efficiently achieved when Use Cases 2 and 3 have been completed.

Assessment of Thesaurus Management Software

As an extension of the assessment of the TRT, NCHRP Project 20-109 included an assessment of TMS tools. During the assessment of future opportunities for the TRT, it became apparent that the known requirements are focused on maintaining the status quo (Use Case 0) and enabling the ISO-compliant TRT (Use Case 1), which is foundational for all other use cases. In addition, Use Case 1 addresses the need for improved dissemination of the TRT to the screen, in print, and as exportable files to encourage and enhance use of the TRT. Therefore, Use Case 1 was used as the practical benchmark for the tool assessment.

Research Approach

The research team identified commercial and open source TMS products based on the expertise of team members and a search of the literature. The list of 52 products was reduced to a working list of 11 based on whether the software had recently been updated, whether its web site looked to be active, and whether it met the basic criteria of ISO compliance. Enhancement to the current TMS was also included as an option.

Following a review of all the assessment criteria, it was determined that all criteria required for the business requirements for the TMS could be taken from Dimension E (Thesaurus Architecture). This resulted in 48 business requirements (see Appendix C). The team then used its knowledge of ISO compliance, the information from the TRT assessment, and interviews with TRT management and staff to assign tentative mandatory/optional weights to each business requirement. In the actual evaluation, the mandatory and optional designations were replaced with a 2 or a 1, respectively, in the "Importance" column in order to calculate a weighted score for each requirement. Those products that met the requirement were given a score of 2. Those that could not meet the requirement were given a score of 0. In this assessment, the research team did not score those that partially met a requirement, but this could be done in the future. The research team was unable to evaluate all requirements using the available sources. In those cases, no score was given for the requirement for that product.

The 12 TMS options were then evaluated against the business requirements. Information was gathered based on vendor web sites, documentation, literature reviews, the expertise of team members, and vendor contacts. The assessment of the current TMS was based on interviews with the development contractor and TRT management and on access to the system. Scores were then calculated based on the weight (mandatory or optional) of the requirements multiplied by the score given to each product.

In addition to evaluating the 12 options based on the business requirements, the team gathered information concerning non-functional factors such as the estimated cost (both initial and recurring), learning curve, staffing requirements, hardware/operating system environment,

and licensing options and terms. Whenever possible, the team reviewed demonstrations or documentation of the software to gather sample screenshots to show the diversity of the interface look and feel and navigation across the products.

It should be noted that this was a market assessment and was not data collection for a procurement. Once a strategic direction is set and procurement activities begin, it will be possible to eliminate much of the uncertainty in this market assessment. The research team suggests that all requirements, as well as additional questions related to the non-functional factors, be provided to the vendors to complete as part of the procurement process.

Findings

Because an ISO-compliant thesaurus is foundational to TRT enhancement and to implementation of other future TRT strategies, as represented by the use cases, the products on the working list were filtered to include only those products that are ISO compliant with respect to the relationships that are supported; only the current TMS is not fully ISO compliant. ISO compliance is an overarching requirement for enhancing the TRT since it sets the stage not only for rectifying some of the current gaps in the TRT's capabilities but also is a requirement for any enhancements and advanced use cases, including Use Case 1 and beyond. ISO compliance brings with it a number of functional requirements around the standard relationships that are met by most products. The current system can perform some of these functions, although it is not fully ISO compliant.

Despite the lack of complete information about many of the products, all open source and commercial products rated significantly higher than the current TMS, even if enhancements were made to the current TMS where possible. All commercial products evaluated are ISO compliant to the extent that they support the standard ISO 25964 relationships; only the current TMS is not fully ISO compliant.

Gaps in the current TMS include not only the lack of standard relationing, but also the lack of a variety of reporting capabilities and a general reliance on the system developer to produce these reports. All open source or commercial products have some level of reporting capability—either to the screen, in print, or both—with the function in the hands of the thesaurus manager or other non-developer roles.

Another gap involves the lack of embedded workflow/governance processes. These exist in the current TRT environment, but they are external to the TMS. Workflow/governance processes are closely related to a number of the requirements. For example, products have statuses that can be used to control workflow, support additional governance and review-related note fields, permit various levels of access, and generate reports that may be used in the workflow and governance processes. Open source or commercial products have more or less formal workflow capabilities. If the TRT community desires to change, formalize, or more highly structure the current governance process, this may be an area of differentiation among the products.

Beyond the compliance requirements, the differentiators among the commercial products were found to be primarily the non-functional requirements such as cost (including recurring costs), staffing, training, and licensing.

The initial costs for the products on the working list vary widely—from freeware to $295 for a single user/desktop version to $129,000 for a perpetual license with 50 users. All on-premise versions of software have options for perpetual licensing. This is distinguished from recurrent licenses, which come at a lower initial cost but with higher annual payments. The perpetual licenses usually include recurring maintenance, update, or support costs that are paid annually

for the following year. The distinctions that drive differences in cost are desktop versus network versions and the number of thesaurus manager "seats" or simultaneous accesses. Other factors may contribute to the final cost including the particular server environment and any customization or technical support required for installation, which is not included in the license. For example, migrating the current TRT content might be a cost factor if the current ISO-compliant master version created for the TRT assessment is not usable. Options for cloud-based TMS as a service are becoming available. While this option eliminates some installation and maintenance costs, hosting concerns, and system administration, it must be reviewed carefully for budget impacts, alignment with desired growth strategies against the use cases, and licensing constraints.

The relatively inexpensive, single-user, desktop products may be viable alternatives, particularly for Use Case 1, where there is limited integration with other systems or users. This initial small-scale implementation, or "convenience copy" would allow staff to learn about the structure and function of an ISO-compliant TMS and to perform some of the initial clean-up of the TRT content. It would be feasible to start with the "convenience copy" and grow into the higher-end products when needed. This decision requires an assessment of the degree to which the TMS would be integrated with external governance processes, such as reviews by the TRT Subcommittee or SMEs. The more distributed the involvement, the more likely the TRT is to need a network/web-based TMS or at least some way to publish updates or extracts of the TRT to the web for review.

High-end products have many features that can take a significant amount of time to learn. However, these products can also be used in very low-level approaches to perform the majority of the tasks that are needed while providing the ability to grow into more complex TRT functionality in the future. The level of staff expertise and training would depend on what functionality is to be exploited and how the TRT is to be used.

While two products may receive the same score for a particular requirement, there can be differences in the way the products implement the requirement. The most desirable implementation (especially look and feel) may depend on the characteristics of the TRT content (as outlined in Dimension A, Thesaurus Content), the roadmap for TRT content enhancement, and workflow and governance approaches. Therefore, demonstrations by the vendors, preferably including sample TRT data, are suggested as part of any procurement strategy.

Many of the decisions regarding the TMS will be subjective, having to do with the look and feel of various TMS products and their usability. Therefore, the research team encourages those who will be using the products under consideration to participate in demonstrations, take advantage of trial versions, and attend the webinars often hosted by the vendors. Purchase of one or more copies of an inexpensive desktop, single-user product as a "convenience copy" is suggested as a way to learn more about the ISO-compliant data model and the functionality of a compliant tool and to fine-tune the requirements.

Conclusions

As a result of the TMS assessment against criteria and comparison of additional context from the use cases, the research team offers the following general conclusions regarding the selection of a TMS.

Final decisions about the appropriate TMS are closely linked to the strategic decisions regarding the future of the TRT as represented by the use cases. If Use Case 1 is the desired approach, the next step is to review the optional requirements to determine which ones the TRT Subcommittee

and the TRT staff consider to be mandatory. Even if the initial focus is on Use Case 1, the TMS selection can be made with an eye to the more advanced use case scenarios, since Use Case 2 requirements are met by the majority of the TMS products on the working list.

Non-functional factors such as initial and recurring costs, technical platform, staffing, licensing, usability, and training will be major factors in the final decision process and are closely tied to the strategic direction.

If the decision is to move more in the strategic direction represented by Use Cases 2 through 4, it will be important to identify additional requirements to support the required functionality. Each forward-looking use case brings additional requirements or modifications to current requirements as well as new opportunities for the TRT.

Strategic and Immediate Action Plan Development

As part of the initial stakeholder interviews and other discussions regarding the TRT, the research team sought to identify goals and objectives. The following were specifically documented during the stakeholder interviews and the in-person meeting with the project panel and have been considered in development of the draft strategic and immediate action plans provided as Appendixes A and B, respectively, of this report:

- Produce derivative products, some of which may be revenue generators.
- Support state DOTs. Some are developing their own thesauri and the TRT should minimally be able to be used as a starting point.
- Support a larger pipeline of candidate terms.
- Proactively incorporate terms to ensure literary warrant is aligned with current research practice.
- Support a broader information community including operational terms.
- Incorporate liberal use of equivalent terms, particularly acronyms.
- Support multilingual capabilities, including the integration of non-English terms from the ITRD Thesaurus.
- Integrate information across TRB systems (the current 32 subject areas are inadequate).
- Provide a management capability for the TRT while not putting the thesaurus functionality of TRID at risk.
- Provide more management and governance support through out-of-the box delivery of searching, reporting, and exporting capabilities.
- Improve the capability of the thesaurus management team to use the TMS product to trace the lifecycles of proposed, candidate, provisional, approved, and not-approved terms.
- Improve the control and management of non-preferred terms within the TMS.
- Use the TRT for discovery by non-information professionals.
- Serve more customers and conduct outreach or the TRT's usage will not be significant enough to warrant even the current investment.
- Add relationships and include more associative and equivalence terms to improve the semantic richness.
- Improve usability, including a better visual experience for SMEs.
- Provide the ability to incorporate other transportation information beyond TRID, which might include state data to allow for better state comparisons and performance, enabling TRID and other resources to be used together.
- Create collaborative partnerships.
- Increase transportation community engagement by providing a domain-level governance model.
- Better support the maintenance process, making it easier to add and change terms to keep the TRT up-to-date.

- Support the TRT's reuse by other systems.
- Integrate with mobile technologies.
- Build metrics into the system to support metrics-driven decisions.
- Include all use cases in the vision.

The following constraints were identified for consideration when determining strategies and activities for achieving these goals and objectives:

- The limitations of long-term staff budget constraints make workflow and governance important.
- Increased involvement by the transportation information, data, knowledge, and research management communities could provide a mechanism for increased efficiency and effectiveness in managing the TRT.
- Increased scope and coverage for the TRT means more terms that require improved full lifecycle management, from proposed through final approval/non-approval, and increased control over non-preferred terms.
- Consider the more detailed equivalent or USE/UF relationships (e.g., spelling variants, common misspelling, generic nouns, trade names, dialectical variants, antonyms, etc.) that might be helpful in the future. (Minimally, the list from the ISO standard could serve as a checklist for the inclusion or enhancement of related terms.)
- Consider who the TRT's power users are (or will be). Who will want to export or access the TRT to create designer applications? It is important to learn about their interests and requirements.

Conclusions

In Chapters 1 through 6, the NCHRP Project 20-109 research team laid out the project plan and research approach, provided historical background to place the current TRT in context, and summarized the analysis and findings for each dimension of the assessment. The team provided five use cases as context for understanding the impacts of the assessment findings for the strategic future and longer term development of the TRT.

In this chapter, the research team offers conclusions founded on the detailed data collection and criteria level assessments, the findings from that data, and the team's understanding of possible futures for the TRT.

Conclusion 1

The research team proposes that the TRT's future value as an information management tool with application beyond indexing of TRID depends fundamentally on achieving Use Case 1. Unless this transformation is accomplished, there is no future scenario for the thesaurus beyond its current state. The TRT may not be able to compete with other knowledge organization systems that may become available on the open source or commercial markets. Transformation to an ISO-compliant thesaurus and thesaurus management application is achievable in Phase 2 of NCHRP Project 20-109.

Conclusion 2

As a result of the TMS assessment against criteria and comparison of additional context from the use cases, the research team offers these general conclusions regarding the selection of a TMS:

- Final decisions about the appropriate TMS are closely linked to the strategic decisions regarding the future of the TRT, as represented by the use cases.
- If Use Case 1, the ISO-compliant TRT, is the desired approach, the next step is to review the optional requirements to determine which ones the project panel, the TRT Subcommittee, and TRT staff consider to be mandatory.
- Even if the initial focus is on Use Case 1, the TMS selection can be made with an eye to the more advanced use case scenarios since Use Case 1 requirements are met by the majority of the TMS products on the working list.
- Non-functional factors such as initial and recurring costs, technical platform, staffing, licensing, usability, and training will be major factors in the final decision process and are closely tied to the roadmap that will be developed following the in-person meeting with the project panel.

51

- If the decision is to move more in the strategic direction represented by Use Cases 2 through 4, it will be important to identify additional requirements to support the required functionality. Each forward-looking use case brings additional requirements or modifications to current requirements as well as new opportunities for the TRT.

Conclusion 3

If an investment in Use Case 0 (the status quo) is the only possible option, the research team suggests that TRB and the transportation community limit the investment to tasks that will contribute to interoperability and correction of non-standard practices.

Conclusion 4

Engage the transportation community, in particular state DOTs, research institutes, university technology centers, and transportation libraries in drafting a long-term strategy based on a consensus across the field. This can be included as a Phase 2 task for this project (NCHRP Project 20-109).

TRT Strategic Plan

The Transportation Research Thesaurus (TRT) is a tool developed and maintained by the Transportation Research Board (TRB) to support access to and discovery of information produced by the members of the TRB. The potential value and role of the Thesaurus have grown over the last decade as the products and services of TRB have expanded and as information management methods and tools have developed. In 2016, the National Academies of Sciences, Engineering, and Medicine and the state departments of transportation (DOTs) sponsored a project to assess the current state of the TRT and to define its strategic future development. The objective of this project was to enhance the TRT by determining the vision and requirements for the future of the TRT. It includes an assessment of the ability of the current thesaurus content, procedures, and products to meet those requirements and the establishment of a strategic plan and short-term actions to achieve the requirements.

Research Strategy

The objective of this project was to enhance the TRT by defining future strategies that will help TRB to achieve its vision and mission. To accomplish this, the project team conducted an assessment of the current thesaurus. This strategic plan lays out a set of strategies and action items that will ensure the TRT continues to be a vital element of TRB's mission and vision that meet the expectations of TRB members and the transportation community. The assessment is presented as a high-level situational analysis and is presented in the context of TRB's mission, vision, and future strategies.

The TRT was initially developed under NCHRP Project 20-32 (2001) to provide a tool to improve the indexing and retrieval of transportation information. This project produced an updated structure and content, as well as a user's guide to support the implementation and use of the TRT. The initial project leveraged the Viewer software developed by CDB Enterprises in 1994/95. While the original design was intended to be ANSI/NISO Z39.19 compliant, the use of the Viewer software and its limited application as a component of search resulted in a suboptimal thesaurus application.

In 2007, NCHRP Project 20-70 was undertaken to develop a database at TRB that would allow TRB to maintain the authoritative version of the TRT in-house (e.g., to eliminate the external contractor role), to integrate the TRT with the then-new Transportation Research Information Services (TRIS) system for creating catalog records, and to publish the TRT on the Internet. Seven specific tasks were identified and completed. This project resulted in the current version of the TRT (http://trb.trt.org). When this project was launched, the architectural design was guided by the TRIS database. Alignment with ANSI/NISO Z39.19 and/or ISO 25964 standards was not undertaken. These standards were not provided to the contractor as design specifications.

53

In 2008, NCHRP Project 20-79 was launched. This project focused on the addition of AASHTO definitions to the TRT and was completed in 2009. This project was implemented in the digital version of the TRT, a component of TRIS. While the project resulted in the integration of some existing definitions, a comprehensive strategy for the use and inclusion of definitions was not developed.

Since 2009, two additional NCHRP research projects have addressed or mentioned the TRT, including NCHRP Project 20-90 and NCHRP Project 20-97. NCHRP 20-90 was launched in 2011 and completed in 2013 by Cambridge Systematics. The objective of the research was to prepare guidance describing the practices that state DOTs could use for capture, preservation, search, retrieval, and governance of transportation data and information, and strategies and actions a DOT could follow to implement such practices. As part of this broader project, there was a small targeted review of the TRT as a cataloging tool. While some suggestions were offered to enhance the TRT, we note that the goal was not to provide a comprehensive review of the thesaurus. NCHRP Project 20-97 referenced the TRT as a tool that was designed to support findability of transportation information. The goal of this project, though, was to improve the findability of state DOT information by defining an information management framework, describing good practices for organizing and classifying information, and developing federated search procedures. The TRT was mentioned but was not treated or leveraged in depth for this research project. Since 2011, the state DOTs have also provided feedback and comments on the TRT, in particular the relevance of TRT's content and usability and ease of access to the resource.

All of these earlier projects and reports were leveraged in NCHRP Project 20-109. It is the first comprehensive assessment of the TRT. The assessment criteria are included in Appendix C of this report and the full assessment can be found in NCHRP Project 20-109 Interim Report 1. The assessment and its contextualization in TRB's broader mission, vision, goals, and core services are the foundation for this strategic plan.

Mission, Vision, Goals, and Core Services of the Transportation Research Board

The current state and future direction of the TRT is best understood in the context of TRB's mission, goals, and strategies. According to TRB's Strategic Plan, its mission is

> . . . to increase the benefits that transportation contributes to society by providing leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal.

TRB, as an organization, has been a major source of information and research results pertaining to highways since the 1920s (although the organization has been known as TRB only since 1974). Information and research has been continuously disseminated through meetings of standing committees, the production of publications, and convening an annual meeting. The shift to a broader focus on transportation in 1974 was recognition of the increasingly multimodal nature of transportation. Over the years, TRB's portfolio of services has continued to grow. The organization annually engages more than 7,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise as volunteers by participating on TRB committees, panels, and task forces. TRB provides services and products that represent all major modes of transportation and a range of technical disciplines including engineering, planning, research, operations, policy, and more. The strategic direction of the TRT should align with and support the mission and strategic goals of the organization. The TRT must have a broad scope and coverage to match the organization's mission, and it must represent the language and vocabulary used by many diverse stakeholders.

TRB's vision statement tells us that "TRB is where the nation's leaders and the transportation community turn for research, innovation, collaboration, and advice on emerging and critical transportation issues" (www.trb.org/AboutTRB/StrategicPlan1.aspx). The intent is for TRB to be impactful, informative, and resourceful. To be *informative*, TRB generates and disseminates information on long-term issues of strategic, national, and international importance, as well as on practical, technical information and shorter-term policy issues. To be *resourceful*, TRB delivers information and research efficiently and effectively to inform decisions on investments, practices, and policies and promotes collaboration across the diversity of participants in TRB programs and activities. To be *impactful*, TRB identifies emerging issues, brings practitioners and researchers together to promote innovation in transportation, and provides government with advice on major transportation issues. The TRT does not currently have a vision statement. A vision statement has been proposed that aligns with TRB's vision and is tied to the TRT's future strategies.

TRB achieves its mission and vision through eight distinct goals. Of the eight, three pertain directly to the TRT (Goals 2, 4, and 7). Two other goals leverage the value and support of the TRT indirectly (Goals 3 and 5).These goals are described as the following:

- **Goal 2.** Conduct and promote knowledge creation and dissemination, especially on innovative practices and technologies in the transportation sector. *The TRT supports this goal by representing and promoting the language of transportation across the TRB and representing all variations of that language across the globe. The TRT's strategic direction has a direct relationship to TRB's efforts to achieve this goal.*
- **Goal 3.** Provide timely and informed advice on transportation and transportation-related issues to decision makers and others who are responsible for multimodal transportation systems. *The TRT can serve as a tool for bridging the languages of multiple perspectives, roles, and modes of transportation. The TRT's strategic direction has a direct relationship to TRB's ability to achieve this goal.*
- **Goal 4.** Act as an effective and impartial forum for the exchange of knowledge and information, including transportation and its relationship with social, economic, environmental, and other issues. *The TRT supports this goal by ensuring that all aspects of the field of transportation—its theory and practical applications—are represented in the language of transportation. The TRT's strategic direction has a direct relationship to TRB's efforts to achieve this goal.*
- **Goal 5.** Promote collaboration on transportation research, education, and technology transfer at international, national, regional, state, and local levels; across public and private sectors; and with transportation providers, customers, and other stakeholders. *The TRT can play a vital role in supporting effective communication across these functions, geographical locations, and sectors.*
- **Goal 7.** Conduct and promote communications efforts to enhance the awareness of transportation research and its contributions to innovation and progress in transportation. *Communication is grounded in language, and the language of transportation includes all stages of the life cycle of transportation, all modes, and all points of integration with and impact on society. The TRT can contribute to TRB's ability to achieve this goal.*

The role of information is also evident in TRB's core services, which are essential to achieving its goals. TRB's core services include (1) providing a forum for information exchange, (2) managing research programs, (3) developing policy analysis based on objective data and research, and (4) disseminating transportation research results. Each of these four core services is heavily focused on information and research results. The TRT is a critical enabling tool supporting access to and discovery of TRB's information. It is directly relevant to all four core services:

- **Providing a forum for information exchange.** The TRT must serve as an enabling tool for access and discovery of the information generated by the organization's 200 standing

committees, annual meeting, and specialty conferences (Core Service 1). The members and participants of these forums represent global and local perspectives on transportation issues, multiple languages, and different ways of talking about common issues. Information is also exchanged among TRB members, volunteers, affiliates, and other interested parties through webinars, e-newsletters, and social media. This presents a rich information management environment for TRB and a challenging role for the TRT.

- **Managing research programs.** TRB administers federally funded applied research programs on behalf of other organizations (Core Service 2), including NCHRP, FHWA, ACRP, TCRP, FTA, the Congressional Research Program (related to freight, rail, and hazardous materials), the Innovations Deserving Exploratory Analysis (IDEA) programs, and the Strategic Highway Research Program (SHRP 2). These research programs represent a broad scope of transportation issues, a rich set of stakeholders, and a diverse array of domain-specific languages. The TRT is critical to achieving this goal as it provides a tool to integrate and translate across these domains.

- **Developing policy analysis based on objective data and research.** TRB has conducted more than 100 studies at the request of Congress, federal agencies, the states, and other organizations on complex and sometimes controversial transportation topics of national importance (Core Service 3). Research requires access to TRB and non-TRB data, information, and expertise. The TRT is a potentially valuable tool for finding and accessing these resources regardless of where they reside.

- **Disseminating transportation research results.** The TRT is a key information access and discovery tool for TRB members and others looking for TRB products and services (Core Service 4). TRB disseminates transportation research results in several different types of documents and formats. TRB manages and maintains TRID—the world's largest and most comprehensive bibliographic resource for transportation information. The TRT does and can play an increasingly effective role in facilitating search and discovery in TRIS.

The TRB's Strategic Plan and its core services make a clear and obvious case for strong information management, information access, and discovery. The TRT is a critical enabler for TRB's achievement of its mission, vision, and goals. It is clear to the stakeholders consulted in this project that the TRT is a critical success factor in TRB's future. The challenge is to ensure that the strategic direction of the TRT aligns with the strategic direction of the TRB.

TRB Situational Analysis and Role of TRT

As part of the strategic planning process, TRB recently undertook an extensive review of the transportation environment and TRB's current activities. An environmental scan was conducted to identify the factors influencing the strategic direction of the organization. An examination of the strengths, opportunities, limitations, and challenges (rewording the classic SWOT analysis of strengths, weaknesses, opportunities, and threats) was conducted in consultation with TRB stakeholders. A survey and a set of focus groups involved the full array of TRB leaders, including TRB Executive Committee and Technical Activities Council members, volunteer chairs of TRB standing committees and research panels, representatives of the University Transportation Centers, and TRB staff. As TRB moves forward, the TRT must also move forward. Four of the opportunities in the TRB Strategic Plan pertain to the TRT:

1. **The TRB can uniquely address critical issues in transportation that are crosscutting and that have far-reaching impact.** The TRT is a language and vocabulary tool that can facilitate the crosscutting nature of transportation in its expanded scope and coverage.
2. **The emergence of new services, service providers, and institutional paradigms can bring in new partners and financial sponsors. Increasing international interest and engage-**

ment in TRB can facilitate the diversification of the organization and the breadth and depth of its dialog. The project team understands that there is interest in working with the TRT in both the private sector (i.e., transportation industry, transportation manufacturing, insurance, etc.) and the public sector. There are opportunities for building new partnerships across these different communities. As a core language tool, the TRT is pivotal for facilitating multilingual conversations and for enabling access to transportation information that is published in all of the languages of transportation.

3. **Leveraging technology to create and market new products and to improve communications with its volunteers and customers.** It is not unusual for a 20th century traditional thesaurus to transition to a 21st century semantic tool to support smart search, automated categorization, automated indexing and abstracting, and other intelligent applications.

4. **Implementing new revenue-generating services and activities.** As a 21st century semantic tool, the TRT provides a foundation from which to expand current and future products and services in an efficient and effective way. The TRT can also provide a foundation for community-based development of new products and services.

The TRT can support two TRB challenges identified in the situational analysis:

1. **Continuing competition with other organizations and increasing competition with other emerging avenues for information dissemination.** The project team observed this challenge during the assessment. There has been an explosion of smart search systems and of commercial databases producing and promoting transportation-related information since the TRT was first conceived and developed. In addition, there has been an explosion of relevant transportation information at the state and local levels that is of value to the broader community. This presents two opportunities for the TRT—it must be increasingly diverse in order to represent and connect local and regional variations of language, and it must be increasingly granular to represent the type of applied, operational, and administrative language used by transportation practitioners. It must go beyond the "research" perspective.

2. **Society's growing desire and expectation for quick answers to complex questions.** At the present time, the TRT is used for formal human indexing, but there is potential for it to serve a role in intelligent discovery systems.

Transportation Research Thesaurus— Mission, Vision, and Goals

The TRT is supported by two publicly available versions of a mission statement. The first is from the TRT home page (http://trt.trb.org) and reflects the historical intent from the late 1990s: "The Transportation Research Thesaurus (TRT) is a tool to improve the indexing and retrieval of transportation information. The thesaurus covers all modes and aspects of transportation. The TRT's purpose is to provide a common and consistent language between producers and users of transportation information." The second mission statement is found on the web page "About the TRT" section, which describes the TRT's mission this way: ". . . to provide a common and consistent language between producers and users of the Transportation Research Information Services (TRIS) Database (https://trid.trb.org). The thesaurus is currently also used as an indexing tool for federal, state and university collections." This second mission statement is more comprehensive and better suited to a strategic future for the TRT.

While the TRT has a mission statement, it has neither a published vision nor goals. Over time, the development of the TRT has not kept pace with the mission, vision, and strategies of TRB and the transportation community that it serves. The original warrant, as represented in the first mission statement, was very well defined and suited to its context within TRB. Over the years,

though, the interpretation and implementation of that warrant has varied and is not consistent with the second mission statement. In addition, while the field of information management and discovery has made significant advances, the TRT has remained largely an operational tool for indexers and catalogers. The potential for the TRT in the 21st century information environment is significant. NCHRP Project 20-109 identifies four strategic futures for the TRT. These strategic futures constitute a vision for the TRT that is consistent with the vision of the TRB and reflects the current and future state of information management, search, and discovery (Figure A-1).

Strategy 1 positions TRB to contribute directly to TRB's future vision. Strategy 2 enhances TRB's vision to be an informative organization. Strategy 3 enables TRB to be resourceful and effective in delivering services. In Strategy 4, the TRT is directly impacting TRB's information role. Each of these strategies adds exponentially to TRB's vision and mission.

Situational Analysis of the TRT and All Previous Evaluations

As of December 2016, the TRT contained a total of 12,125 terms, 9,591 of which were approved for indexing in TRIS and 2,534 of which were non-approved, lead-in terms. These terms were spread across 21 facets or categories (Figure A-2).

The most heavily developed and used vocabularies are Vehicles and Equipment (Facet Q), Materials (Facet R), and Facilities (Facet P). No current facets are aligned with or focused on modes of transportation. Many of the vocabularies are underdeveloped in terms of 21st century transportation issues (e.g., management topics, intelligent transportation topics, economic and social aspects, etc.). In order to gain a sense of the TRT's current status and its capabilities to support the mission of TRB, the project team defined an assessment framework that included five dimensions. The five dimensions included (a) thesaurus content, (b) thesaurus access and use, (c) thesaurus governance processes, (d) thesaurus scope and coverage, and (e) thesaurus
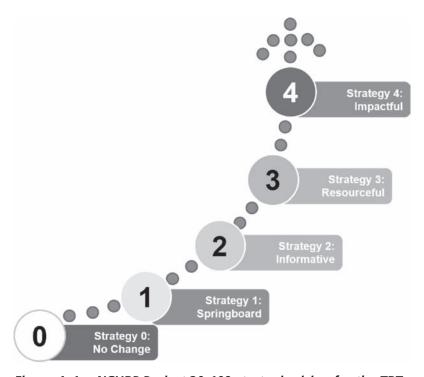
Figure A-1.   NCHRP Project 20-109 strategic vision for the TRT.

```
Transportation (Facet A)
Transportation Operations (Facet B)
Management and Operations (Facet C)
Communications and Control (Facet D)
Planning and Design (Facet E)
Construction and Maintenance (Facet F)
Testing (Facet G)
Safety and Security (Facet H)
Environment (Facet J)
Economic and Social Factors (Facet K)
Persons and Personal Characteristics (Facet M)
Organizations (Facet N)
Facilities (Facet P)
Vehicles and Equipment (Facet Q)
Materials (Facet R)
Physical Phenomena (Facet S)
Disciplines (Facet T)
Mathematics (Facet U)
Areas and Regions (Facet V)
Time (Facet W)
Information Management (Facet X)
```

Figure A-2.   TRT's 21 facets.

application architecture. Based on established practices in the peer-reviewed literature and the research team's experience, 134 criteria were identified across the dimension. In addition, as the data collection progressed, an additional 32 criteria were identified pertinent to the TRT's current context. Data to support the 166 criteria were collected from five sources, including:

- A transformed working copy of the TRT that enabled the research team to generate reports and statistics,
- Stakeholder interviews and feedback,
- Online data collection from Transportation Librarians (in lieu of interviews),
- Targeted input on information sources and use practices from Subject Matter Experts, and
- Focused searching of peer transportation information sources.
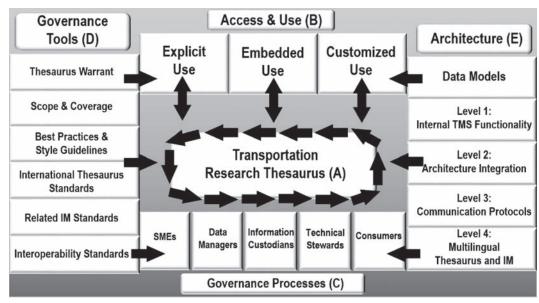
A high-level summary of the assessment is provided here in the form of a situational analysis. The situational analysis identifies strengths, opportunities, and challenges or weaknesses that should be addressed. The strategies proposed in this plan derive from the assessment. The detailed assessment methodology and criteria are presented in Appendix C of this report.

The project team identified strengths of the TRT across all five dimensions of the assessment framework (Figure A-3). Specifically, we found:

- A rich foundation of thesaurus terms focused primarily on road transportation,
- Strong use by and support from indexers and catalogers of transportation information,
- Strong participation by the library community in the current governance model,
- Strong interest in the potential of the TRT among subject matter experts who were previously unaware of the TRT.

The project team identified several opportunities for the TRT, primarily focused around scope and coverage and architecture. These opportunities are described below:

- Search tests suggest there are significant opportunities to expand the scope and coverage of the TRT by leveraging literary warrant more extensively through the integration of transportation-focused terms from search logs, commercial sources, and other transportation vocabularies.

IM = information management

Figure A-3. Assessment framework for TRT.

- Gaps in coverage when we compared the TRT terms to other transportation thesauri and controlled vocabularies. While the overall "no match" rate was high, there is opportunity in the potential links between terms from other thesauri and terms in the TRT.
- Shift the focus of the management of the TRT from term levels to facet level and ensure that there are clearly defined roles and responsibilities for managing and developing each facet. Consult with subject matter experts as part of the development process to ensure that current topics and issues are included in the thesaurus when they are discussed.
- There are many functions that are not currently provided within this context. The opportunity here is to review these functions, especially those that support interoperability, export, and governance reporting to determine how the architecture might be supplemented to address more of the requirements in this dimension.

The assessment surfaced challenges across all five dimensions of the framework. While these are categorized as weaknesses, the project team treated them as opportunities in developing future strategies. The TRT content challenges include

- Non-standard use of thesaurus relationships, particularly NT/BT and RT/Hierarchical.
- Insufficient metadata for individual terms to support the management of terms.
- Enumerated hierarchy as fundamental structure of the TRT.
- Lack of functional thesaurus management capabilities.
- Lack of lifecycle management/term workflow capabilities in the system.

Challenges in the access and usability dimension were noteworthy and speak to the TRT's ability to serve as an effective tool for supporting the information and knowledge-related goals of TRB. These weaknesses include

- Low usage rates of the TRT terms in searching within TRID and low incidence of these terms in other transportation-related repositories and sources.
- Lack of awareness of the TRT among subject matter experts.
- Use of the TRT-guided searching in TRID is low.
- TRID content is discovered through Google redirects—indexing terms do not generally align with terms in search logs.

- Varied rates of use of terms across facets. Terms with greater specificity appear to more often align with search terms than the more general terms.
- Terms that are not transportation related or transportation specific (contextualized for transportation perspectives) are used for indexing but not for searching. Disconnect between the rate at which terms are used for indexing and the rate at which they are searched.
- Rigid design of the TRT for use in TRID constrains its use outside of its current context. As the state level IM/KM environment has developed, there are more opportunities to use the TRT outside of TRID. The structure, though, constrains this.
- Hierarchy is difficult to navigate by end users. Hierarchical structure is rarely used by indexers, librarians, or searchers. Subject matter experts also commented that it does not reflect their view of their domains.
- Posted functionality does not work—Key Word In Context and Key Word Out of Context are simple launched searches and do not reflect traditional thesaurus functions.

While governance is an area of strength in its engagement of the library and information science community, it also exposes a weakness in the lack of representation of transportation subject matter experts. The architecture and design factors also introduce challenges to managing the TRT at a facet rather than a term level. The project team offers several challenges that should be treated as opportunities for improvement:

- Because the TRT is designed to support TRID as an explicit indexing and guided search tool, it lacks basic thesaurus management functions. Without these functions, it is difficult to address the usability weaknesses.
- Engagement with subject matter experts or the general membership of TRB was a weak area.
- Level of focus of governance is on individual terms. Governance to ensure that the TRT was aligned with the TRB language was an area of weakness.
- Found that the functional design of the TRT was not conducive to managing at the facet level.
- Change management is difficult in the current architecture, and this constrains the decisions made by governance groups.
- Governance bodies do not include transportation experts but rather draw primarily from library and information science. There is a need for transportation expertise to be drawn into the governance and oversight process.
- There is a term suggestion page on the TRT home page, but the decision process and the feedback loop are not tightly integrated. There is no formal mechanism to respond to individuals who have submitted terms. And, there is no formal explanation of why a term was not accepted. Also, the current application does not allow the tracking of proposed but not approved terms.
- Placement of a proposed term in the hierarchy is sometimes a key decision point.
- The current application does not provide easy generation of management reports that might inform governance processes or serve to communicate with the TRB community.

A major weakness is the lack of adherence of the TRT to ISO 25964, the international standard for thesaurus structures and functions. This presents a challenge for exchanging content with the broader community. The current scope and coverage of the TRT is another noteworthy weakness. At this time, the TRT does not provide consistent coverage of the transportation topics and functions that are addressed by TRB's annual and specialty conferences. Specifically, the project team made the following observations:

- Literary warrant was well defined when the TRT was created but the implementation of the TRT has not maintained the spirit of that warrant. The TRT's scope and coverage is not comprehensively aligned with TRB's current coverage of the field. The research team notes

that the overall results of this section suggest a shift to a proactive literary warrant strategy to support the TRT.

- Confusion among the library and information science community supporting the TRT as to whether it is a thesaurus or a classification scheme. A classification scheme defines "classes" or groups of things. It is used to create collections of content that treat the class in a substantive way (i.e., a library collection is physically organized into classes on the shelf). A thesaurus is a controlled vocabulary of concepts and terms that is intended to index the concepts that are treated inside a single content object (i.e., think of the concepts that are included in a back-of-the-book index). In addition to simply exposing the concepts, a thesaurus tells us something about how the concepts relate to one another in a particular context. The confusion in the library and information science community derives from the confounded use of terms and classes as subject headings in traditional library catalogs. As the constraints of those catalogs are no longer an overriding factor in the design of search systems, we can use thesauri and classification schemes to their best advantage.
- Lack of modal-specific vocabularies was a clear weakness identified by subject matter experts.
- The TRT aims to be modal agnostic, but the result is a non-modal treatment of transportation information and research. Modal agnostic in a different architecture can support modal-specific, multimodal, economic, and social perspectives, as well as life cycle stages, etc. However, the current hierarchical architecture precludes this flexibility.
- Search tests and Google Analytics reports suggest that there are gaps in coverage between what searchers are looking for and what the TRT is covering.
- Search tests also highlighted the generic nature of many terms in the TRT and gaps of transportation-specific terms. Of particular note was the lack of alignment of the TRT's overall structure with TRB Annual Meeting program areas.
- A simple review of the 21 facets of the TRT surfaced only two that would be recognized as transportation-focused from their names. Nineteen of the facets are generic and might be understood to be part of any other controlled vocabulary.

While there is only one general weakness or challenge described for architecture, the architecture has broad impacts for the current and future development of the TRT. In essence, the current architecture is constrained by the tight coupling of the TRT to TRID indexing and searching. This is a 1990s architectural solution, which is not aligned with current commercially available thesaurus management applications. This architecture negatively impacts all other dimensions of the framework.

TRT Future Strategies

The project team prepared five use case scenarios as part of the assessment and to frame the discussion of future goals and objectives. The use cases have been transformed to strategies for this stage of the project. Four of the five strategies propose enhancements to the TRT that will address weaknesses and challenges and offer opportunities to support TRB's information mission, vision, and goals (Figure A-4). Strategy 0 represents a "no change" status quo option that essentially freezes the TRT in its current architecture and role as a search component of TRID. Strategy 1 is a springboard strategy which provides a new platform from which the TRT may generate additional value for the TRB community. Strategy 1 presents a first step forward in developing additional noteworthy value in the area of information access and discovery. This strategy has the potential to expand the coverage of transportation information and to increase the sophistication of access and discovery. Strategy 2 describes an expanded role for TRT in the broader context of access to and discovery of transportation data and information. Strategy 3 describes additional products and services that may
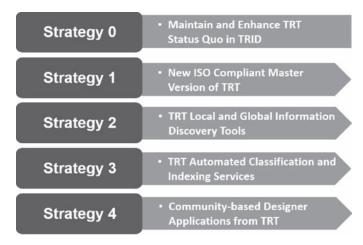
Figure A-4.   Future strategies for the TRT.

be offered by TRB, leveraging the Thesaurus as a core component of automated indexing, categorization, and abstracting applications. Finally, Strategy 4 describes a situation where the Thesaurus is a catalyst for community-based product and service development.

Strategy 0—No Change, No Enhancement

This strategy focuses on small changes that can be made to the current version of the TRT as it is designed within TRID (Figure A-5). This is a "no change" option, which essentially does not speak to a strategic direction for the TRT's future. It assumes that in the future the TRT may not continue to play a role. This strategy does not propose an expanded role for the TRT in the TRB information environment in the future. The enumerated hierarchy remains the main architectural and functional design feature of the TRT. Any new value that is achieved is outside of the context of the TRT in TRID. Strategy 0 aligns with Use Case 1 in Interim Technical Report #1.

Proposed action items are the following:

1. Develop a fully elaborated Guiding Principles and Best Practices document that details policies and practices pertaining to (1) individual thesaurus term selections and the basis for
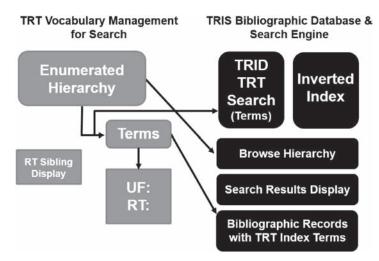
Figure A-5.   Strategy 0—no change, no enhancement to TRT.

inclusion/exclusion; (2) the use of ISO-standard thesaurus relationships; (3) forms of terms (i.e., concepts rather than single words, exclusion of stand-alone adverbs and adjectives, etc.); and (4) business rules imposed by TRID search.

2. Develop a strategy for the inclusion and use of synonyms, specifically regional synonyms, full and quasi-equivalence, abbreviations, acronyms, and lexical variants.
3. Review all previously non-approved terms for inclusion as synonyms.
4. Reduce or eliminate the use of parenthetical qualifiers.
5. Reduce or eliminate the use of structural terms.
6. Review selected facets for transportation-specific enhancement (Facet A, B, C, D).
7. Develop New Term Tracker to support suggestion and feedback on new terms.
8. Develop an on-site and online workshop on standard thesaurus practices for the transportation community.
9. Develop a strategy to improve awareness of the TRT throughout TRB.
10. Develop an ISO-compliant, exportable version of the TRT for other TRB members to work with. This involves mapping the TRT/TRID database structure and the hierarchy to relationships that are definable and understandable to other thesaurus management tools.

Strategy 1: New ISO-Compliant Master Version of TRT

Strategy 1 is a springboard that moves the TRT to a standards-based platform from which it can pursue a range of opportunities (Figure A-6). This strategy assumes that the TRT has a strong and vibrant future supporting TRB's mission and vision. This strategy shifts the TRT master version from its current function in TRID to an ISO-compliant thesaurus management application. The relationship with TRID remains strongly in place, and the functionality of TRID is maintained. In this scenario, TRID becomes a consuming system that draws content from the TRT through a new web services function. There are several options for supporting web services, ranging from simple extract/transfer/load of content from TRT into TRID to the more sophisticated direct and on-demand query of TRT content by TRID via APIs (e.g., application programmer interfaces). The former may be more appropriate for a proof of concept and the latter for a full implementation, when a formally procured thesaurus management tool is in place. We understand that the
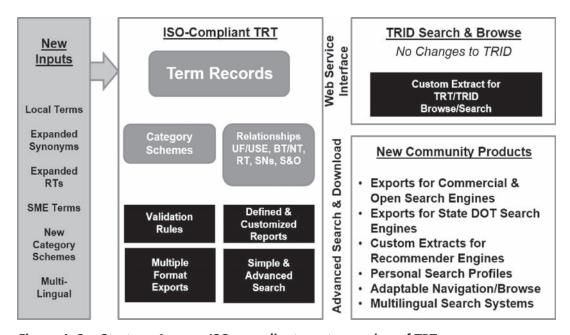
Figure A-6.   Strategy 1—new ISO-compliant master version of TRT.

preference by the TRT team is for the interactive query option. The current functionality of the TRT within TRID will be decommissioned and archived. A new ISO 25964 version of the TRT is implemented in a new commercial off-the-shelf thesaurus management application. A new master version of the TRT is provided by the project team in a proof of concept phase. Strategy 1 aligns with Use Case 2 in Interim Technical Report #1.

This strategy provides new functional capabilities that will facilitate management of the thesaurus at the facet or domain level; will provide the TRT support team with a range of tools that are currently lacking; and will make it much easier to provide full, partial, or custom extracts and exports from the TRT for TRB community members. This strategy also provides simple and advanced search of the TRT, the ability to expand its scope and coverage without adversely impacting TRID, and the option to implement an expanded categorization scheme to support all modes of transportation. This scenario also gives the TRT management team and the consuming community time to explore options, to work with a full-function thesaurus management tool, and to explore which features and functions are required for future use.

Proposed action items are the following:

1. Acquire an ISO 25964 compliant thesaurus management application tool "for convenience." The selection of such a "convenience copy" supports learning and requirements formalization and the working copy of the TRT.
2. Prepare a Guiding Principles and Best Practices baseline document to manage and include all current and any changes to policies and practices as TRB works through the future strategies for TRT. This action item should also include a renewed discussion of the warrant of the TRT, given a new vision of the future.
3. Configure an ISO 25964 compliant working tool, including (a) define standard relationships; (b) define administrative and note fields, including date created, language base value, unique term identifiers; (c) define local note fields to track Source and Origin for locally contributed terms; (d) define standards-based exports; (e) define term status values; (f) define term workflow stages and set values; (g) review standards and advanced search options; (h) configure current TRT Enumeration as an internal categorization scheme; and (i) update the Guiding Principles document to reflect configuration decisions and actions.
4. Transform the current version of the TRT to an ISO-compliant version including the import of all approved and equivalence (non-controlled) terms, creation of formal records with administrative fields for each TRT term, creation of formal BT/NT relationships, the move of the current TRT enumeration codes to a controlled category field, and the integration of previously non-approved terms with "not-approved" status and notes as to reasons for rejection.
5. The indexer interface of TRID will also be linked to the new thesaurus management system through either extraction-transformation-load or interactive web services strategies. The shift in underlying TRT source would be invisible to the indexers. At the TRID system level, a "cut over" date would be defined for switching the behind-the-scenes access from the TRT component in TRID to the new thesaurus management system. Consultation of the TRT by indexers will be targeted to the new thesaurus management system on the "cut over" date. The new system will provide for enhanced access, searching, and discovery of terms for manual indexing.
6. Conduct a quality control review of the new working copy at the facet level.
7. Define and prioritize changes to the working copy of TRT, including all of those defined in Strategy 0, eliminating the RT Hierarchical relationship type, conducting a detailed evaluation of BT/NT relationships, and expanding the coverage of synonyms and associative terms.
8. Select two current facets for a proof of concept for implementing changes in the working copy, including a detailed assessment and rebalancing and working with the detail provided in Technical Report 1.

9. Define and integrate new category structure and begin filling in vocabularies for new facets, including expansion of modal-specific and cross-modal vocabularies. At the present time, the TRT is modal agnostic.
10. Continue to review and update the remaining facets of the TRT.
11. Develop and test web services integration to TRID, the current TRIS/TRT business rules, including review of fields required for search, review of the impact of expanded synonyms and associative terms on search expansion, and the impact on current and past indexing practices.
12. Release the new automatically generated KWIC, KWOC, term browse, and high-level category browse features. These functions are built into the thesaurus management system as standard predefined reports and functions.
13. Enable TRT search and reporting functions via the web to support external user access and downloads.
14. Review the governance models proposed by the research team and establish a community-wide dialog on (1) shifting governance to the facet level; (2) expanding the involvement of subject matter experts and external stakeholders in the governance process; (3) ensuring that governance processes are fully specified and operationalized; (4) ensuring that all supporting guidelines and decision support tools are in place; and (5) establishing strong communication channels to and from stakeholders. This action item would also consider the role of the TRT Subcommittee in the larger governance model.
15. Review the need for additional roles, responsibilities, training, and skills building among the TRT support team and user community including formal roles for project management, managing facets of the TRT, and expanded community engagement.
16. Prepare for formal procurement—defining requirements, reviewing the market, determining the procurement strategy, organizing the procurement committee, including considerations of state and other stakeholder involvement, and identifying funds to support the procurement.

Strategy 2—TRT Local and Global Information Discovery Tools

Strategy 2 aligns with TRB's vision to be informative (Figure A-7). Strategy 2 enhances search and discovery in TRID, at the state DOT level, in commercial sources, and in the open web. This strategy improves search and discovery of transportation information by introducing and
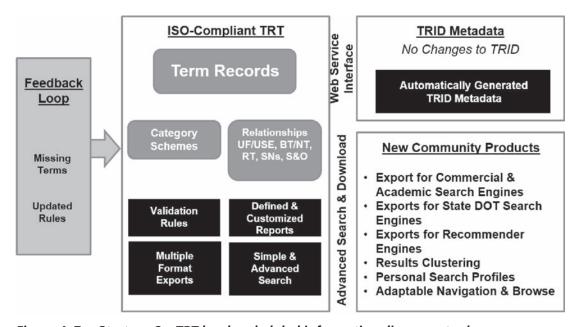
Figure A-7.   Strategy 2—TRT local and global information discovery tools.

enabling smart search capabilities. It sets the stage for multilingual search. This strategy shifts the focus of managing the TRT from individual terms as entities to facet vocabularies. It provides the opportunity to ensure that the TRT is reflecting the language of a transportation domain and all of those who are engaged in that domain. It also enables the TRT users to leverage individual terms as digital entities across sources—with the addition of DOIs for individual terms. This provides a foundation for the National Transportation Library, TRB, and other consumers of the TRT to implement linked data solutions. This strategy also facilitates the inclusion, sharing, and tracing of local transportation terms—variant synonyms as well as unique terms—in the TRT. It also enables TRB and the TRT consumers to recommend the inclusion of more granular terms and terms that reflect the priorities of practitioners across the country (e.g., performance indicators, KPIs, etc.) All of these enhancements to the TRT provide a basis for more intelligent search. This strategy also provides the opportunity to review, configure, and implement multilingual search architecture and to begin creating or integrating other language versions of the TRT terms. This strategy presumes the engagement of the broader transportation community. Strategy 2 aligns with Use Case 3 in Interim Technical Report #1.

Proposed action items are the following:

1. Review the current functionality and value of the TRID search interface including search and browse structures and the ability of TRID search to leverage the expanded TRT content and functionality.
2. Encourage and enable external use of the TRT including the development of standard and customized exports of the content to support external use throughout the transportation community, encourage local use and feedback, and encourage publishers to use the TRT in their search systems.
3. Develop documentation and an online tutorial for external web service users.
4. Encourage the inclusion of local terms and configure the TRT database to track source and origin of local terms.
5. Develop strategy for multilingual architecture and multilingual search.
6. Expand the external user community around the TRT to support the expansion and improvement of transportation search and discovery.
7. Create DOIs for individual TRT terms by leveraging the unique identifiers created in Strategy 1.
8. Reframe governance to include external subject matter experts and the TRT user community and assign new responsibilities for overseeing and developing facets of transportation.
9. Augment the Guiding Principles and Best Practices document to reflect the new search and discovery functionality and to reflect any changes in coverage.
10. Review the impact of Strategy 2 on TRB TRT competencies, roles, and staffing.
11. Update KPIs as needed.

Strategy 3—TRT Automated Classification and Indexing Services

Strategy 3 speaks to the resourcefulness element of TRB's vision (Figure A-8). This strategy focuses on enabling new products and services, grounded in the TRT, and leveraging existing resources more efficiently and effectively. Strategy 3 enhances the new Master Version of the Thesaurus by providing services for Automated Categorization and Indexing across the transportation community. This strategy supports new products and services from TRB and also from the state or local levels. Strategy 3 is best implemented after Strategy 2 has been realized. However, TRB can begin discussions around Strategy 3 while Strategy 2 is in progress. Strategy 3 aligns with Use Case 4 in Interim Technical Report #1.

At the end of Strategy 2, TRB is positioned to provide revenue-generating, automated categorization, indexing, and abstracting services to members at the state level, to research institutes,
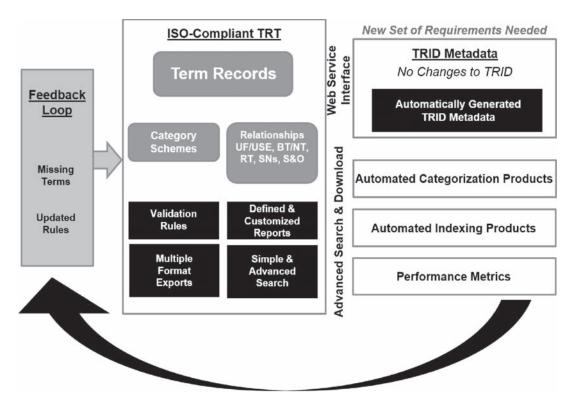
Figure A-8.   TRT automated classification and indexing services.

to other information providers, and to industry clients. Without Strategy 2, though, the TRT's content may not be sufficient to provide quality services. This strategy also represents a shift in approach from manual to automated indexing and categorization. It allows TRB to expand both indexing and categorization, compared to the current subject description. This strategy does require new staff competencies and roles. Its impact on information management and discovery in the field of transportation can be significant. It is important, though, for TRB decision makers and the TRT management team to have a strong understanding of the functionality underlying these tools. This strategy will require a new procurement. This strategy also presents opportunities for cooperation and collaboration with states, with the National Transportation Library, with other international transportation partners, and with those in the private sector transportation and transportation-related industries.

Action items are the following:

1. Explore and develop models for these potential new services.
2. Gauge the demand for these services and gauge the costs and revenue generation.
3. Assess the role of the National Transportation Library in providing these automated services.
4. Review the TRT's ability to support these new services by assessing current facets for categorization and assessing the terminology for indexing.
5. Develop profiles and rules for new applications, including building the categorization profiles, building new abstracting rules as required by the abstracting engines, and defining threshold levels for conceptual indexing.
6. Prepare for procurement, and review the market for relevant tools.
7. Develop a proof of concept project working with two state DOTs.
8. Configure and implement the selected tool.
9. Develop a transition plan from manual to automated services.
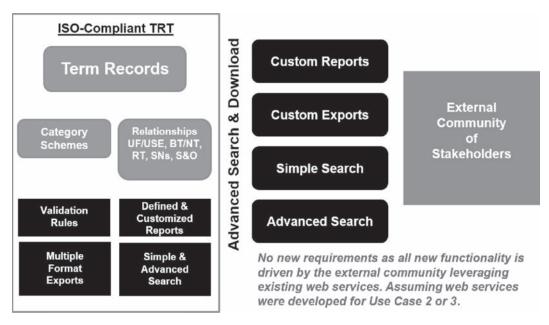
Figure A-9.   Community-based designer applications from TRT.

10. Assess KPIs for manual and automated indexing, updating previous KPIs as needed.
11. Review the impact of the use case on TRB competencies, roles, and staffing.

Strategy 4—Community-Based Designer Applications from TRT

Strategy 4 draws from enhancements made in Use Case Scenarios 2, 3, and 4 and sets the stage for community-based development of new information products and services grounded in the TRT (Figure A-9). This is an impactful strategy, which shifts the value of the TRT from direct to enabling and leverages the TRT as a tool. The expectation is that enhancements and innovations will also be identified for the TRT as a semantic analysis tool. TRB's role in this strategy is as an enabler. TRB's responsibilities would be to ensure that custom exports and extracts are available to external developers and to provide the search and browse capabilities that allow external developers to see the TRT content from different perspectives. Governance expands to include one new area—the promotion of Creative Commons agreements that support development but also enable back feeding and sharing across the community. Strategy 4 aligns with Use Case 5 in Interim Technical Report #1.

Action items are the following:

1. Update web services and ensure that external users can define and launch custom extracts.
2. Design and implement Creative Commons licensing with full attribution and acknowledgement of source and obligations to share back.
3. Create a virtual storefront for sharing custom-built applications.
4. Expand and support an open TRT information focused user community.
5. Ensure that the online tutorial and documentation reflects current strategy.
6. Seed the storefront with existing thesaurus-based, user-developed applications, such as custom transportation dictionaries, mouse-over definitions for transportation terms, web page developers' link for term lookup, "Did you mean?" disambiguation services, digital thesaurus applications, visual thesaurus displays, and use in business rule repositories, etc.
7. Update KPIs to reflect new functionality delivered by this strategy.
8. Assess the impact of the use case on the competencies, roles, and staffing.

APPENDIX B

Immediate Action Plan

Purpose of the Action Plan

The research team was asked to devise an immediate action plan that those responsible for the TRT's management might use to initiate implementation of a strategic plan. The purpose of this action plan is to provide a practical path forward to enhance the TRT based on the outcome of the TRT assessment strategic plan developed in the research described in this report. The plan's scope is based in Use Case 1.

Project Plan and Timeline

The project plan and timeline (Table B-1) shows tasks and subtasks comprising the immediate action plan; a calendar (by week) and duration for each task; dependencies between/among tasks (represented as a predecessor task number); responsible roles; and other resources, such as IT support and training, needed to accomplish the task. Estimates of approximate cost were provided as well, but depend on assumptions about when and how work would be accomplished; these estimates are not presented here.

Roles and Responsibilities

A key component to moving from the strategic plan to the action plan is the organization of the project. The research team has identified the following roles and responsibilities for project personnel:

- **TRB Project Director**—has primary responsibility for the direction of the project, serves as the project's champion, interacts with other TRB staff regarding funding and resources, and receives reports from other project members.
- **Project Manager**—has primary responsibility for the day-to-day management and coordination of the project, monitoring of schedule and resources, and submitting deliverables. The research team suggests the use of outside resources for this position to provide support to the Director without jeopardizing the ongoing work of the TRT group.
- **Senior Thesaurus Analyst**—responsible for identifying issues and suggesting best practices and the most efficient and effective approaches to dealing with technical- and content-related tasks. Provides thesaurus management expertise to support the Project Director and the Project Manager in instituting these best practices within the environment and constraints of the project. The research team suggests the use of outside resources for this position to provide support to the Director without jeopardizing the ongoing work.
- **Thesaurus Analyst**—provides support to the Senior Thesaurus Analyst and other team members. May perform studies under the guidance of others and make changes to the content

70

under supervision. Performs quality control on the content under the direction of the Senior Thesaurus Analyst and the Project Manager.

- **Software Developer**—makes changes to the current TRT architecture and code to accommodate Strategy 1. Develops web services/APIs as needed to integrate the Master TRT with TRID and to make the TRT available to others. Provides expertise in the relevant programming languages, the current TRT architecture, and the TRB/NAS IT environment.

A leadership team would have responsibility for advising the technical team on the various aspects of the work. This team should include representatives from TRB, the TRT Subcommittee, and other interested stakeholders.

Table B-1 presents the components of this proposed immediate action plan and suggests which members of the development team would be assigned responsibility. These tasks and responsibilities might be revised to suit the specific situation—e.g., balance between TRB staff and contractor work—when implementation is initiated.

Table B-1.   Project plan and timeline.

Dependencies	Task Number	Task Name	Calendar	Duration	Project Personnel	Other Resources
	0	**Identify the Strategy 1 team members and the TRB Project Director**	Week 1	1 week	Project Manager, TRB Project Director	
	1.0	**Acquire an ISO 25964 compliant convenience tool to support Strategy 1**	Weeks 2 - 3	2 weeks		
	1.1	Determine number of copies needed for TRB team (fee expected)	1 day		Project Manager, TRB Project Director	Software
	1.2	Determine number of copies for other interested stakeholders (no fee expected)	1 day		Project Manager, TRB Project Director	
IR2	1.3	Identify low cost ISO-compliant desktop tool	5 days		Project Manager, Senior Thesaurus Analyst	
1.3	1.4	Download and install convenience copies	5 days		TRB Project Director, Software Developer	
	2.0	**Prepare a new Guiding Principles and Best Practices document**	Weeks 4 - 5	2 weeks		
IR1	2.1	Review and restate/reaffirm TRT warrant	5 days		Project Manager, Senior Thesaurus Analyst	
IR1	2.2	Develop table of contents and structure for guiding principles document	1 day		Senior Thesaurus Analyst	
IR1	2.3	Create placeholder section for term selection and inclusion policies	1 day		Senior Thesaurus Analyst	
IR1	2.4	Create placeholder section for term relationships	1 day		Senior Thesaurus Analyst	
IR1	2.5	Create placeholder section for categorization schemes	1 day		Senior Thesaurus Analyst	
IR1	2.6	Create placeholder section for use/consumption models	1 day		Senior Thesaurus Analyst	
IR1	2.7	Create placeholder section for annual evaluations and key performance indicators	5 days		Senior Thesaurus Analyst	
IR1	2.8	Create placeholder section for governance models	5 days		Senior Thesaurus Analyst	
2.1 - 2.8	2.9	TRB review of guiding principles strawman structure		2 weeks	Contracting Officer	
1.0	**3.0**	**Configure ISO 25964 compliant convenience tool**	Weeks 4 - 9	5 weeks		

(continued on next page)

Table B-1.   (Continued)

Dependencies	Task Number	Task Name	Calendar	Duration	Project Personnel	Other Resources
1.4	3.1	Install convenience copy of tool on TRB server	5 days		Software Developer, Senior Thesaurus Analyst	Software, server
1.4	3.2	Install convenience copies on desktops	5 days		Software Developer, Senior Thesaurus Analyst	
1.4	3.3	Install web version (if part of the convenience tool)	5 days		Software Developer, Senior Thesaurus Analyst	
3.1	3.4	Define standard relationships (BT/NT; USE/UF; RT/RT)	1 day		Software Developer, Senior Thesaurus Analyst	
3.1	3.5	Define variations of equivalent relationships (e.g., abbreviations, acronyms, initialisms, regional variants, etc.)	10 days		Project Manager, Senior Thesaurus Analyst	
3.1	3.6	Decide whether to continue to create Related Term (hierarchical) relationships	1 day		Project Manager, TRB Project Manager	
3.1	3.7	Propose and define variations of associative relationships (i.e., review of standards discussions)	10 days		Project Manager, Senior Thesaurus Analyst	
3.1	3.8	Define administrative fields (e.g., Date Created, Unique ID Format, Language Base Value, etc.)	5 days		Project Manager, TRB Project Manager, Senior Thesaurus Analyst	
3.1	3.9	Define note fields (e.g., Scope Notes with strategy, Definition field with strategy for content and specifications, Source & Origin field with strategy for values, Validation field with strategy for values, and personal notes)	5 days		Project Manager, TRB Project Manager, Senior Thesaurus Analyst	
3.1	3.10	Review standard term workflow values (e.g., approved, candidate, provisional, not approved) and accept or adapt	5 days		Project Manager, TRB Project Manager, Senior Thesaurus Analyst	
3.1, 3.2	311	Review existing simple and advanced search capabilities	5 days		TRB Project Director, TRB Others	
3.1	3.12	Review options for maintaining TRT enumeration codes (e.g., controlled categorization field or uncontrolled note field)	5 days		Senior Thesaurus Analyst	
3.1	3.13	Define new strategy for defining new TRT enumeration codes for new terms added to TRT	5 days		Project Manager, TRB Project Manager, Senior Thesaurus Analyst	
2	3.14	Revise Guiding Principles and Best Practices to reflect new term and thesaurus structures	5 days		Senior Thesaurus Analyst, Thesaurus Analyst	
3.14	3.15	TRB review of ISO 25964 compliant installation		2 weeks	Contracting Officer	
	4.0	**Transform the current version of TRT to an ISO-compliant version**	Weeks 9 - 17	8 weeks		
1.3	4.1	Establish simple term records using simple MS Excel table report from TRT, including approved terms and lead-in terms	1 day		Senior Thesaurus Analyst	Updated reports from TRT database, including new records added since 12/2016
4.1	4.2	Load in TRT enumeration codes for each record	1 day		Senior Thesaurus Analyst	
4.1	4.3	Copy records for every term to an import file—including all displayed relationships, definitions, scope notes—for every facet	20 days		Senior Thesaurus Analyst	

Table B-1. (Continued)

Dependencies	Task Number	Task Name	Calendar	Duration	Project Personnel	Other Resources
3.4, 3.5, 3.6, 3.7	4.4	Transform all relationship and field labels to import formats and add repeatable labels for import	20 days		Senior Thesaurus Analyst	
3.4	4.5	Establish new formal BT/NT relationships	5 days		Senior Thesaurus Analyst	
4.4	4.6	Import all existing records	3 days		Senior Thesaurus Analyst	
4.6	4.7	Generate administrative reports to track records	1 day		Senior Thesaurus Analyst	
	5.0	**Conduct a quality control review of the working copy at the facet level**	Weeks 18 - 21	3 weeks		
4.7	5.1	Generate reports of full records for each facet as working copies for quality control exercise	1 day		Senior Thesaurus Analyst	
5.1	5.2	Review full facet for inadvertent duplicates/variants and consolidate	5 days		Thesaurus Analyst(s)	
5.1	5.3	Review full facet for intentional duplications (terms with multiple TRT enumeration codes) and resolve	5 days		Thesaurus Analyst(s)	
5.1	5.4	Review individual records in the working copy and compare to the web published version in historical TRT	10 days		Thesaurus Analyst(s)	
5.2, 5.3, 5.4	5.5	Make a backup copy of the working copy for archive purposes	1 day		Senior Thesaurus Analyst	
5.5	5.6	TRB review of new ISO-compliant master version of TRT		3 weeks	Contracting Officer	
	6.0	**Define and prioritize changes to working copy of TRT**	Weeks 21 - 25	4 weeks		
IR1	6.1	Determine strategy for parenthetical qualifier	5 days		Thesaurus Analyst(s)	
IR1	6.2	Determine strategy for use of container terms	5 days		Thesaurus Analyst(s)	
IR1	6.3	Determine strategy for stand-alone adverbs and adjectives	5 days		Thesaurus Analyst(s)	
IR1	6.4	Determine strategy for synonyms	5 days		Senior Thesaurus Analyst	
IR1	6.5	Define strategy for use of associative terms	5 days		Senior Thesaurus Analyst	
3.14, 6.5	6.6	Revise Guiding Principles and Best Practices to reflect new policies and practices	5 days		Senior Thesaurus Analyst	
6.6	6.7	TRB review of updated guiding principles document		2 weeks	Contracting Officer	
	7.0	**Proof of concept for implementing changes in working copy**	Weeks 26 - 36	10 weeks		
IR1	7.1	Select two current facets for a full assessment and rebalancing, working with the detail provided in Technical Report 1	10 days		TRB Project Director, Project Manager	
IR1	7.2	Define impacts and expected improvements for each facet	10 days		Senior Thesaurus Analyst	
IR1	7.3	Assess the balance and second-level category values for each facet	10 days		Senior Thesaurus Analyst, Thesaurus Analyst	
IR1	7.4	Revise category structure for top level facets and identify gaps in coverage	10 days		Senior Thesaurus Analyst, Thesaurus Analyst	
3	7.5	Reconcile use of parenthetical qualifiers	10 days		Thesaurus Analyst(s)	
3	7.6	Reconcile use of container terms	10 days		Thesaurus Analyst(s)	
3	7.7	Reconcile use of stand-alone adverbs and adjectives	10 days		Thesaurus Analyst(s)	
3	7.8	Implement and test new strategy for synonyms, including all types of variants	20 days		Senior Thesaurus Analyst, Thesaurus Analyst	

Table B-1.   (Continued)

Dependencies	Task Number	Task Name	Calendar	Duration	Project Personnel	Other Resources
4.5	7.9	Review use of BT/NT relationships and make changes to bring variations in compliance with ISO 25964	20 days		Thesaurus Analyst(s)	
6.4, 6.5	7.10	Review all proposed but not accepted terms for inclusion as equivalence or associative terms	20 days		Thesaurus Analyst(s)	
6.5	7.11	Implement and test new strategy for associative terms	20 days		Senior Thesaurus Analyst, Thesaurus Analyst	
7.4, 7.5, 7.6, 7.7, 78, 7.9, 7.10, 7.11	7.12	Generate working reports for each facet	1 day		Senior Thesaurus Analyst	
7.12, 6.6	7.13	Formalize the review process for future evaluation and review	10 days		Senior Thesaurus Analyst	
	7.14	Assess the impacts and improvements of the changes	5 days		TRB Project Director, Project Manager	
IR1	7.15	Identify and prioritize additional facets for improvements	5 days		TRB Project Director, Project Manager, Senior Thesaurus Analyst	
7.15	7.16	TRB review of enhancements to new master version of TRT		3 weeks	Contracting Officer	
	8.0	**Define and implement the new category structure**	**Weeks 37 - 45**	**8 weeks**		
IR1	8.1	Review existing facets and determine which to retain in current form, which to combine and which to eliminate	5 days		TRB Project Director, Project Manager, Senior Thesaurus Analyst	
IR1	8.2	Identify facets in need of focus or expansion and write new scope notes	5 days		TRB Project Director, Project Manager, Senior Thesaurus Analyst	Continued access to SMEs for scope and coverage advice
IR1	8.3	Identify new modal-facets and prepare scope notes	5 days		TRB Project Director, Project Manager, Senior Thesaurus Analyst	Continued access to SMEs for scope and coverage advice
IR1	8.4	Identify topics that are not yet covered by TRT and write scope notes	5 days		TRB Project Director, Project Manager, Senior Thesaurus Analyst	Continued access to SMEs for scope and coverage advice
8.2, 8.3, 8.4	8.5	Write scope notes for all retained facets	15 days		Thesaurus Analyst(s)	Continued access to SMEs for scope and coverage advice
8.5	8.6	Create new category scheme and add all values	5 days		Senior Thesaurus Analyst	
8.6	8.7	Assign values to existing records using report function and import formats	15 days		Thesaurus Analyst(s)	
3.3, 8.7	8.8	Implement new browse structure on TRB TRT website for testing	20 days		Software Developer, TRB Project Director, Project Manager, Senior Thesaurus Analyst	

Table B-1.   (Continued)

Dependencies	Task Number	Task Name	Calendar	Duration	Project Personnel	Other Resources
8.8	8.9	TRB review of enhanced and expanded web browse structure		3 weeks	Contracting Officer	
	9.0	**Develop and test web services integration to TRIS**	Week 37 - 41	4 weeks		
3	9.1	Define required fields for use by TRIS	5 days		Software Developer, Senior Thesaurus Analyst	
7.14	9.2	Review the expansions to synonyms and impact to TRID search	5 days		Software Developer, Senior Thesaurus Analyst, Thesaurus Analyst	
7.14	9.3	Review the expansions of associative terms and impact to TRID search	5 days		Software Developer, Senior Thesaurus Analyst, Thesaurus Analyst	
	9.4	Define the API/web service that will query TRT for terms and their relationships	10 days		Software Developer	API connectors for convenience copy
7.14	9.5	Review the impact of changes on indexing practices	5 days		Senior Thesaurus Analyst, Thesaurus Analyst	
7.14	9.6	Review the potential impact of changes on full-text and open web search	10 days		Senior Thesaurus Analyst, Thesaurus Analyst	
9.6	9.7	TRB adjustments and/or signoff of new web services			Software Developer, TRB Project Director	
	10	**Review the need for additional training and skills building**	Week 9 - 13	4 weeks		
Final Report	10.1	Competencies for managing development of facet vocabularies	5 days		TRB Project Director, Project Manager, Senior Thesaurus Analyst	
Final Report	10.2	Competencies for working with thesaurus management systems and standards	5 days		TRB Project Director, Project Manager, Senior Thesaurus Analyst	
IR1	10.3	Competencies for SME outreach and networking	10 days		TRB Project Director, Project Manager, Senior Thesaurus Analyst	
Final Report	10.4	Project management competencies for TRT	10 days		TRB Project Director, Project Manager	
10.4	10.5	TRB review of TRT role and competency requirements		3 weeks	Contracting Officer	
	11.0	**Prepare for formal procurement**	Week 46 - 50	4 weeks		
IR3	11.1	Define requirements based on proof of concept and convenience copy	10 days		TRB Project Director and TRB Others	
IR3	11.2	Review the market for TMS tools	10 days		TRB Project Director, Project Manager	
	11.3	Define budget levels	5 days		TRB Project Director	
IR3	11.4	Define licensing constraints	1 day		TRB Project Director, Project Manager	
	11.5	Define procurement strategy - open RFP, single source, bake-off, etc.	10 days		TRB Project Director, Project Manager	

(continued on next page)

Table B-1. (Continued)

Dependencies	Task Number	Task Name	Calendar	Duration	Project Personnel	Other Resources
11.3, 11.5	11.6	Launch procurement with TRB procurement office			TRB Project Director	
	12.0	**Continue to revise and update remaining facets**	Week 37 - 45	8 weeks		
2.7, 7.13	12.1	Prepare evaluation guides drawing from Interim Report 1 results	5 days		Senior Thesaurus Analyst, Thesaurus Analyst	
12.1	12.2	Undertake the reviews	20 days		Senior Thesaurus Analyst, Thesaurus Analyst	
12.2	12.3	Document the assessment results	5 days		Senior Thesaurus Analyst, Thesaurus Analyst	
12.3	12.4	Suggest changes and revisions	10 days		Senior Thesaurus Analyst, Thesaurus Analyst	
12.4	12.5	Make changes as directed	TBD	TBD	Senior Thesaurus Analyst, Thesaurus Analyst	

TRT Assessment Criteria

This appendix identifies the individual criteria used to assess the TRT. The criteria are organized by the five dimensions of the assessment framework described in Figure 1-1 of this report. The criteria are traced to peer-reviewed literature on thesaurus assessment and evaluation. Full citations for the literature are provided in the bibliography for this report.

Assessment Criteria for Content (Dimension A)

The team proposed 35 criteria to assess the content of the thesaurus. During the analysis, 12 additional criteria were added and 2 were combined. Thesaurus content is defined to include thesaurus terms, definitions, classes, relationships, and core functions that support its management and maintenance. We include both descriptive statistics of terms, relationships and essential management and maintenance attributes, as well as comparative statistics across and within current facets. This dimension of the reference model is one of the richest in terms of peer-reviewed references to assessment criteria. Each of these criteria is traceable to the peer-reviewed literature on thesaurus assessment and evaluation (Aitchison, Gilchrist, and Bawden 2000; Clarke 2001; Cleverdon and Keen 1966; Coates 1973; Davis 1968; Foskett 1972; Höpker 1972; Jones 1977; Jones 1967; Kim 1973; LaBarre 2010; Lancaster 1968; Mai 2006; Mandersloot, Douglas, and Spicer 1970; Miller 1997; Milstead 2001; Rada et al. 1988; Schirmer 1967; Travis and Fidel 1982; Tudhope, Alani, and Jones 2006; Willetts 1975).

- Total number of terms in the TRT
- Total number of lead in (non-preferred terms)
- Total number of hierarchical relationships (paradigmatic relationships)
- Total number of associative relationships (syntagmatic relationships)
- Total number of structural terms [New]
- Percentage total terms by facet [New]
- Thesaurus rank of terms by facet [New]
- Inclusion in the TRT or not-approved terms
- Inclusion in the TRT of candidate terms
- Inclusion in the TRT of provisional terms
- How many top descriptors are there in the TRT (second-level terms)?
- Unique identifiers for individual terms
- Terms with definitions
- Terms with traceable source and origin
- Terms with scope notes
- Number of parenthetical qualifiers used in the TRT
- Total number of terms added annually since the TRT creation—all facets
- Ratio of descriptors to identifiers

77

- Terms over three words
- Terms with multiple broader terms
- Errors in assigning hierarchical relationships
- Errors in assigning equivalent relationships
- Errors in assigning associative relationships
- Inclusion of varied word forms as synonyms (spelling variations, singular/plurals, misspellings)
- How are homographs of different meanings treated?
- What percentage of general and specific terms linked through hierarchical relationships?
- What is the level of specificity of terms in the TRT? (Hopker's method of analyzing a class of terms)
- What is the number of coordinate terms (combining multiple concepts to define a class) found in the TRT?
- What is the rate of occurrence of terms no longer used in transportation that are still found in the TRT (e.g., accidents)?
- What is the current rate of productivity of candidate terms that were not accepted for inclusion in the TRT?
- How many top descriptors are there in the TRT (second-level terms) in each facet?
- Number of changes in status of terms over time (when terms change from preferred to non-preferred)
- Retrospective indexing practices for the TRT term changes and deprecations
- Number of terms within each current facet of the TRT
- Number of equivalence terms within each current facet of the TRT
- Number of hierarchical terms within each current facet of the TRT
- Number of associative terms within each current facet of the TRT
- Use of stand-alone adjectives and adverbs [New]
- Total population of RT and RTL terms [New]
- Ratio of RTL to RT relationships [New]
- Number of Level 2 terms [New]
- Depth of levels below Level 2 [New]
- Number of single branching terms and classes [New]
- Number of uncontrolled terms [New]
- Number of lateral relationships (RTLs) [New]
- Number of terms added annually in each facet since the TRT creation

Assessment Criteria for Access and Use (Dimension B)

The team proposed eight criteria to assess the access and use of the thesaurus. During the analysis, 18 new criteria were added. Thesaurus access and use are defined to include instances of direct access to the thesaurus by different stakeholders; embedded use in search, workflow, and navigation applications; and customized use in a range of other applications. The metrics for this dimension focus on the current and extensibility of use of the thesaurus in these different contexts. Because this dimension is represented in later generations, there are fewer peer-reviewed references to evaluation criteria. In most cases, the evaluation criteria are drawn from the open or gray literature and shared by recognized experts (Greenberg 2004; Pinto 2008).

- Number of additional external systems currently using the TRT as an embedded search or indexing application
- Number of additional external systems currently using the TRT terms for search or indexing as manual updates

- Number of additional external systems currently using the TRT for search or indexing as a batch upload
- Number and nature of other uses of the TRT beyond TRID, TRBPI, TRBRIP
- Depth of placement of most frequently used terms (defined through search logs)
- Number of information assets assigned to each level of the hierarchy
- Number of closely related terms assigned to the same document
- Ability to display to the web the full set of terms assigned to a facet
- Use of the TRT to select terms for indexing in TRID or other information sources [New]
- Use of the TRT to select terms for searching in TRID or other information sources [New]
- Use of the TRT to understand the transportation information landscape [New]
- Use of the TRT to discover definitions for transportation terms [New]
- Index term + keyword searches in TRID (Google Analytics) [New]
- Keyword searches in TRID (Google Analytics) [New]
- Ratio of three types of searches [New]
- Rate of use of TRT terms as index terms in TRID [New]
- Frequency of top term use as index terms in TRID [New]
- Frequency of Level 2 term use as index terms in TRID [New]
- Facet definition and frequency of use of second and lower level term use in TRID for each facet [New]
- Display of full thesaurus term record [New]
- Use of the TRT search functionality [New]
- Use of navigation structures [New]
- Ease of navigation of enumerated structure [New]
- Ease of navigation of facet structure [New]
- Effectiveness of display of KWIC [New]
- Effectiveness of display of KWOC [New]

Assessment Criteria for Governance Processes (Dimension C)

The team used 12 criteria to assess the governance processes supporting the thesaurus. Thesaurus governance processes are defined to include engagement by different stakeholders, as well as governance processes (e.g., candidate term discovery, review and evaluation, disposition, and tracing). While governance spans generations, it has generally been tailored to the needs of individual organizations. Less is published in the peer-reviewed literature on this dimension. However, there are authoritative sources from the open and gray literature which can be used as guides.

- How do subject matter experts provide input into term selection and adoption? Into thesaurus evaluations and audits?
- How are "not accepted" terms suggested by subject matter experts represented in the thesaurus?
- How are data managers and stewards involved in discussions of changes to thesaurus terms?
- How are information custodians (librarians/information professionals) who are currently using the thesaurus involved in the governance process?
- How are technical stewards/architects involved in thesaurus governance and decisions? This includes technical stewards and architects outside of TRB Headquarters.
- What are the opportunities for the TRT consumers to suggest terms for inclusion in the TRT?
- How are historical terms managed in the TRT? How are not-approved terms managed in the TRT?

- Are scope notes created for historical terms that are not approved managed in the TRT?
- How often is the TRT evaluated? Who is involved in the evaluation?
- How are stakeholders informed of decisions on terms? Can stakeholders search the TRT to discover the status of their suggested terms?
- How does the governance group know if a new candidate term has been previously considered?
- Are suggestions for terms for structured data access received and considered by the TRT governance group?

Assessment Criteria for Compliance and Standards (Governance Tools) (D)

The team proposed 31 criteria to assess the governance processes supporting the thesaurus. During the analysis, one new criterion was added. Governance principles and tools are defined to include thesaurus user warrant, thesaurus literary warrant, scope and coverage, availability and adherence to best practices and style guidelines, alignment with thesaurus standards, alignment with relevant information management standards, and adherence to interoperability standards. This dimension has both longstanding evaluation criteria dating back to the 1970s and updated criteria based on changing context of operation (Antelman, Lynema, and Pace 2006; Austin 1974; Blocks, Cunliffe, and Tudhope 2006; Blocks et al. 2002; Cleverdon and Keen 1966; Cleverdon 1960; Farradane 1967; Farradane 1970; Farradane, Russell, and Yates-Mercer 1973; Jones 1971; Keith 1970; Klingbiel 1973a; Klingbiel 1973b; Klingbiel 1974; Lancaster 1972; Miller 1997; Montague 1965; Neville 1972; Pickford 1968; Pickford 1971; Rada et al. 1991; Rada and Bicknell 1989; Rada and Martin 1987; Rolling 1970; Rolling 1965; Salton 1972; Saracevic 1971; Stokolova 1977; Tudhope et al. 2006; Turski 1971; Vickery 1971).

- Level of readability of the TRT content (e.g., a mixture of upper and lower case used in terms)
- Consistency of term style guidelines with design guidelines in standard search system architectures
- Consistency of terms with stated TRT style guidelines
- Comparison of the TRT terms to terms in Research in Progress search logs
- Comparison of the TRT terms to terms in TRB Publications Index Search Logs
- Comparison of the TRT terms to terms in TRID Search Logs
- Coverage of terms from Index and Table of Contents in *Highway Safety Manual* in the TRT
- Coverage of terms from Index and Table of Contents in transportation design manuals included in the TRT
- Coverage of TRB research publications author-generated keywords in the TRT
- Inclusion in TRB author-assigned keywords as lead-in terms in the TRT
- The TRT hierarchy levels assigned to author-generated keywords
- Inclusion of research terms used in additional authoritative transportation research resources (e.g., research briefs, dictionaries, glossaries)
- Number of popular (e.g., non-expert) transportation terms are included in the TRT
- Number of Research in Progress index for aviation terms included in the TRT
- Number of terms in Research Results Reports (RIP) included in the TRT
- Number of RIP index terms that are lead-in terms in the TRT
- Rate of match of TRB search log terms to "parts" of a TRT term
- Comparison of coverage to other transportation thesauri
- Comparison of the TRT coverage to semantically generated terms from TRB publications
- Comparison of the TRT facets to topics used by other transportation associations (state, city, regional, local)
- Comparison of the TRT facets to TRB Annual Meeting subject classes

- Comparison of the TRT terms to terms used by transportation research institutes, organizations, and vendors
- Coverage of facets of the field of transportation as defined by peer-reviewed journals
- Coverage of future and "hot topic" terms in the TRT. How well positioned is the TRT to support this research?
- Facet-level treatment of modes of transportation in the TRT
- General search productivity of the TRT terms in the TRB and transportation domains
- Scope of coverage of transportation life cycle stages
- Currency of coverage of transportation methods and tools
- Subject matter expert assessment of depth of coverage of a facet in their area of expertise
- Productivity of the TRT terms in related sources of transportation information (Google, Google Scholar, state DOT search systems)
- Productivity of the TRT coordinate terms in transportation-related sources
- Clearly stated and followed warrant [New]

Assessment Criteria for Thesaurus Architecture (Dimension E)

The team used 48 criteria to assess the architecture of the thesaurus management application. Thesaurus architecture is defined to include underlying data models and four levels of functionality including internal TMS functions, architectural integration with other applications, support for communication protocols, and multilingual capabilities. There are a few peer-reviewed references for these criteria (Ganzmann 1990; Hines and Harris 1971). However, this is an emerging aspect of thesaurus management, so most criteria will come from the field of enterprise and application architecture.

- Does the current thesaurus management application support batch upload capability?
- Does the current thesaurus management application support the definition of extended types of relationships (building upon existing ISO 25964 defined relationships)?
- Does the current thesaurus management application support the maintenance and assignment of multiple term status values (e.g., candidate, approved, not approved, provisional)?
- Does the current thesaurus management application support the creation of multiple categorization schemes?
- Does the current thesaurus management application allow for "ownership" of terms?
- Does the current thesaurus management application support individual terms?
- Does the current thesaurus management application support a Create Date attribute?
- Does the current thesaurus management application support a Relationship Add capability?
- Does the current thesaurus management application support the ability of end users to generate reports?
- Does the current thesaurus management application support the ability of non-developers to parameterize reports?
- Does the current thesaurus management application support the following export formats (e.g., text, xml, rdf/skos, delimited text, custom html, rtf, csv Cartesian, csv + extra delimiters, CSV in columns)?
- Does the current thesaurus management application support on-demand generation and updating of thesaurus statistics?
- Does the current thesaurus management application support output of search results and reports to both screen displays and print formats?
- Describe all of the types of standard reports the current application can generate
- Does the current thesaurus management application support Term Update/Edit by thesaurus managers?

- In the current thesaurus management application, are there any constraints to the levels of hierarchical relationships supported in the current application?
- Does the current thesaurus management application support a Created By attribute?
- Does the current thesaurus management application support a network/consortium version for use by multiple thesaurus stewards?
- Does the current thesaurus management application support the definition of different kinds of notes fields—scope notes, editors notes, personal notes, and definitions?
- Does the current thesaurus management application support archiving of thesaurus terms?
- Does the current thesaurus management application support dynamic advanced search or in context search?
- Can the current TRT application manage and assign multiple levels of permissions and privileges?
- Does the current thesaurus management application support a global rebalance function?
- Does the current application have the ability to support multiple definitions for a single term record?
- Does the current application have the capability to identify relationship errors or conflict upon batch or manual entry of new terms (e.g., multiple entries of same term, multiple relationships, incomplete relations, relationing errors, and consistency checks)?
- Does the current thesaurus management application support the ability for a term to appear in more than one facet (e.g., category)?
- Does the current thesaurus management application support the existence of distinct records for individual terms?
- Can the current thesaurus management software display both fully spelled out labels and label abbreviations?
- Is there a Global Replace capability in the TRT today?
- Are there any limitations or constraints applied to the length of thesaurus terms in the current thesaurus management application?
- Level 1—reports capabilities—ability to define label formats in reports
- Does the current thesaurus management application support the selection of fields to be included in export reports?
- Does the current thesaurus management application support the ability to generate alphabetical listings of terms?
- Does the current thesaurus management application support the ability to generate hierarchical listing?
- Does the current thesaurus management application support the ability to generate a KWIC report?
- Does the current thesaurus management application support the ability to generate a KWOC report?
- Does the current thesaurus management application support the ability to select terms by types of relationships?
- Does the current thesaurus management application support simple search capabilities?
- Does the current thesaurus management application support advanced search capabilities?
- Does the current thesaurus management application impose any character limitations on the size or occurrence of scope notes fields?
- Does the current thesaurus management application support the direct extraction and use of unique term identifiers for use in other systems?
- Does the current thesaurus management application extract full term names for use in automated indexing and automated categorization applications or is translation required?
- Does the current thesaurus management application provide a web-based SDK for publishing a thesaurus on the web?

- Can the TRT in its current form be integrated into other information management architectures today and in the future?
- What level of effort or translation is required to use the thesaurus content in other non-TRB applications (e.g., amount of transformation and manipulation required)?
- Does the current thesaurus management application support the exchange of updated terms with other users through standard communication protocols (XML, RDF/SKOS)?
- Can the case of terms be prespecified for output? Does the current thesaurus management application support the creation of other language versions of individual terms? (i.e., and full functionality of thesaurus for other language versions)?
- Is there an upper level constraint for the number or unique identification of terms? Is there any upper limit to the unique ID sequence?

Interview Guide and Data Collection Forms

1. Background and Purpose of the Interviews

The purpose of this guide is to describe the process the research team proposes to conduct interviews from stakeholders, as outlined in the project proposal. The guide serves as a critical reference source for the research team in completing the schedule of interviews and in conducting the interviews in a consistent and comprehensive manner. Working from a guide that has been reviewed, edited, and supported by both the research and the research team will ensure that we consistently collect, document, and archive interviews and any information that is referenced or shared during the interviews. The interview documentation will serve as an important knowledge base for the research team and the research team in conducting its assessment of the TRT and in formulating findings and conclusions.
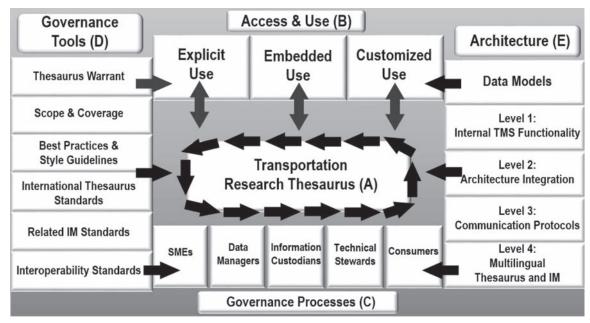
2. Context for Understanding Current Challenges and Opportunities for the Transportation Research Thesaurus

To gain a better understanding of the current environment, challenges, and opportunities, the research team will conduct interviews with five stakeholder groups, including: (1) department of transportation staff who have critical roles pertinent to the TRT; (2) transportation subject matter experts who are the primary intended users of the TRT; (3) the TRT administrative staff, TRB project panel, and the TRT Subcommittee; (4) the TRT System Architects and Administrators; and (5) information science professionals in state transportation department libraries, academic transportation libraries, and transportation knowledge networks.

The research team has prepared a preliminary set of questions to guide the interviews with the roles described above. For each interview, the research team will record information in a distinct document. Each document will be filed and archived as part of the research effort. In addition, reference materials shared by the interviewee will be logged and filed in the research archive. Interview responses will be reviewed by the team as they are recorded, to ensure that the entire team can share insights and fold the knowledge gained into their tasks.

The information gathered through interviews is intended to supplement the quantitative assessment data generated by the research team working directly with the TRT. The five stakeholder groups were identified to ensure that we have comprehensive supplementary information for all parts of the assessment framework (Figure D-1).

The goal of the interviews is to collect and synthesize subject matter expert opinions and experiences related to the TRT and to information seeking and discovery habits. The information collected through the interviews will contribute to the body of knowledge that serves as a foundation for our findings and conclusions. To ensure that we include this knowledge in the foundation,

IM = information management

Figure D-1.   Assessment framework.

the research team will conduct a thorough interview process. The interviews will be conducted simultaneously with the literature review. As knowledge is gained through the interview process, it will be fed back to those team members who are supporting the formal and gray literature review.

3. Proposed Interview Schedule and Strategy

Who We Plan to Interview

Table D-1 describes the stakeholder groups referenced in the project proposal. Table D-1 also highlights how stakeholders' feedback will be factored into the assessment framework.

Stage 1: Department of Transportation Critical Roles and Responsibilities

This stage focuses on gathering information from key stakeholders within the department of transportation. The purpose of this stage is to gain an understanding of the purpose, expected

Table D-1.   Stakeholder feedback alignment with assessment framework.

Stakeholder Group	Framework A	Framework B	Framework C	Framework D	Framework E
DOT Key Roles	X	X	X	X	X
Transportation SMEs	X	X	X	X	
TRT Administrators	X		X	X	X
TRB Project Panel Member	X	X	X	X	X
TRT Subcommittee Members	X	X	X	X	X
TRT System Architects & Developers		X			X
SDOT Librarians	X	X	X	X	
Academic Librarians	X	X	X	X	X
Transportation Knowledge Network Members	X	X	X	X	X
Publishers	X			X	

value, and intended use of the TRT. The information gathered in this stage of the project pertains to Framework Dimensions A, B, C, D, and E. The strategy for this stage is described in Section 4.

Stage 2: Subject Matter Experts on Advisory Panel and Across State Transportation Departments

This stage focuses on gathering information from subject matter experts in the field of transportation. The purpose of this stage is to gain an understanding of the value and use of the TRT by domain experts. The information gathered in this stage of the project pertains to Framework Dimensions A, B, C, and D. The strategy for this stage is described in Section 5.

Stage 3: The TRT Administrative Staff, TRB Panel Members, and TRT Subcommittee Members

This stage focuses on gathering information from the core staff whose roles and responsibilities involve the creation, management, and maintenance of the TRT. The information gathered in this stage of the project pertains to Framework Dimensions A, B, C, D, and E. The strategy for this stage is described in Section 6.

Stage 4: System Administrators and Implementers Who Have Experience Working with the TRT

This stage focuses on gathering information from those individuals whose responsibilities involve the TRT system level architecture, design, and functionality. The information gathered in this stage of the project pertains to Framework Dimensions B and E. The strategy for this stage is described in Section 7.

Stage 5: Transportation Librarians, Knowledge Networks, and Publishers

This stage of the interview focuses on the state and network level value of the TRT. In this stage, we gather information from state transportation department libraries, from academic libraries that currently or have historically used the TRT to support information management, and transportation knowledge networks for which the TRT may have current or potential value. The information gathered in this stage of the project pertains to Framework Dimensions A, B, C, D, and E. The strategy for this stage is described in Section 8.

At the state transportation department level, we propose to conduct in-person interviews with a minimum of four states, including Michigan, Ohio, Virginia, and Washington. In addition, we propose an additional 10 virtual interviews with states known to have a rich transportation research base, including:

- California
- Colorado
- Georgia
- Massachusetts
- Minnesota
- Missouri
- Louisiana
- Arizona
- Mississippi
- Texas

In addition, we propose to contact university, college, and research institute libraries that may use the TRT at the present time or have a future intent to use the TRT for institutional purposes. This set of interviews will draw from the membership of the four regional components of the Transportation Knowledge Network. Interviews for these stakeholders will draw upon Framework Dimensions A, B, C, D, and E.

Table D-2.   Proposed interview schedule.

Interview Stage	Start Date	Completion Date
Stage 1	Oct. 1	Oct. 31
Stage 2	Oct. 31	Nov. 15
Stage 3	Oct. 1	Nov. 15
Stage 4	Nov. 1	Nov. 30
Stage 5	Nov. 1	Nov. 30

In addition to information science professionals, the research team will also consult with publishers of major transportation research journals and proceedings to gain an understanding of the transportation vocabulary used to support those sources of information. Interviews for these stakeholders will draw upon Framework Dimensions A and D.

Interview Schedule and Formats

The interview schedule strives to make the most efficient use of the research team's time and resources while maximizing our contact time with subject matter experts. A high-level schedule for the completion of interviews is provided (Table D-2).

Interview Design and Format

Each interview will be structured around a consistent set of questions. The research team has compiled a draft set of questions for each of the four roles we intend to interview. Those questions are included in this Guide. We invite the research team to review and provide comments on the interview questions. Pretests of the questions suggest that the interviews can be conducted in an hour or less. The interviews are designed to be conducted by one team member on site and multiple team members when conducted virtually.

How We Plan to Capture Interview Content

Interview content will be recorded as the conversation is taking place, with the full permission of the interviewees. It will be transcribed from recordings and shared among the team. Recordings will be disposed of immediately after transcription. Information collection, synthesis, and analysis will be facilitated by the use of a structured interview guide. All data collected will be treated like research data, with full sourcing and tracing.

4. Interview Objectives and Questions for Key Department of Transportation Individuals

The goal of this round of interviews is to gain a broad understanding of stakeholders' understanding of the goals and objectives of the TRT. Questions will also address current levels of satisfaction with, and future use of, the TRT in current and other transportation information contexts. Table D-3 describes the questions for this round of interviews.

5. Interview Objectives and Questions for Subject Matter Experts

Subject matter expert interviews will be iterative. We intend to follow a five-step process that will make efficient use of their time and gather targeted information about their use of the TRT, its scope and coverage of their subject domains, and their vision for future use of the TRT.

Table D-3. Interview questions for Stage 1.

Framework Dimension	Question Text	Interviewee Response	Referenced Materials
A	Please tell us about the TRT's current strengths.		
A	Please tell us about the TRT's current weaknesses.		
A	How would you describe the content and structure of the TRT?		
A	Please tell us what you know about the history and evolution of the TRT. Have there been other assessments or recommendations related to TRT that the research team should be aware of?		
A	Please tell us about your experience working with other thesauri.		
B	Please tell us what you know about current uses of TRT beyond the TRB research databases.		
B	Please share with us any ideas you have for the future or potential uses of TRT within the field of transportation.		
A	Please tell us about your experiences working with semantic search tools.		
B	Please tell us how you think TRT supports access to transportation research in general.		
B	How effective is TRT for your searching of TRID? of TRIP?		
B	Are you aware of any projects to use or integrate TRT into other information management or search systems?		
B	If you have had any experiences integrating TRT into other applications, please tell us how you accomplished that.		
B	Do you use TRT in your local information systems? Please tell us how you are using it.		
B	Do you use TRT terms to search other applications? If so, please explain how you select and search them.		
B	Have you used TRT to create any other discovery tools?		
C	Please tell us what you see as the strengths and the weaknesses of the current TRT governance process.		
C	Do you think your current access to TRT enables you to participate fully in governance? If not, what access or capabilities would you like to see TRT afford?		
C	Please tell us about your experience recommending changes to TRT.		
D	Please tell us about the scope and coverage of TRT relative to your areas of interest.		
D	Please tell us about the authoritative glossaries, dictionaries, and encyclopedias in your field. Also please tell us about the most important peer-reviewed journals, conferences, and professional associations which promote or through which research passes in your fields.		
D	How would you rate TRT in terms of a representation of the full field of transportation? Are there gaps in coverage or varying levels of treatment that you think should be addressed? Are there other areas which have received too much emphasis?		
D	Please tell us how you think TRT aligns with the coverage and focus of TRB.		
E	What are your expectations for the future use of the TRT?		
E	What routine reports or descriptive information would you like to receive about the status and content of the thesaurus?		

The five steps are outlined below. Questions and their alignment to the assessment framework are also provided.

This set of interviews is more elaborate than the others because the subject matter experts constitute a formal advisory group. The advisory group members have committed to contributing up to 10 hours over the course of the project.

Step 1. First Questions to Subject Matter Experts—30 Minutes

The first questions to the subject matter experts will come in the form of a simple Qualtrics form. The form will ask them to electronically list the major sources of information in their areas of expertise, including dictionaries, glossaries, encyclopedias, peer-reviewed journals, professional associations, and conferences. This should take about 30 minutes of their time, and the results will feed directly into the definitions and the TRT scope, coverage, and warrant assessment. The results may also contribute to the project work on definitions.

Step 2. Virtual Interview on General Discovery Habits—1 Hour

Step 2 takes the form of a virtual/in-person interview (depending on where the individual lives). This will take about 1 hour. These interviews will take the form of conversations around the Interview Guide questions. We will send the questions to them ahead of time.

Step 3. Review of the TRT Term Clusters for Subject Areas—1 Hour

At this point, we will have done a high-level comparison of the TRT coverage in subject areas with the authoritative sources in the field; we will have an idea of how subject matter experts discover information and how they assess the TRT. We can generate the TRT term clusters for their areas (drawing from the working copy of the TRT) and ask them to do a quick review of what is there that should not be, what is missing, and what should receive greater/lesser emphasis. These interactions should take place about the middle of November and will feed back into the facet level assessments of the TRT. We anticipate this exercise will take about an hour per subject area—some subject matter experts may be covering more than one subject area.

Step 4. Short Query on Current and Future Uses of the TRT—30 Minutes

Again, this is not a formal public survey but just a way to collect information from the subject matter experts in an unobtrusive way. We have a few questions in the Interview Guide that focus on current and future use of the TRT in the subject matter experts' discovery and business environments. These questions ask for feedback on pros/cons and visions, which will help us to speak to the current TRT's ability to support their future visions.

Step 5. Review of the Task 2 TRT Assessment Report—1 Hour

Just to make sure that we've represented all of the subject matter experts' input accurately, we intend to provide preview copies of the Task 2 assessment results to the subject matter experts for review. It also provides a check across subject matter experts on the feedback we received.

Questions related to Framework Dimension A are not included in this listing. They will be covered in the source and term cluster reviews described above. Table D-4 describes the questions for this round of interviews.

Table D-4.   Interview questions for Stage 2.

Framework Dimension	Interview Question	Interviewee Response	Documents Referenced
B	When you assign keywords to your publications, do you consult and use TRT terms?		
B	Do you suggest new terms or changes to terms in TRT?		
B	Do you use TRT? If yes, how and when?		
B	If you do not use TRT, please tell us why not.		
B	Please share with us your thoughts on the currency of terms in TRT. If out-of-date terms are included, how would you suggest they be changed?		
B	When you search TRID, please tell us how you search. By keywords, by title, etc.		
B	When you search TRID, how often do you find what you're looking for?		
B	Are you more successful searching one of the other of these sources—TRID, TRIP?		
B	Please tell us about the TRT terms that you most often choose to search.		
B	How do you search for information in your domain?		
B	Where do you search?		
B	Please tell us about the kind of information you work with in an average week—structured data, unstructured data, graphics, images….		
B	Please tell us about the topical focus of the information you work with in an average week.		
C	Have you suggested terms for inclusion in TRT that were not accepted? Were those terms added as entry terms?		
C	If you have suggested a term for inclusion in TRT, have you been informed of the decision to include or not to include?		
C	Is there a way for you to discover the terms that have been suggested for inclusion in TRT?		
C	Do you know how to submit a term for inclusion in TRT?		
C	Are your TRB author suggested TRB terms included in TRT?		
D	What is the most trusted resource for research results in your field?		
D	What are the standard glossaries and encyclopedias in your field?		
D	Referring to the cluster map, what do you think is missing in coverage?		
D	Referring to the cluster map, what is included that should not be included?		
D	Referring to the cluster map, what has received too much emphasis?		
D	Please tell us your thoughts on the currency of coverage of transportation methods and tools in TRT.		
D	Please tell us how you would characterize the coverage of your domain in TRT.		
E	Do you see additional uses for the TRT in your working environment? Please describe.		
E	Please tell us about the environments in which you see future use of TRT.		
E	Is there a need for a multilingual version of TRT in your subject area?		
E	Have you used TRT—in whole or in part—in other applications? Please describe.		

6. Interview Questions and Objectives for TRB Subcommittee, Project Panel, and Thesaurus Managers

The goal of this round of interviews is to gain a deep understanding of the functionality, structure, and governance of the TRT from those who are most actively involved with its construction, maintenance, and use. These interviews will be conducted after we've had an opportunity to work with the TRT in depth. Table D-5 describes the questions for this round of interviews.

Table D-5.   Interview questions for Stage 3.

Framework Dimension	Question Text	Interviewee Response	Documents Referenced
A	Please tell us what you think are TRT's areas of strengths.		
A	Please tell us what you think are TRT's areas of weakness.		
A	Please share with us your knowledge of TRT's history and evolution.		
A	What do you know about TRT's original warrant? Please share with us your thoughts on whether the warrant established in the 1990s is relevant today. If not, what changes would you suggest?		
A	Please tell us what you know of TRT's style guidelines? Do you think these style guidelines are relevant today? If not, what changes would you suggest?		
A	In the current TRT, there is a category of relationships called Related Term (Hierarchical). Would you please tell us what this means to you? Please tell us how often you use these relationships to expand your search.		
A	Please share with us your thoughts on the inclusion of definitions in TRT. How do you choose sources for definitions? How are they verified?		
A	What changes or augmentations to TRT would you like to see in the future?		
A	From your experience, how often are TRT terms changed?		
B	Please describe the research process you follow when you are deciding to include/not to include a new term.		
B	What is the most valuable aspect of TRT to your work?		
B	Please share with us any earlier projects to either assess or to create other versions of the TRT that you think are important for the research team to know of.		
B	Please tell us how well you think TRT supports access to TRB research.		
B	Please share with us your thoughts on the effectiveness of TRT as a search tool in TRID? In TRIP?		
B	Please tell us who does the indexing for these important transportation systems. Are there any indexing guidelines?		
B	Please tell us whether you are aware of any previous or ongoing efforts to embed TRT in another information system.		
B	Please tell us whether you are aware of any projects that leverage a subset of the TRT, perhaps by facet. Please describe the process behind these projects.		
B	Please tell us how you use the Web version of TRT. What works well? What would you like to see improved and how?		

(continued on next page)

Table D-5.   (Continued).

Framework Dimension	Question Text	Interviewee Response	Documents Referenced
B	If you use TRT for indexing, please tell us how you work with TRT and select terms. Do you use the TRID interface, the TRT web interface, a local version?		
B	Of the various access points for TRT (via the website), which do you use most often? Least often? Explain your choices.		
B	Please tell us how you would use TRT in your library if you had a full exported copy.		
C	From your experience, how often do subject matter experts suggest terms for inclusion in TRT? How often are suggested terms included in the thesaurus?		
C	What TRT tools help you with your governance responsibilities? Are there other tools you wish you had?		
C	In your experience, when was the last time a formal assessment of TRT was undertaken? Either by TRB, an individual state or a library?		
C	What sources or processes do you use to identify new terms for inclusion in TRT? Search logs?		
C	If you were to conduct an annual assessment of TRT, what would be the most important aspect?		
C	In your experience, please tell us who is involved in the TRT governance process. The core group, stakeholders and custodians?		
C	In your experience, how are historical terms or newly non-preferred terms handled in TRT?		
C	Please tell us how you check to see whether a suggested term has been considered in the past?		
C	Have you had experience working with the TRT XML export?		
D	Please tell us what you know about the structure of TRT and its alignment with ANSI/NISO standards.		
D	Please tell us about the sources you most frequently search to find transportation information. Please tell us how you search those sources. Do you ever use TRT terms to search those sources?		
D	Based on your experience, please tell us what you think of the balance in coverage of aspects of transportation. In particular, we would like to hear your thoughts on the coverage of modes of transportation.		
D	From the perspective of your library users or those SMEs you support, how well does TRT currently reflect their needs and interests?		
D	How does TRT's current scope and coverage compare to the search logs for sources in your libraries?		
D	Please tell us whether you are familiar with any national or international thesaurus standards. If so, please tell us where you think TRT aligns with thesaurus standards, and where it diverges. Please tell us whether you think the variations influence the use of TRT.		
E	Please tell us what other thesaurus management applications you have worked with.		
E	Please tell us what capabilities available in other applications you would like to have available in TRT.		
E	What reports from TRT would you like to have to support your work as a librarian or indexer?		
E	Please tell us your thoughts on the hierarchical structure of TRT.		
E	When was the last time a new facet was added to TRT? How was it integrated into the hierarchical structure?		

7. Interview Questions and Objectives for the TRT System Administrators and Implementers

This stage represents a more limited and targeted group of stakeholders. The goal of this stage of the interviews is to ensure that we have a full and accurate understanding of the system architecture, design, and functionality. These interviews will contribute to a deeper understanding of the constraints and opportunities that result from the TRT's embedded status in TRID. Table D-6 describes the questions for this round of interviews.

8. Interview Questions and Objectives for Transportation Networks and State Department of Transportation Librarians

This round of interviews is more extensive and focuses on intermediaries and information professionals who are working with or would like to work with the TRT. Table D-7 describes the questions for this round of interviews.

Table D-6.   Interview questions for Stage 4.

Framework Dimension	Question Text	Assessment Mapping
A	Based on your experience working with TRT, please describe the database structure. Does each term have a full and complete record in the structure?	
A	Is each relationship explicitly called out and understandable from the view of a single term record?	
B	Please explain to us how TRT is integrated into TRID. What would be the impact of changing the current TRT application?	
B	Please tell us whether TRT has a batch uploading capability. In your view, could such a batch upload capability be developed? What would you estimate to be the LOE for that work?	
B	Please describe the functionality that is available to thesaurus managers that is not available to TRT consumers.	
C	Please describe the report capabilities designed into TRT.	
C	Please tell us how new terms are presented to you. What information is provided to include them in TRT? Please describe how you enter these new terms into TRT.	
C	Please tell us how you make changes to terms in TRT. Please describe what happens to the terms whose status is being changed from preferred to non-preferred. Please tell us what happens to the relationships from that term to other terms.	
C	Please tell us whether you have ever worked with a consumer of TRT to integrate or design TRT into another non-TRB application.	
C	If you could change the design of TRT, what would be your primary area of focus? What do you see as the greatest constraint of TRT? What do you see as the greatest asset of TRT?	
E	In your view, what other thesaurus applications could be used as substitutes for the current TRT application?	
E	Under what circumstances could the current TRT application be viewed as a consumer of a new TRT application? Particularly where such a new application could support thesaurus use and consumption by states and other transportation information consumers?	
E	Dimension E questions related to the application's architecture	
E	Please tell us about the design of the search in TRID and the integration of TRT into that design. What would happen if either TRID or TRT were redesigned?	

Table D-7.   Interview questions for Stage 5.

Framework Dimension	Question Text	Assessment Mapping
A	Please tell us what you think are TRT's areas of strength.	
A	Please tell us what you think are TRT's areas of weakness.	
A	Please share with us your knowledge of TRT's history and evolution.	
A	What do you know about TRT's original warrant? Please share with us your thoughts on whether the warrant established in the 1990s is relevant today. If not, what changes would you suggest?	
A	Please tell us what you know of TRT's style guidelines. Do you think these style guidelines are relevant today? If not, what changes would you suggest?	
A	In the current TRT, there is a category of relationships called Related Term (Hierarchical). Would you please tell us what this means to you? Please tell us how often you use these relationships to expand your search.	
A	Please share with us your thoughts on the currency of terms in the TRT. What efforts are routinely taken to make sure that TRT reflects the current work in the field of transportation?	
A	Please share with us your thoughts on the inclusion of definitions in TRT. What is the quality? How often do you consult these definitions?	
A	What is the most valuable aspect of TRT to your work?	
A	What changes or augmentations to TRT would you like to see in the future?	
A	From your experience, how often are TRT terms changed?	
B	Please share with us any earlier projects to either assess or to create other versions of the TRT that you think are important for the research team to know of.	
B	Please tell us how well you think TRT supports access to TRB research.	
B	Please share with us your thoughts on the effectiveness of TRT as a search tool in TRID? In TRIP?	
B	Please tell us whether you are using TRT terms to describe or manage transportation information in other systems.	
B	Please tell us whether you are using TRT search terms to search for transportation research in other applications. Please tell us how you select them and whether you have created direct links across systems.	
B	Please tell us whether you are aware of any previous or ongoing efforts to embed TRT in another information system.	
B	Please tell us whether you are aware of any projects that leverage a subset of the TRT, perhaps by facet. Please describe the process behind these projects.	
B	Please tell us how you use the web version of TRT. What works well? What would you like to see improved and how?	
B	If you use TRT for indexing, please tell us how you work with TRT and select terms. Do you use the TRID interface, the TRT web interface, a local version?	
B	Of the various access points for TRT (via the website), which do you use most often? Least often? Explain your choices.	
C	From your experience, how often do subject matter experts suggest terms for inclusion in TRT? How often are suggested terms included in the thesaurus?	
C	What TRT tools help you with your governance responsibilities? Are there other tools you wish you had?	
C	In your experience, when was the last time a formal assessment of TRT was undertaken? Either by TRB, an individual state or a library?	
C	What sources or processes do you use to identify new terms for inclusion in TRT? Search logs?	
C	If you were to conduct an annual assessment of TRT, what would be the most important aspect?	
C	In your experience, please tell us who is involved in the TRT governance process? The core group, stakeholders, and custodians?	
C	In your experience, how are historical terms or newly non-preferred terms handled in TRT?	

Table D-7.   (Continued).

Framework Dimension	Question Text	Assessment Mapping
C	Please tell us how you check to see whether a suggested term has been considered in the past.	
C	Please tell us how you would use TRT in your library if you had a full exported copy.	
C	What modifications would you make to TRT if you could use it locally?	
C	What other uses—beyond libraries—would be made of TRT if there were an exportable copy?	
C	Have you had experience working with the TRT XML export?	
C	If you are working with a local version of TRT, please tell us what applications you are using to support and maintain it.	
D	Please tell us what you know about the structure of TRT and its alignment with ANSI/NISO standards.	
D	Please review the following examples and tell us how these terms are related (give about 5/6 examples).	
D	Please tell us about the sources you most frequently search to find transportation information. Please tell us how you search those sources. Do you ever use TRT terms to search those sources?	
D	Based on your experience, please tell us what you think of the balance in coverage of aspects of transportation. In particular, we would like to hear your thoughts on the coverage of modes of transportation.	
D	From the perspective of your library users or those SMEs you support, how well does TRT currently reflect their needs and interests?	
D	How does TRT's current scope and coverage compare to the search logs for sources in your libraries?	
D	Please tell us whether you are familiar with any national or international thesaurus standards. If so, please tell us where you think TRT aligns with thesaurus standards and where it diverges. Please tell us whether you think the variations influence the use of TRT.	
D	Please tell us whether you review author keywords in TRB papers for inclusion in TRT.	
E	Please tell us what other thesaurus management applications you have worked with.	
E	Please tell us what capabilities available in other applications you would like to have available in TRT.	
E	What reports from TRT would you like to have to support your work as a librarian or indexer?	
E	Please tell us your thoughts on the hierarchical structure of TRT.	
E	When was the last time a new facet was added to TRT? How was it integrated into the hierarchical structure?	

9. Subject Matter Expert Information Sources and Use Data Collection Form

NCHRP Project 20-109 Subject Matter Expert Information Sources

Q1. Please tell us about your field of transportation. Please select all that apply. If a field of transportation is missing from the list, please feel free to write it in to a text field.

❏ Highway Transportation
❏ Air Transportation
❏ Maritime Transportation
❏ Rail Transportation
❏ Freight Transportation
❏ Intelligent Transportation Systems
❏ Multimodal Transportation
❏ Transportation Policy

- ❏ Transportation Planning
- ❏ Transportation Maintenance
- ❏ Transportation Operations
- ❏ Safety
- ❏ Transportation Security
- ❏ Transportation Asset Management
- ❏ Transportation Engineering
- ❏ Transportation Materials Science
- ❏ Other (Please specify) _____
- ❏ Other (Please specify) _____
- ❏ Other (Please specify) _____
- ❏ Other (Please specify) _____

Q2. Please refer us to the most authoritative dictionaries and glossaries in your field of transportation. Please list as many sources as you wish.

Q3. Please refer us to any encyclopedias which are important sources of definitions or which provide important descriptions of your field of transportation. Please list as many sources as you wish.

Q4. Please tell us the top five peer-reviewed journals in your field of transportation.

- ❏ Click to write Choice 1 _____
- ❏ Click to write Choice 2 _____
- ❏ Click to write Choice 3 _____
- ❏ Click to write Choice 4 _____
- ❏ Click to write Choice 5 _____

Q5. Please tell us the top five trade publications in your field of transportation

- ○ Click to write Choice 1 _____
- ○ Click to write Choice 2 _____
- ○ Click to write Choice 3 _____
- ○ Click to write Choice 4 _____
- ○ Click to write Choice 5 _____

Q6. Please list the professional associations that are important to your field of transportation.

Q7. Please list the professional conferences that are important to your field of transportation.

Q8. Please list any trade shows that are important to your field of transportation.

10. State DOT and Academic Librarian Data Collection Form

Section 1. Respondent Demographics

Q1. Please tell us which of the following describes your primary work sector? (More than 50% of your work hours)

❍ U.S. Department of Transportation
❍ Other U.S. Federal Government
❍ Non-U.S. Government
❍ State Highway/Transportation Department
❍ Other State Government
❍ Local or Regional Government Agency (City, County, MPO)
❍ Aviation
❍ Port Authority
❍ Transit Agency
❍ Railroad
❍ Academia
❍ Research Institution
❍ Industry
❍ Contractor/Consultant
❍ Service/Trade Organization or Association
❍ Other (Please Specify) _____

Q2. Please tell us which transportation domains are most relevant to your work? (Please select all that apply)

❏ Road and Highway Transportation
❏ Air Transportation
❏ Maritime and Water Transportation
❏ Rail Transportation
❏ Intelligent Transportation Systems
❏ Freight Transportation
❏ Pipeline Transportation
❏ Transportation System Planning and Policy
❏ Transportation System Design and Construction
❏ Transportation System Asset Management and Maintenance
❏ Transportation Systems Operations Research
❏ Transportation Safety and Security
❏ Transportation Materials Science
❏ Intermodal Transportation
❏ Transportation Economics and Mobility
❏ Other (Please Specify) _____

Q3. Please tell us which of the following research and information services your organization provides (Please select all that apply)

❏ Access to local collections of transportation related information, publications, and data
❏ Access to local collections of scholarly information covering all disciplines
❏ Access to commercial databases pertaining to transportation
❏ Access to general scholarly commercial databases
❏ Access to TRB publications, databases, and products
❏ Search services for staff

- ❏ Current awareness services for staff
- ❏ Reference services for staff in your organization
- ❏ Reference services for the public in your state or locale
- ❏ Cataloging and indexing services for your organization's library
- ❏ Cataloging and indexing services for other information sources
- ❏ Information services and architecture support for projects at your organization
- ❏ Other (Please specify)

Q4. Please tell us about the types of information your organization manages. (Please select all that apply)

- ❏ Print Books
- ❏ Digital Monographs and Books
- ❏ Print Peer-Reviewed Journals
- ❏ Electronic Journals and Articles
- ❏ Conference Proceedings
- ❏ Physical Audio
- ❏ Digital Audio
- ❏ Physical Images
- ❏ Digital Images
- ❏ Grey Literature
- ❏ Newspapers
- ❏ Maps
- ❏ General Government Documents
- ❏ Legal Sources (e.g., Statutes, Legal Codes, Legislation)
- ❏ Research Data Sources
- ❏ Other (Please specify)

Section 2. Transportation Research Thesaurus Awareness and Use

Q5. Please describe your awareness of and familiarity with the Transportation Research Thesaurus (TRT). (Please select all that apply)

- ❏ I am aware of the TRT
- ❏ I have been introduced to the TRT through a training or orientation program
- ❏ I have been trained to use the TRT for searching
- ❏ I have been trained to use the TRT for cataloging and indexing
- ❏ I was not aware of the TRT
- ❏ Other (Please explain)

Q6. Please tell us how frequently you use the Transportation Research Thesaurus (TRT) (Matrix question—options by frequency rates/columns)

- ❏ I use the TRT for searching TRB research and publications
- ❏ I use the TRT for cataloging and indexing
- ❏ I use the TRT for understanding an area of transportation
- ❏ I use the TRT to look up definitions of transportation terms
- ❏ Other uses (Please specify)

Q7. Please tell us about how your organization has used the TRT to support special or one-time projects (Please select all that apply)

- ❏ To select keywords for research publications
- ❏ To construct indexes for books
- ❏ As subject descriptors for the local library catalog

❏   To support machine assisted indexing and/or classification
❏   As a reference source embedded in workflow or decisions systems
❏   As a search tool in a local search system
❏   Other (Please describe)

Q8. When you use the TRT, please tell us which of the following you use. Please also tell us about how frequently you use these options.

❏   Browse the hierarchical listing of terms for individual facets
❏   Browse the enumerated (e.g., alphabetical coding) list of terms
❏   Search for the term I am looking for
❏   Use the Keyword In Context (KWIC)
❏   Use the Keyword Out of Context (KWOC)
❏   Other (Please describe)
❏   I do not use the TRT

Q9. If you selected any of the options to Question 7, please describe the process you followed to extract terms and relationships from the TRT.

Section 3. Scope, Coverage, and Currency of the Transportation Research Thesaurus

Q10. Please tell us what you think are the TRT's strengths.

Q11. Please tell us what you think are the TRT's weaknesses.

Q12. Warrant is defined as the rationale for deciding to include terms in a thesaurus. Please tell us which "warrant" you think best describes how decisions are made for the TRT.

❏   General Literary Warrant (e.g., the term must appear in the published peer-reviewed literature)
❏   Extended Literary Warrant (e.g., the term must have appeared in the published peer-reviewed literature for several years)
❏   General User Warrant (e.g., the term is in use by professionals and practitioners in the field)
❏   Extended User Warrant (e.g., the term must be in common parlance within the community for several years)
❏   Don't Know
❏   Other (Please explain)

Q13. Please tell us what changes or enhancements would contribute to your use of the TRT.

<div style="border:1px solid black; height:150px;"></div>

Section 4. Participation in Governance and Development of Transportation Research Thesaurus

Q14. Please tell us in which of the following ways you have contributed to or participated in the TRT development. Please select all that apply.

- ❏ Submitted terms for inclusion
- ❏ Recommended other transportation vocabularies for inclusion
- ❏ Suggested changes or edits to current terms
- ❏ Suggested changes or edits to current relationships
- ❏ Suggested changes or edits to current facets
- ❏ Other (Please explain)

Q15. Please tell us in which of the following ways you have participated in the governance of the TRT. Please select all that apply.

- ❏ Researched terms which are under consideration
- ❏ Participated in the review of candidate terms
- ❏ Participated in the decision making process
- ❏ Contributed to the review of the TRT facets and/or terms
- ❏ Other (Please explain)

Section 5. Thank You and Follow-Up

Q16. Please use the space below to share any other thoughts on the TRT that you think will be helpful to the community in moving forward.

<div style="border:1px solid black; height:150px;"></div>

Q17. Would you be willing to share your thoughts with a member of the research team in an interview? If so, please advise what format would be most convenient for you.

- ❏ Phone or Skype Conversation
- ❏ Online Virtual Meeting
- ❏ In Person Interview
- ❏ Other (Please Specify)

Bibliography

Aitchison, J., Gilchrist, A., and Bawden, D. 2000. *Thesaurus Construction and Use: A Practical Manual*. 4th edition. Routledge.

Antelman, K., Lynema, E., and Pace, A. K. 2006. "Toward a Twenty-first Century Library Catalog." *Information Technology and Libraries*, Vol. 25, No. 3, p. 128.

Austin, D. 1974. *PRECIS: A Manual of Concept Analysis and Subject Indexing*. Council of the British National Bibliography. London, UK.

Batty, D. 2001. *NCHRP Report 450: Transportation Research Thesaurus and User's Guide*. Transportation Research Board, National Research Council, Washington, D.C.

Blocks, D., Binding, C., Cunliffe, D., and Tudhope, D. 2002. "Qualitative Evaluation of Thesaurus-based Retrieval." In *International Conference on Theory and Practice of Digital Libraries*, September 16–18, 2002, Rome, Italy, pp. 346–361.

Blocks, D., Cunliffe, D., and Tudhope, D. 2006. "A Reference Model for User-System Interaction in Thesaurus-based Searching." *Journal of the American Society for Information Science and Technology*, Vol. 57, No. 12, pp. 1655–1665.

Busch, J. 2011. "Governance Maturity Model." Unpublished presentation.

Cambridge Systematics, Inc. 2013. *NCHRP Report 754: Improving Management of Transportation Information*. Transportation Research Board of the National Academies, Washington D.C.

Clarke, S. G. D. 2001. "Thesaural Relationships." In *Relationships in the Organization of Knowledge*. Springer, Netherlands, pp. 37–52.

Cleverdon, C. W. 1960. "The ASLIB Cranfield Research Project on the Comparative Efficiency of Indexing Systems." *Aslib Proceedings*, Vol. 12, No. 12, pp. 421–431.

Cleverdon, C. W., and Keen, M. 1966. Aslib Cranfield Research Project—Factors Determining the Performance of Indexing Systems: Volume 2, Test Results.

Coates, E. J. 1973. "Some Properties of Relationships in the Structure of Indexing Languages." *Journal of Documentation*, Vol. 29, No. 4, pp. 390–404.

Davis, C. H. 1968. "Integrating Vocabularies with a Classification Scheme." *American Documentation*, Vol. 19, No. 1, p. 101.

Farradane, J. 1967. "Concept Organization for Information Retrieval." *Information Storage and Retrieval,* Vol. 3, No. 4, pp. 297–314.

Farradane, J. E. L. 1970. "Analysis and Organization of Knowledge for Retrieval." *Aslib Proceedings*, Vol. 22, No. 12, pp. 607–616.

Farradane, J., Russell, J. M., and Yates-Mercer, P. A. 1973. "Problems in Information Retrieval: Logical Jumps in the Expression of Information." *Information Storage and Retrieval*, Vol. 9, No. 2, pp. 65–77.

Foskett, A. C. 1972. *Subject Approach to Information*. 6th Revised Edition. Facet Publishing.

Ganzmann, J. 1990. "Criteria for the Evaluation of Thesaurus Software." *International Classification*, Vol. 17, No. 3-4, pp. 148–157.

Greenberg, J. 2004. "User Comprehension and Searching with Information Retrieval Thesauri." *Cataloging and Classification Quarterly*, Vol. 37, No. 3-4, pp. 103–120.

Hines, T. C., and Harris, J. L. 1971. "Columbia University School of Library Service System for Thesaurus Development and Maintenance." *Information Storage and Retrieval*. Vol. 7, No. 1, pp. 39–50.

Hodge, G., and Busch, J. 2012. "Vocabulary Governance Examples." Unpublished presentation. Jan. 2012.

Höpker, W. W. 1972. "Criteria for Comparing Various Systems of Classification." *Methods Archive*, Vol. 11, pp. 144–151.

Jones, A. D. 1977. "A Comparative Evaluation of the Thesaurus of Engineering and Scientific Terms and the DDC Retrieval and Indexing Terminology." Master's Thesis, City University of London. Accessed online on August 1, 2017 http://www.dtic.mil/dtic/tr/fulltext/u2/a142607.pdf

Jones, K. P. 1971. "Compound Words: A Problem in Post-Coordinate Retrieval Systems." *Journal of the American Society for Information Science*, Vol. 22, No. 4, pp. 242–250.

Jones, K. P. 1967. "The Use of Links and Roles on a Pre-co-ordination Basis in Optical Coincidence Systems." *Aslib Proceedings*, Vol. 19, No. 6, pp. 195–199.

Keith, N. R. 1970. "A General Evaluation Model for an Information Storage and Retrieval System." *Journal of the American Society for Information Science*, Vol. 21, No. 4, pp. 237–239.

Kim, C. 1973. "Theoretical Foundations of Thesaurus Construction and Some Methodological Considerations for Thesaurus Updating." *Journal of the American Society for Information Science*, Vol. 24, No. 2, pp. 148–156.

Klingbiel, P. H. 1973a. "A Technique for Machine-aided Indexing." *Information Storage and Retrieval*, Vol. 9, No. 9, pp. 477–494.

Klingbiel, P. H. 1973b. "Machine-aided Indexing of Technical Literature." *Information Storage and Retrieval*, Vol. 9, No. 2, pp. 79–84.

Klingbiel, P. H. 1974. *Multimillion Word Data Bases: A Preliminary Report. Volume 2* (No. DDC-TR-74-1-Vol-2). Defense Documentation Center, Alexandria, VA.

La Barre, K. 2010. "Facet Analysis." *Annual Review of Information Science and Technology*, Vol. 44, No. 1, pp. 243–284.

Lancaster, F. W. 1968. "On the Need for Role Indicators in Post-coordinate Retrieval Systems." *American Documentation*, Vol. 19, No. 1, pp. 42–46.

Lancaster, F. W. 1972. *Vocabulary Control for Information Retrieval*. 2nd Edition. Information Resources Press.

Mai, J. E. 2006. "Contextual Analysis for the Design of Controlled Vocabularies." *Bulletin of the American Society for Information Science and Technology*, Vol. 33, No. 1, pp. 17–19.

Mandersloot, W. G., Douglas, E., and Spicer, N. 1970. "Thesaurus Control—The Selection, Grouping, and Cross-referencing of Terms for Inclusion in a Coordinate Index Word List." *Journal of the American Society for Information Science*, Vol. 21, No. 1, pp. 49–57.

Miller, U. 1997. "Thesaurus Construction: Problems and Their Roots." *Information Processing and Management*, Vol. 33, No. 4, pp. 481–493.

Milstead, J. L. 2001. "Standards for Relationships Between Subject Indexing Terms." In *Relationships in the Organization of Knowledge*. Springer, Netherlands, pp. 53–66.

Montague, B. A. 1965. "Testing, Comparison, and Evaluation of Recall, Relevance, and Cost of Coordinate Indexing with Links and Roles." *Journal of the American Society for Information Science*, Vol. 16, No. 3, pp. 201–208.

Neville, H. H. 1972. "Thesaurus Reconciliation." *Aslib Proceedings*, Vol. 24, No. 11, pp. 620–626.

Pickford, A. 1968. "An Objective Method for Generation of an Information Retrieval Language." *Information Scientist*, Vol. 2, No. 1, pp. 17–37.

Pickford, A. G. A. 1971. "Some Problems of Using an Unstructured Information Retrieval Language in a Co-ordinate Indexing System." *Aslib Proceedings*, Vol. 23, No. 3, pp. 133–138.

Pinto, M. 2008. "A User View of the Factors Affecting Quality of Thesauri in Social Science Databases." *Library and Information Science Research*, Vol. 30, No. 3, pp. 216–221.

Rada, R., Barlow, J., Potharst, J., Zanstra, P., and Bijstra, D. 1991. "Document Ranking Using an Enriched Thesaurus." *Journal of Documentation*, Vol. 47, No. 3, pp. 240–253.

Rada, R., and Bicknell, E. 1989. "Ranking Documents with a Thesaurus." *Journal of the American Society for Information Science*, Vol. 40, No. 5, pp. 304–310.

Rada, R., and Martin, B. K. 1987. "Augmenting Thesauri for Information Systems." *ACM Transactions on Information Systems (TOIS)*, Vol. 5, No. 4, pp. 378–392.

Rada, R., Mili, H., Letourneau, G., and Johnston, D. 1988. "Creating and Evaluating Entry Terms." *Journal of Documentation*, Vol. 44, No. 1, pp. 19–41.

Rolling, L. 1965. "The Role of Graphic Display of Concept Relationships in Indexing and Retrieval Vocabularies." In *Classification Research: Proceedings of the 2nd International Study Conference*, September 14-18, 1964, Elsinore, Denmark, P. Atherton, Ed. Munksgaard, Copenhagen, pp. 295–325.

Rolling, L. N. 1970. "Compilation of Thesauri for Use in Computer Systems." *Information Storage and Retrieval*, Vol. 6, No. 4, pp. 341–350.

Salton, G. 1972. "A New Comparison Between Conventional Indexing (MEDLARS) and Automatic Text Processing (SMART)." *Journal of the American Society for Information Science*, Vol. 23, No. 2, pp. 75–84.

Saracevic, T. 1971. "Selected Results from an Inquiry into Testing of Information Retrieval Systems." *Journal of the American Society for Information Science*, Vol. 22, No. 2, pp. 126–139.

Schirmer, R. F. 1967. "Thesaurus Analysis for Updating." *Journal of Chemical Documentation*, Vol. 7, No. 2, pp. 94–98.

Spiteri, L. 1998. "A Simplified Model for Facet Analysis: Ranganathan 101." *Canadian Journal of Information and Library Science*, Vol. 23, No. 1-2, pp. 1–30.

Stokolova, N. A. 1977. "Elements of a Semantic Theory of Information Retrieval—I. The Concepts of Relevance and Information Language." *Information Processing and Management*, Vol. 13, No. 4, pp. 227–234.

Travis, I. L., and Fidel, R. 1982. "Subject Analysis." *Annual Review of Information Science and Technology*, Vol. 17, pp. 123–157.

Tudhope, D., Alani, H., and Jones, C. 2006. "Augmenting Thesaurus Relationships: Possibilities for Retrieval." *Journal of Digital Information*, Vol. 1, No. 8. Accessed online on August 1, 2017. http://jodi.ecs.soton.ac.uk/Articles/v01/i08/Tudhope/.

Tudhope, D., Binding, C., Blocks, D., and Cunliffe, D. 2006. "Query Expansion via Conceptual Distance in Thesaurus Indexed Collections." *Journal of Documentation*, Vol. 62, No. 4, pp. 509–533.

Turski, W. M. 1971. "On a Model of Information Retrieval System Based on Thesaurus." *Information Storage and Retrieval*, Vol. 7, No. 2, pp. 89–94.

Vickery, B. C. 1971. "Structure and Function in Retrieval Languages." *Journal of Documentation*, Vol. 27, No. 2, pp. 69–82.

Willetts, M. 1975. "An Investigation of the Nature of the Relation Between Terms in Thesauri." *Journal of Documentation*, Vol. 31, No. 3, pp. 158–184.

Acronyms, Abbreviations, and Initialisms

ANSI	American National Standards Institute
BT	Broader term
DOT	Department of transportation
IM	Information management
ISO	International Organization for Standardization
ITRD	International Transport Research Documentation
KPI	Key performance indicator
KWIC	Key Word in Context
KWOC	Key Word Out of Context
NISO	National Information Standards Organization
NT	Narrower term
RDF	Resource Description Framework
RT	Related term
SKOS	Simple Knowledge Organization System
SME	Subject matter expert
SN	Scope note
S&O	Source and origin
TMS	Thesaurus management software
TRID	Transport Research International Documentation
TRIS	Transportation Research Information Services
TRT	Transportation Research Thesaurus
USE/UF	Use/Use For

Abbreviations and acronyms used without definitions in TRB publications:

A4A	Airlines for America
AAAE	American Association of Airport Executives
AASHO	American Association of State Highway Officials
AASHTO	American Association of State Highway and Transportation Officials
ACI–NA	Airports Council International–North America
ACRP	Airport Cooperative Research Program
ADA	Americans with Disabilities Act
APTA	American Public Transportation Association
ASCE	American Society of Civil Engineers
ASME	American Society of Mechanical Engineers
ASTM	American Society for Testing and Materials
ATA	American Trucking Associations
CTAA	Community Transportation Association of America
CTBSSP	Commercial Truck and Bus Safety Synthesis Program
DHS	Department of Homeland Security
DOE	Department of Energy
EPA	Environmental Protection Agency
FAA	Federal Aviation Administration
FAST	Fixing America's Surface Transportation Act (2015)
FHWA	Federal Highway Administration
FMCSA	Federal Motor Carrier Safety Administration
FRA	Federal Railroad Administration
FTA	Federal Transit Administration
HMCRP	Hazardous Materials Cooperative Research Program
IEEE	Institute of Electrical and Electronics Engineers
ISTEA	Intermodal Surface Transportation Efficiency Act of 1991
ITE	Institute of Transportation Engineers
MAP-21	Moving Ahead for Progress in the 21st Century Act (2012)
NASA	National Aeronautics and Space Administration
NASAO	National Association of State Aviation Officials
NCFRP	National Cooperative Freight Research Program
NCHRP	National Cooperative Highway Research Program
NHTSA	National Highway Traffic Safety Administration
NTSB	National Transportation Safety Board
PHMSA	Pipeline and Hazardous Materials Safety Administration
RITA	Research and Innovative Technology Administration
SAE	Society of Automotive Engineers
SAFETEA-LU	Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (2005)
TCRP	Transit Cooperative Research Program
TDC	Transit Development Corporation
TEA-21	Transportation Equity Act for the 21st Century (1998)
TRB	Transportation Research Board
TSA	Transportation Security Administration
U.S.DOT	United States Department of Transportation

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

The nation turns to the National Academies
of Sciences, Engineering, and Medicine for
independent, objective advice on issues that
affect people's lives worldwide.
www.national-academies.org

ISBN 978-0-309-44685-3

90000

9 780309 446853