



NRC-CMRC

Office of Audit and Evaluation

Evaluation of NRC's Initiative under the Roadmap for Canada's Official Languages 2013-2018

October 2016



National Research
Council Canada

Conseil national de
recherches Canada

Canada 

Prepared by:

Office of Audit and Evaluation
National Research Council Canada

Approval:

This report was approved by NRC's President on 31 October, 2016

TABLE OF CONTENTS

Acronyms and Abbreviations	ii
Executive Summary	iii
1. Introduction.....	1
1.1 Evaluation Overview	1
2. Initiative Overview	2
2.1 The Roadmap for Canada’s Official Languages.....	2
2.2 NRC’s Initiative under the Roadmap	2
2.3 Financial Resources	2
3. Relevance.....	3
3.1 Continued Need.....	3
3.2 Alignment with NRC and Federal Priorities	5
3.3 Alignment with Federal Roles and Responsibilities	6
4. Performance.....	8
4.1 Contributions to the Growth and Competitiveness of Canadian Industry	8
4.2 Resource Utilization	12
5. Conclusion	14
6. Recommendations.....	15
7. Management Response	16
Appendix A: Methodology.....	A-1
Appendix B: Selection of Documents Reviewed	B-1

ACRONYMS AND ABBREVIATIONS

AILIA	Association de L'Industrie de la Langue / Language Industry Association
CAS	Courts Administration Service
CST	Countering Security Threats using Natural Language Processing
DARPA	Defense Advanced Research Projects Agency
GALE	Global Autonomous Language Exploitation
ICT	Information and Communication Technology
ILT	Interactive Language Technologies
MATS	Multimedia Analytics Tools for Security
MTP	Multilingual Text Processing
NRC	National Research Council
OAE	Office of Audit and Evaluation
Portage	NRC's statistical machine translation software
PCH	Department of Canadian Heritage
R&D	Research and development
RTO	Research and Technology Organization
SEC	Senior Executive Committee
SME	Small- and medium-sized enterprise
SMT	Statistical Machine Translation
UQO	Université du Québec en Outaouais
U.S.	United States

EXECUTIVE SUMMARY

This report presents the results of the 2016 evaluation of National Research Council's (NRC) research, development and commercialization activities funded under the Roadmap for Canada's Official Languages 2013-2018 (hereafter referred to as the "Roadmap"). NRC's initiative under the Roadmap is one of 28 initiatives being implemented by 14 federal partners. As with the two previous five-year action plans, NRC received \$10M over five years to contribute to the Roadmap. This funding is used by the Information and Communication Technology (ICT) Portfolio for R&D into natural language processing technologies in support of the growth and competitiveness of the Canadian language industry and other Canadian industries.

This evaluation was conducted by an independent evaluation team from the NRC Office of Audit and Evaluation (OAE). It provides information on the relevance and performance (including effectiveness and efficiency) of NRC's initiative under the Roadmap. Given the relatively small size of the NRC initiative and the fact that NRC previously conducted a comprehensive evaluation of the initiative in early 2012¹, the OAE undertook a calibrated evaluation of the funded initiative. The evaluation was streamlined to narrow the scope and employ a limited number of methods compared to traditional evaluations. Calibrated evaluations allow for the effective utilization of evaluation resources, while maintaining the credibility and usability of the evaluation results. It is also worth noting that an evaluation of the broader ICT Portfolio is planned for 2017-18.

The evaluation methods included a review of internal and external documents (key sources included in the bibliography) as well as administrative data. In addition, in-depth interviews were conducted with internal (4) and external (4) stakeholders as well as with clients and collaborators (6).

Relevance:

The following are the findings of the evaluation with regards to the relevance of the initiative:

1. There is demand for Statistical Machine Translation (SMT) technology by the Canadian public sector and from industry domains outside the traditional focus of language translation. However, there is little evidence of significant industry pull for SMT technology from the principle target of NRC's initiative under the Roadmap - the Canadian translation industry.
2. NRC's commitments under the Roadmap are consistent with its strategic outcomes as well as with federal priorities. NRC's initiative is also broadly aligned to the ultimate outcome of the Roadmap. However, the alignment of NRC's activities and outcomes to the objectives of the education pillar, where it has been placed within the Roadmap's results architecture, could be better articulated.
3. NRC's role in research and development of natural language processing technologies is appropriate, given that it aligns with NRC's mandate and that internationally Canada's

¹ A previous evaluation of the initiative was conducted in 2012; however, as the Department of Canadian Heritage (PCH) is currently conducting a horizontal evaluation of the broader Roadmap, in order to meet the terms of the Horizontal Coordination Framework for Roadmap 2013-2018, an evaluation covering the three-year period of 2013-14 to 2015-16 is needed. This current evaluation will serve as a source of data to support the horizontal evaluation.

peer governments are making similar investments. NRC's ICT Portfolio has a critical mass of experts in natural language processing, who are recognized world-wide, supporting the notion that NRC's role in delivering this initiative is conducive to strong results.

4. NRC plays a vital role in maintaining Canada's domestic capacity in natural language processing. In its absence, there is no other organization, public or private, willing or able to take on this role.

Performance:

Performance was considered in the context of:

- the research team's contribution to the state-of-the-art in natural language processing;
- impact on Canadian language industry and other industries; and
- efficiency.

The following are the findings of the evaluation with regards to performance:

1. The success of the initiative in terms of its excellence in research and development has not translated to significant uptake by the private sector. Presently, the receptor capacity for the technologies developed has been most apparent in the public sector.
2. NRC's SMT software, developed under the initiative, has had a significant impact on the public sector.
3. NRC's suite of language technologies, developed with assistance from Roadmap funding, has had a broader application beyond the Canadian language industry.
4. NRC resources are being used efficiently in responding to the organization's commitments under the Roadmap.

Recommendations:

The evaluation put forth one recommendation:

NRC should consider its fit within the education pillar of the Roadmap and seek to better articulate how its activities and expected outputs align with the Roadmap's results architecture.

1. INTRODUCTION

This report presents the results of the 2016 evaluation of National Research Council's (NRC) Initiative under the Roadmap for Canada's Official Languages 2013-2018 (hereafter referred to as the "Roadmap"). NRC's initiative under the Roadmap is one of 28 initiatives being implemented by 14 federal partners. As was the case in the two previous five-year action plans, NRC committed \$10M over five years to the Roadmap. This funding is used by the Information and Communication Technology (ICT) Portfolio to contribute to the growth and competitiveness of the Canadian language industry and other Canadian industries.

The Department of Canadian Heritage (PCH) is conducting a horizontal evaluation of the broader Roadmap. The current evaluation of NRC's initiative under the Roadmap is required under the terms of the Horizontal Coordination Framework for Roadmap 2013-2018 and will serve as a source of data in support of the horizontal evaluation.

1.1 Evaluation Overview

This evaluation was conducted by an independent evaluation team from the NRC Office of Audit and Evaluation (OAE). It covered the three-year period of 2013-14 to 2015-16 and gathered information on the relevance and performance (effectiveness and efficiency) of NRC's initiative under the Roadmap. Given the relatively small size of the NRC initiative and the fact that a comprehensive evaluation was conducted in early 2012, the OAE conducted a calibrated evaluation of the funded initiative. The evaluation was streamlined to narrow the scope and employ a limited number of methods compared to traditional evaluations. Calibrated evaluations allow for the effective utilization of evaluation resources, while maintaining the credibility and usability of the evaluation results. It is also worth noting that an evaluation of the broader ICT Portfolio is planned for 2017-18.

The evaluation questions, outlined at the start of each section, were developed following consultations with portfolio management and a review of key documents during the planning stage. The evaluation design meets NRC senior management needs as well as the requirements of the 2009 Treasury Board *Policy on Evaluation*. It also addresses certain issues that are common to all Roadmap partners, as per requirements stated in the Horizontal Coordination Framework for Roadmap 2013-2018.

The specific methods used in the study included:

- Internal and external document review
- Administrative and performance data review
- Semi-structured interviews
 - 5 internal interviews
 - 10 external interviews (including clients and language industry experts)

A more detailed description of the evaluation design, including the evaluation questions, methods, limitations and challenges is provided in **Appendix A: Methodology**.

2. INITIATIVE OVERVIEW

2.1 The Roadmap for Canada's Official Languages

The Roadmap for Canada's Official Languages 2013-2018 is the Government of Canada's official languages strategy for the period between April 1, 2013 and March 31, 2018. It follows the Roadmap for Canada's Linguistic Duality 2008-2013, and reaffirms the Government of Canada's commitment to promote official languages and enhance the vitality of official-language minority communities. The Roadmap includes 28 initiatives, which are implemented by 14 federal organizations. At an aggregate level, the Roadmap aims to achieve the following strategic outcome: "Canadians live and thrive in both official languages and recognize the importance of French and English for Canada's national identity, development and prosperity."

The Roadmap's initiatives and associated expected outcomes are grouped according to three broad pillars: Education, Immigration and Communities. NRC's commitment falls under the Education pillar.

2.2 NRC's Initiative under the Roadmap

NRC's initiative under the Roadmap seeks to contribute to the growth and competitiveness of the Canadian language industry and other Canadian industries through research and development (R&D). To achieve this, NRC conducts R&D to advance the state-of-the-art in natural language processing, the computer-based processing of information in multiple languages. NRC's core research activities in this area revolve primarily around the development of:

- Statistical Machine Translation (SMT) - the techniques that allow computers to learn how to automatically translate text from one human language to another from an existing body of bilingual documents.
- Text Analytics - the techniques that allow computers to extract and synthesize useful information from large collections of textual sources.

NRC's activities related to the Roadmap initiative are conducted by two resource teams within the Information Technologies directorate of the ICT Portfolio: the Multilingual Text Processing (MTP) team and the Text Analytics team.

2.3 Financial Resources

The Roadmap itself represents a government-wide commitment of \$1.15 billion over 5 years (2013-2018). Since 2013, all federal initiatives under the Roadmap have on-going, A-base funding. As was the case in the previous five-year action plans, NRC received \$10M over five years to contribute to the Roadmap.

3. RELEVANCE

The relevance of NRC's Initiative under the Roadmap was examined from the perspectives of: whether the initiative responds to an identifiable industry need (section 3.1); its alignment with NRC strategic outcomes and federal priorities (section 3.2); and its alignment with federal roles and responsibilities (section 3.3).

3.1 Continued Need

Evaluation Question 1: Does NRC's initiative under the Roadmap respond to an identifiable industry need?

Finding 1: *There is demand for SMT technology by the Canadian public sector and from industry domains outside the traditional focus of language translation. However, there is little evidence of significant industry pull for SMT technology from the principle target of NRC's initiative under the Roadmap - the Canadian translation industry.*

In order to assess the extent to which the NRC initiative under the Roadmap responded to an identifiable industry need, the evaluators examined demand from the Canadian language industry (section 3.1.1), demand from other industry sectors (section 3.1.2), and the ability to meet these needs in the absence of the NRC's Roadmap initiative (section 3.1.3).

3.1.1 Demand from the Canadian language industry

The Roadmap for Canada's Official Languages 2013-18 identifies the Canadian language industry as the principal target for NRC R&D supported under the Roadmap initiative. Evidence from interviews with internal and external stakeholders suggests that the 'language industry' is hard to define and is used in practice as a synonym of the translation sector.

Based on interviews with experts and a review of documents, NRC language technologies funded under the Roadmap have the potential to increase the productivity and competitiveness of Canada's translation industry. They also offer the technological foundation for Canadian firms to realize market opportunities that are emerging in the translation space.² To remain competitive, Canadian translation firms are looking to adopt new technologies in order to increase productivity, while preserving translation quality, as the industry is faced with an ever-increasing translation demand. By producing and transferring leading-edge technologies to Canadian language technology firms and private sector translation firms, NRC has the potential to increase their capacity for innovation and their ability to compete nationally and internationally.

An external translation industry expert, who was interviewed as part of the evaluation, noted that machine translation, in particular, is very important to the translation space. This expert highlighted the "explosive" success of Google Translate (which uses machine translation), but noted that Google Translate is not suited to all uses. As noted by both internal and client interviewees, Portage, NRC's statistical machine translation system, has an advantage over cloud-based systems (such as Google Translate) for companies with particular security requirements since, with Portage, documents and information being translated remain on the

"Machine translation is very important to the translation space. There are not enough translators to meet the demand. The idea is not to replace human translators, but to accelerate their work."

- Language industry expert

² SLATOR. (2016) Slator's State of Translation and Localization Demand 2016

company's secure servers. As well, Portage generates better quality translations in cases where the client operates in a particular domain, with consistent and limited terminology since the system can be trained on this particular vocabulary.

However, although overall it is evident that SMT is an important component of a modernized translation industry and the NRC SMT is of high quality and addresses concerns posed by other firms, the capacity and appetite for the investment required is not apparent in the private sector in Canada. The industry is small and fragmented – lacking the financial or technical capacity to invest in NRC's SMT technology or the critical volume of translation to justify the investment. There is also a certain degree of resistance from some translators, who are culturally opposed to machine translation. Overall, significant demand from the Canadian private sector is not evident.

As will be discussed in the section on performance, NRC has been successful in licensing Portage to Canadian translation organizations such as the Translation Bureau (public sector), as well as to their resale partner Terminotix. There are also a handful of Canadian translation companies that have, in turn, licensed Portage through Terminotix. However, the main opportunities for the exploitation of NRC's language technologies lie in sectors other than translation.

Despite the shortcomings in receptor capacity within the Canadian translation industry it is apparent that NRC's work in the area of language technologies responds to current needs as well as future trends. All five client representatives that were interviewed as part of the evaluation noted that NRC's work aligns with the needs and priorities of their organization and all reported a high likelihood that they would work with NRC again in the future.

3.1.2 Demand from other industry sectors

Notwithstanding the traditional focus on the language industry, within NRC's initiative under the Roadmap, the projects that are currently being sought are predominantly originating from the security, defense and intelligence sectors. This was corroborated by internal and external interviewees and by internal data and documentation. Organizations in these domains need to ensure that the systems they use are secure, which is a key advantage of Portage over competitors. NRC's language technologies have also been applied in the health sector, in areas of health information processing and analytics. Finally, there are also many projects being undertaken and future opportunities being explored with public sector organizations.

3.1.3 Ability to meet these needs in the absence of the initiative

In most cases, the clients that were interviewed as part of the evaluation noted that there were other options available to them with respect to machine translation or other natural language processing. In the absence of NRC technology, most clients indicated that they would have been able to acquire similar technology elsewhere, but that this likely would have been more expensive, more complex (in cases where the clients benefited from having multiple language technologies available through NRC), and likely would have come from outside of Canada. As stated in the 2012 evaluation, "the language technology and translation industries in Canada lack the R&D capabilities and financial resources to make a significant investment in the development of innovative language tools and technologies."

3.2 Alignment with NRC and Federal Priorities

Evaluation Question 2: Are NRC activities and objectives under the Roadmap aligned to federal government priorities and NRC strategic outcomes?

Finding 2: *NRC's commitments under the Roadmap are consistent with its strategic outcomes as well as with federal priorities. NRC's initiative is also broadly aligned to the ultimate outcome of the Roadmap. However, it is unclear how NRC's activities and outcomes align to the objectives of the education pillar, where it has been placed within the Roadmap's results architecture.*

3.2.1 Alignment with NRC strategic outcomes and federal priorities

NRC's Roadmap commitments (see the text box, at right) are well aligned with NRC's first strategic outcome, that is, "Canadian businesses prosper from innovative technologies". They also reflect NRC's focus on industry-relevant R&D, which has been identified as a priority in recent federal budgets.³

The Roadmap for Canada's Official Languages 2013-2018 is a Government of Canada policy statement, which constitutes the Government of Canada's official languages strategy for the period between April 1, 2013 and March 31, 2018. The Roadmap also aligns with federal commitments under Part VII of the *Official Languages Act*. As an initiative under the Roadmap, NRC's activities and their associated outcomes are fully aligned with the federal government's priorities around bilingualism and official languages.

NRC's Roadmap Commitments

Activity: Strengthening of the language industry and language technologies

Outcome: Contribution to the growth and competitiveness of the Canadian language industry and other Canadian industries through research and development.

3.2.2 Alignment with the Roadmap's ultimate outcome and pillar objectives

In addition to examining the extent to which NRC's activities under the Roadmap align with government priorities and NRC's strategic outcomes, the evaluation also examined alignment of NRC's activities within the Roadmap.

The 28 initiatives under the Roadmap all contribute toward the ultimate outcome: "Canadians live and thrive in both official languages and recognize the importance of French and English for Canada's national identity, development and prosperity." NRC's initiative has the potential to make concrete impacts in this regard and, in some cases, these impacts have already been realized. Implementation of NRC's machine translation tool Portage, within public sector organizations, is increasing the efficiency and productivity of human translators and is allowing native speakers of French and English to use secure translation tools that help them get a general understanding of informal written communications in their second language. Both of these, in turn, contribute directly to maintaining and reinforcing Canada's linguistic duality.

The architecture of the Roadmap is structured so that the results of the various initiatives contribute to achieving objectives under one of three broad pillars (education, immigration and communities). NRC's initiative falls under the education pillar, which seeks to contribute to the following objective: "Canadians benefit from education and training opportunities in their first official language and for learning the other official language of the country, and from access to

³ Government of Canada. (2015) Budget 2015. p.108.

technological tools, taking advantage of the many economic, cultural and national identity advantages resulting from these.”⁴

As a result of the broad scope of the Roadmap, it is understandable that the pillar objectives are similarly broad. However, the direct connection between NRC’s research and development activities and the high level objective of the Education pillar is not clearly articulated by the Roadmap or by NRC. Through the course of the evaluation, both internal and external stakeholders struggled to find tangible connections between the results statement of the education pillar and the work being done at NRC.

Although the long term vision for machine translation is the complete automation of translation across numerous language pairs which would have tangible impacts on accessibility to education in any language; interviewees note that, in the short to medium term, NRC’s language technologies are mainly applicable as a productivity enhancement for traditional human translation as well as for gathering and analyzing large quantities of data for use across a number of industries including security, defense and intelligence, business analytics, health, and legal.

Recommendation: *NRC should consider its fit within the education pillar of the Roadmap and seek to better articulate how its activities and expected outputs align with the Roadmap’s results architecture.*

3.3 Alignment with Federal Roles and Responsibilities

Evaluation Question 3: Is NRC’s role in delivering this initiative appropriate and conducive to strong results?

Finding 3: *NRC’s role in research and development of natural language processing technologies is appropriate, given that it aligns with NRC’s mandate and that internationally Canada’s peer governments are making similar investments. NRC’s ICT Portfolio has a critical mass of experts in natural language processing, who are recognized world-wide, supporting the notion that NRC’s role in delivering this initiative is conducive to strong results.*

3.3.1 Appropriateness of NRC’s role

NRC’s activities and outcomes under the Roadmap are in line with the mandate of NRC, as stated in the *NRC Act*. NRC’s specific role in support of the Canadian language industry and other industries stems from subsection 5 (1) (c) of the *NRC Act* which states that the Council may “undertake, assist or promote scientific and industrial research”. Under the *Act*, NRC is charged with the direction or supervision of research undertaken by or for industrial firms or other organizations [subsection 5 (1) (d)] and with carrying out experimental and developmental work with respect to the above and making the resulting processes, methods and products available for the benefit of manufacturing and other scientific purposes [subsection 5 (1) (k)].

In other countries, national governments are active in financing the development of natural language processing technologies. For example, the Spanish government has recently invested over \$100 million in the development of natural language processing and machine translation technology. The investment is made with similar objectives to those of Canada’s Official Languages Roadmap; specifically, Spain’s objective is to “strengthen the support for Natural

⁴ Horizontal Initiative - Roadmap for Canada's Official Languages 2013-18, <http://canada.pch.gc.ca/eng/1455797494601>.

Language Processing and Machine Translation in Spanish and Spain’s co-official languages,” which are Galician, Basque, and Catalan.⁵ Similarly, the United States (U.S.) Government, through the Defense Advanced Research Projects Agency (DARPA) has invested heavily in natural language processing through the Global Autonomous Language Exploitation (GALE) program.

Finding 4: *NRC plays a vital role in maintaining Canada’s domestic capacity in natural language processing. In its absence, there is no other organization, public or private, willing or able to take on this role.*

External interviewees were asked whether there was a role for government to play in supporting the development of machine translation. Responses were mixed, with one external expert seeing no role for government in “picking winners and losers” and subsidizing efforts by private sector companies engaged in developing these technologies. However, most external interviewees (clients and experts) responded more positively, noting that, in the Canadian context, there are no firms or academic organizations with:

- a) the critical mass to invest in developing these sorts of technologies, and
- b) the necessary skill level to meet the requirements of this specialized area.

As such, NRC was seen to play a vital role in maintaining Canada’s domestic capacity in this important area of intellectual property; in its absence, no other organization, public or private, is willing or able to take on this role. Others also reported that maintaining a national capacity for statistical machine translation is critical for all countries, but even more so for bilingual nations. There are a multitude of scenarios where sending data offshore or “to the cloud” is impractical from a security standpoint.

3.3.2 Evidence of NRC’s world-class team and its outputs

NRC consistently ranks among the top teams at international competitions, including placing first in international competitions in text mining (i2b2) and in statistical machine translation (U.S. National Institute of Standards and Technology 2012 Open Machine Translation Evaluation, or OpenMT12).⁶ Looking specifically at Portage, experiments using the BLEU translation assessment framework demonstrated that when Portage is used as recommended by NRC – to translate texts from the same domain as the given machine has been trained on – it will often yield translations that are of better quality than those obtained from Google Translate. However, Google Translate will yield better translations for out-of-domain or mainly out-of-domain texts⁷.

External language industry experts also confirmed that NRC’s natural language processing team is of a world-class calibre and is recognized worldwide. NRC was selected by SRI International (previously known as the Stanford Research Institute) as their technology partner in the recent GALE DARPA program, and by Philips Research in an FP7 text-mining project. Asked what other teams in Canada were operating at this level, external interviewees noted that there were a number of academics working in the same space with good results but that NRC was Canada’s premier group within this specialized domain. In short, according to external interviewees, NRC’s team is not only the best team in Canada; they are also among the best in the world.

⁵ SLATOR. (2016) Slator’s State of Translation and Localization Demand 2016.

⁶ National Research Council. (2015). ICT Five Year Strategic Plan: Fiscal Year 2015 to 2010. P. 11

⁷ Kuhn, R., Stewart, D., and Islam, R. (2015). A BLEU Comparison of Portage vs. Google Translate on Four Domains. P. 2

4. PERFORMANCE

The evaluation assessed the performance of NRC's Roadmap initiative in terms of: the extent to which the initiative has contributed to the growth and competitiveness of Canadian industry (section 4.1); and the extent to which NRC has demonstrated an efficient use of resources in responding to its commitments under the Roadmap (section 4.2).

The performance of the team funded by the initiative could also be considered in the context of their output as researchers advancing the state-of-the-art in natural language processing technologies. To some extent, section 3.3.2 above outlines some of the success that the team has achieved in that domain. However, given the objectives outlined in the Roadmap for Canada's Official Languages, this evaluation focuses on performance in the context of the growth and competitiveness of Canadian industry.

4.1 Contributions to the Growth and Competitiveness of Canadian Industry

Evaluation Question 4: To what extent has the initiative contributed to the growth and competitiveness of the Canadian language industry and other Canadian industries through R&D?

4.1.1 Impact on the Canadian language industry

Finding 5: *The success of the initiative in terms of its excellence in research and development has not translated to significant uptake by the private sector in the area of language translation. Presently, the receptor capacity for the technologies developed has been most apparent in the public sector.*

NRC's performance in terms of its impact on the Canadian language industry has been marginal. This can be explained by the low uptake of the SMT technology, and internal and external factors which are explored below.

NRC's work in support of the Canadian language industry revolves around its statistical machine translation system, Portage. This software tool allows a computer to translate text between language pairs. The principal application for this technology is improving the productivity of translators by generating a first draft of a given translation. The technology can also be used for 'gysting' – which means using a rough translation to understand the general idea being conveyed within a given text.

Since April 2013, NRC has licensed Portage to a handful of Canadian translation agencies via a reseller called Terminotix, a Montreal-based company specializing in software tools for computer-aided translation. A representative of one of these translation agencies (which is, in fact, one of the larger translation agencies in Canada) was interviewed as part of the evaluation. This representative noted that Portage is now used to produce first drafts of English-to-French and French-to-English translations for approximately 80% of the company's clients. These translations are then post-edited by human translators. She noted that having access to machine translation has increased the firm's competitiveness with clients that specifically request the use of the technology (approximately 10% of all contracts). Although the company anticipates that Portage will also result in increased translator productivity, it is still evaluating whether this is in fact the case. Terminotix reported that some of its clients have reported a 15-20% increase in translator productivity.

Despite these successes in licensing Portage, it is evident that demand from Canadian Industry for this type of product is limited. There are less than five license agreements in place with Canadian language industry clients, although more significant impacts were raised for public sector clients (see also section 4.1.2) and clients in other industry sectors (see also section 4.1.3). Some internal interviewees as well as one external expert admitted that, given the productivity promise of machine translation and the quality of NRC's technology, there is the potential for NRC to have a greater impact on the translation industry.

These findings are consistent with the findings of the previous evaluation in 2012 which found that while the group had undertaken significant efforts to transfer its technologies to its targeted clients, these efforts yielded limited results. The performance of the team in terms of commercialization can be clarified by both external and internal factors outlined below.

External Factors Affecting NRC's Commercialization Performance

The dynamics of the Canadian 'Language Industry' are pertinent to NRC's performance in commercializing Portage. The concept of a 'Language Industry' is the product of Industry Canada's 'Language Industry Initiative' which was launched in 2003 in response to a prediction that demand for 'language services' would dramatically increase and the reality that the sector is fragmented and composed predominantly of micro-industries.⁸

One of the initiative's two outputs was the creation of the Language Industry Association (AILIA⁹) which sought to bring together the translation industry and the language training and language technology industries. However, these 'industries' had few common interests and in 2013 AILIA dropped its representation of the 'training' and 'technology' industries to focus solely on translation.¹⁰ Now, for the most part, the term 'language industry' is used as a synonym for 'translation industry.'

Worldwide, the translation industry is extremely large and growth in the demand for translation services is expected to outstrip supply.¹¹ According to industry reports, machine translation is expected to be an important component in bridging that gap¹². The increasing need for translation from multi-lingual government institutions like the European Union, along with the need for global commerce to make products and services available in the local languages of different markets, are two examples that contribute to the expected growth in demand.¹³ Further, new and lucrative market opportunities for the application of SMT are opening up in the legal, financial and health care markets.¹⁴ Overall, SMT is a technology that is highly relevant to the current market landscape.

Because of the size and dynamic of Canada's translation firms, there is little appetite or capacity for investing in the technologies required to capitalize on the aforementioned market opportunities. The Canadian translation industry is small and is generally comprised of micro-enterprises.¹⁵ According to the 2008 Evaluation of the Language Industry Initiative, the

⁸ Industry Canada. (2008). Final Evaluation of the Language Industry Initiative.

⁹ The AILIA acronym comes from a combination of both the French and English association names - Association de l'industrie de la langue/Language Industry Association

¹⁰ External Interview

¹¹ Choudhury, R., McConnell, B. (2013). TAUS: Translation Technology Landscape Report.

¹² Ibid.

¹³ Ibid

¹⁴ SLATOR. (2016). Slator State of Translation and Localization Demand 2016.

¹⁵ Industry Canada. (2008). Final Evaluation of the Language Industry Initiative.

translation and interpretation sector is made up of approximately 800 firms which employ 6.2 workers on average, with 76.9% of translation companies reporting fewer than five employees.¹⁶

Canada's fragmented translation industry is problematic for the commercialization of Portage. The technology, in its current state, is not suitable for small and medium-sized enterprises (SMEs). NRC's Portage software is geared toward large companies that have the financial and technical capacity to integrate its complex and raw technology into their own custom-built interfaces. Portage produces high quality translations that rival those of its major competitors including Google Translate. However, Google Translate is provided to its clients free of charge through a highly intuitive web-based user interface. Portage has been developed as enterprise software which lacks a user interface and requires the installation of expensive and complex hardware on a client's premises. This is also Portage's strength and its niche. Portage is marketed to firms who are seeking a secure translation service which does not require data to be uploaded to the Internet or off-premises. Portage also offers its enterprise clients the option to train the software on specific technical language that is relevant to their particular industry. This allows Portage to offer translation quality that surpasses that of its principle competitors and which requires less post-editing from human translators. This, then, offers productivity savings to firms that operate at a sufficiently large scale such that productivity gains more than offset the cost of licenses and employing technical resources to integrate and manage the technology. Unfortunately the Canadian translation sector is not characterised by many large translation businesses that have the financial or technical capacity to integrate this technology or operate at a sufficiently large scale to justify the business case for an investment in Portage.

It should be noted that Portage has been licensed to one of Canada's largest private sector translation players as well as by the Translation Bureau, which, although public, is among the largest producers of translation products in Canada. NRC's success in deploying its technology to these two organizations demonstrates that when demand is present, NRC's Portage offers a competitive and viable option.

Internal Factors Affecting NRC's Commercialization Performance

There are also internal factors that have contributed to the lack of commercialization success for Portage. The dynamics of Canada's translation sector and the applicability of Portage for SMEs was known by the NRC researchers and managers interviewed during the evaluation. These internal interviewees expressed a desire to improve Portage's usability for smaller firms but noted that to do so with the current team would be inefficient. It was noted that the current team is over-qualified for the work needed to improve Portage's usability, and is fully engaged in other projects. Therefore, developing a user-friendly technology would require hiring additional resources that could be dedicated to this project, allowing the current team of researchers to focus their time on advancing the state-of-the-art in SMT.

As an alternative to hiring resources to undertake the work of improving the Portage user interface, in an ideal world, an 'integrator' would step forward or be found within the private sector who could take the high-quality but raw technology created by NRC, develop the client-facing user interface and play the role of commercializer. However, internal interviewees note that no such value-chain actor has been identified.

Several of the internal NRC researchers interviewed also referenced NRC's revenue focus as an influence on the team's ability to focus on Portage and the translation industry. As mentioned

¹⁶ Ibid.

above, there is no significant demand from private sector translation companies and thus, projects in this domain are not strong revenue generators.

4.1.2 Impact on public sector organizations

Finding 6: *NRC’s Statistical Machine Translation (SMT) software developed under the initiative has had a significant impact on the public sector.*

NRC has been highly successful in deploying Portage within the Canadian public sector. In 2015-16, the Translation Bureau signed a new license for Portage to over 350,000 desktops across the public service. One stated goal of the project is to sharply reduce the usage of non-secure but freely available tools such as Google Translate, Bing, and Reverso. Currently, these tools receive over a million hits a week (about 60 million hits a year) from IP addresses within the Public Service of Canada, which results in fragments of protected documents being sent to the cloud.¹⁷

Over the past year, the Courts Administration Service (CAS) and NRC, along with an independent evaluator, conducted a pilot study to determine whether the use of Portage would reduce the cost and turnaround times of CAS’ translation workload (8 million words a year). NRC has built eight versions of Portage for the four courts whose judgements CAS is responsible for translating (an English-to-French and French-to-English version for each court). A report released by the independent evaluator in March 2016 concluded that the translations produced are of sufficient quality to increase productivity in the courts. CAS is now interested in implementing Portage more broadly this fiscal year, as it faces pressure to increase productivity.

“A lot of what we’ve been doing in relation to the Roadmap is around finding ways to roll out English/French machine translation to government agencies.”

- Internal interviewee

NRC has also successfully rolled out their Portage technology to the weather broadcasting service operated by Environment Canada. The MÉTÉO project carried out by the federal Translation Bureau has provided Environment Canada with versions of Portage that translate weather reports (English to French and French to English). These systems have replaced older machine translation systems that were previously being used.

National interest - An external interviewee highlighted that having in-country access to statistical machine translation capacity is, to some extent, of national importance. As was previously discussed, Portage is deployed off of a machine installed on the premises of its clients. Competitors such as Google Translate require that translation material is sent to the “cloud” or otherwise transferred off a client’s premises to the servers of the service provider. Portage’s principle niche is the security of client data. This is an attractive option for public sector organizations including the House of Commons that need to translate highly sensitive material and would prefer not to outsource this to foreign firms. From this perspective, Portage offers a benefit to Canada above and beyond its economic impact.

4.1.3 Impact on other industries

Finding 7: *NRC’s suite of language technologies, developed with assistance from Roadmap funding, has had a broader application beyond the Canadian language industry.*

¹⁷ CBC News. (2016). Online translator helps federal workers ‘do their job,’ say defenders.

NRC's natural language processing teams have focused a large part of their efforts toward addressing demand for R&D originating from outside the 'language industry.' These projects are to some extent tangential to the objectives of the Roadmap. However, these projects do contribute to the general advancement of NRC's language processing technologies while generating modest revenue. Listed below are a number of significant projects that demonstrate the team's impact outside of the 'language industry.'

Security intelligence – In November 2015, NRC and its partners successfully completed the Countering Security Threats using Natural Language Processing (CST) project, which provided end-users in the security sector with advanced tools for filtering, summarizing and analyzing multilingual texts to monitor risks to Canada as well as public perceptions of those risks. In the course of this project, NRC – together with two industry partners and a professional intelligence service – developed a prototype system that illustrated the potential of these technologies. The tools used in the CST project are based on multiple NRC language technologies (including information extraction, emotion/ sentiment analysis, anomaly detection, and machine translation), which are combined in a way that is much more powerful than any one of them alone.

One of the industry partners for the CST project is also now collaborating with NRC on a separate project aimed at using NRC's language technologies to enhance the company's social media monitoring and analysis tool (this tool is applicable within the security domain as well as within a number of other industry sectors). This has resulted in a new product and service offering for the company, thereby improving its competitiveness.

Marine security – NRC conducted the Arctic Maritime Awareness for Safety and Security project, in collaboration with a defense contractor and with other federal partners. The project used natural language processing technologies to help human analysts remain aware of marine vessels. A key contribution was a web scraper that collects maritime data from across the Internet. NRC developed a state-of-the-art scraper on a number of vessel-specific webpages, which has more functionality than any one of its open-source competitors. A second contribution is the Named Entity Recognizer, which recognizes individuals, boat names, etc., that are mentioned in natural language text.

Healthcare – NRC also applied its natural language processing technologies within the healthcare sector. For example, NRC undertook a project for the Public Health Agency of Canada to develop a health surveillance tool for global mainstream media. The tool provides health analysts with alerts and summaries of emerging health issues as they appear in media articles from around the world. The system includes automated categorization and annotation of documents in multiple languages, machine translation, summarization (single- and multi-document), and detection of duplicate information. NRC is also working with a private sector company in the healthcare market, offering text analytics tools to assist with health information processing.

4.2 Resource Utilization

Evaluation Question 5: To what extent has NRC demonstrated an efficient use of resources in responding to its commitments under the Roadmap?

Finding 8: NRC resources are being used efficiently in responding to the organization's commitments under the Roadmap.

NRC measures the efficiency of labour resources using a ratio of program utilization as a percentage of total hours recorded. The two groups involved in the Roadmap activities reported a high degree of labour resource utilization. For 2015-16, 84% of MTP group labour resources and 83% of labour resources within the text analytics group were utilized on programs. This significantly exceeds the ICT Portfolio average utilization rate of 72%. For both groups, over 90% of this program time is spent contributing to the Multimedia Analytic Tools for Security (MATS) program.¹⁸

NRC management and staff, interviewed as part of this evaluation, all commented that NRC's activities in support of the Roadmap are very efficient. Many internal interviewees, as well as two experts noted that the group of researchers at NRC working in language technologies has had a great deal of success, particularly given the size of the team. As noted above, Portage consistently scores among the best machine translation technologies in international evaluations. It is worth noting that the other organizations that make up the group of world leaders in this area are larger research teams with access to significantly greater resources than what NRC can provide (for example, the University of Southern California Information Sciences Institute, Stanford University NLP Group, Google Research and BBN Technologies).

Many attributed this to the funding commitment through the Roadmap in that it has allowed for the maintenance of a high calibre research team and has contributed to the development and improvement of the technology, incorporating modern techniques and resulting in significant performance improvements. For example, the most recent version of Portage incorporates the use of neural networks, which improves the quality of the machine translation and allows for more accurate modelling of language details.

Some internal interviewees also pointed out that there are efficiency gains resulting from the scope of activities being undertaken. Although translation was the primary application, capabilities have now shifted to include new applications and new industry sectors. NRC language technologies are now being applied to multiple projects, allowing NRC to leverage additional funding opportunities. For example, project funding and collaborations in the security space are advancing the state-of-the-art in natural language processing, given that the core technologies can be applied to many different areas.

4.2.1 Savings in leasing costs

Prior to 2013, the precursor to the MTP group was located on the campus of the Université du Québec en Outaouais (UQO), in Gatineau, Québec. The 2012 evaluation that examined NRC's initiative under the previous Roadmap (2008-2013) found that there were opportunities for increased economies with regard to the space being leased at the time and included a recommendation that "NRC should ensure that the space rented reflects the group's needs and adjust future leasing agreements accordingly if the group remains in its current facilities". In addition, the 2012 evaluation observed that there were limited synergies between NRC and the UQO and little evidence that collaborative projects would be undertaken in the near future.

After a review of space requirements, NRC relocated the Gatineau-based group working in information and communications technologies to its Montreal Rd. campus in Ottawa. Not only did this reduce costs associated with maintaining a satellite facility, but internal interviewees

"More and more, these applications require a palette of different technologies. There is very good synergy between the two teams."

- Internal Interviewee

¹⁸ National Research Council (2015). Corporate Key Performance Indicators. Internal Document.

also stressed that this resulted in increased synergies with other NRC groups (most notably, the text analytics group). They commented that the number of projects with cross-team involvement has increased since that time. In 2015-16, over a third of the projects being undertaken through the MATS program involved resources from both teams.

4.2.2 Opportunities for improved efficiency

Internal interviewees were asked about opportunities for improved efficiency. The response that was most often given related to the fact that some clients lack the technical capabilities to embed the technology into their product. In these cases, NRC researchers can spend a lot of time doing the programming, user interface design, training of the machine and customer support required to integrate the technology with clients' information technology systems. Although NRC does have developers in-house, interviewees noted that there are too few of them and this work can end up taking away from research activities. Some did note that with some larger revenue projects that are anticipated in the next year, there will be the opportunity to hire additional staff and thereby allow research staff to focus on the research.

Many interviewees also noted that there is a missing link in the value chain -- an integrator who could provide customer service and other aspects of commercialization. To date, NRC has not been able to identify a large industry player who could suitably fill this role.

5. CONCLUSION

Overall, the evaluation found that the work being funded by the Roadmap was aligned to the mandate of the NRC and appropriate to the role of the federal government. The evidence of industry need was more complicated. Global translation industry trends demonstrate the current importance of statistical machine translation technology as well as the emerging opportunities for firms who invest. However, it was also apparent that the small firms that generally comprise the Canadian translation industry lack the receptor capacity to integrate NRC's technology or the scale of operations to justify the investment on the basis of productivity gains. It was also concluded that NRC's activities and outputs were to some extent misaligned with its placement in the results structure of the Roadmap. This led to the recommendation that NRC should consider its fit within the education pillar of the Roadmap and seek to better articulate how its activities and expected outputs align with the Roadmap's results architecture.

The evaluation also found that NRC's performance in terms of strengthening Canadian industry was not strong. However, as touched upon above, this was largely due to the dynamics and receptor capacity of the language industry rather than the quality or market-relevance of NRC's technology. On a positive note, NRC's deployment of Portage to the Translation Bureau as well as to a leading provider of translation services are two examples which demonstrate that where the requisite scale of operations and receptor capacity exists, NRC has been successful in commercialization. Also, while the impact on the language industry has been small, the impacts of NRC's R&D in the public sector and in domains outside of the language industry have been significant. Currently, projects in security and intelligence and health are most prominent. This is driven to some degree by NRC's revenue focus.

Finally, the evaluation found that NRC's researchers working in natural language processing are world-class and are a sought after partner for scientific collaboration. Evidence from

administrative data along with testimony from internal and external stakeholders suggests that resources are being used efficiently to respond to NRC's commitments under the roadmap.

6. RECOMMENDATIONS

In consideration of the evidence and findings of this evaluation, the following recommendation is put forth by NRC OAE.

Recommendation: *NRC should consider its fit within the education pillar of the Roadmap and seek to better articulate how its activities and expected outputs align with the Roadmap's results architecture.*

7. MANAGEMENT RESPONSE

Recommendation	Response and Planned Action(s)	Proposed Person(s) Responsible	Timelines	Measure(s) of Achievement
<p>NRC should consider its fit within the education pillar of the Roadmap and seek to better articulate how its activities and expected outputs align with the Roadmap's results architecture.</p>	<p>Agreed. NRC's work is well aligned with the ultimate outcome of the Roadmap. Although there is a link between NRC's work and the objectives of the Education pillar (relating to access to technological [language] tools), this link is less evident.</p> <p>NRC is currently involved in elaborating the federal government's new official languages plan (the successor to the Roadmap for Canada's Official Languages 2013-2018). As part of the development of this new plan, the VP Emerging Technologies – Platforms will ensure that NRC activities and outputs more clearly align throughout the results architecture.</p>	<p>VP Emerging Technologies - Platforms</p>	<p>September 30, 2017</p> <p>[Note: this timeline is tentative, as it is unknown at this time when the new official languages plan will be finalized]</p>	<p>In the new federal languages plan, NRC's contribution is clearly reflected, both in the plan's ultimate outcome(s) and in lower level (pillar) objectives.</p>

APPENDIX A: METHODOLOGY

The evaluation of NRC's initiative under the Roadmap was led by an independent evaluation team from the NRC Office of Audit and Evaluation (OAE). This appendix details the evaluation methodology used by the evaluation team and includes a discussion on the evaluation design and methods, and on challenges and limitations encountered.

Evaluation Design

The evaluation questions were developed following a review of key documents as well as consultations with ICT management and staff, and corporate staff involved with the initiative. The questions were also designed to meet the requirements of Treasury Board's *Policy on Evaluation* (2009).

Table 1: Evaluation questions

<i>1. Does NRC's initiative under the Roadmap respond to an identifiable industry need?</i>
<i>2. Are NRC activities and objectives under the Roadmap aligned to federal government priorities and NRC strategic outcomes?</i>
<i>3. Is NRC's role in delivering this initiative appropriate and conducive to strong results?</i>
<i>4. To what extent has the initiative contributed to the growth and competitiveness of the Canadian language industry and other Canadian industries through research and development?</i>
<i>5. To what extent has NRC demonstrated an efficient use of resources in responding to its commitments under the Roadmap?</i>

Evaluation Methods

In order to meet its commitments under the Roadmap, NRC conducted a calibrated evaluation of its funded initiative, narrowing the evaluation scope and employing a limited number of methods compared to traditional evaluations. Calibrated evaluations allow for the effective utilization of evaluation resources, while maintaining the credibility and usability of the evaluation results.

A risk assessment was conducted to determine the initiative's characteristics and risk profile, the information needs of NRC Senior Management and those of other stakeholders. Based on the low risk and low materiality of this initiative for NRC, it was concluded that a streamlined approach with a limited scope was appropriate for this evaluation. The methods used are discussed in more detail below.

- **Document Review:** Both internal and external documents were reviewed including NRC planning documents, performance reports, and background documents pertaining to the language industry.
- **Data Review:** Available corporate, portfolio and program administrative and performance data, from 2013 to the present, were reviewed.

- **Client / collaborator / stakeholder interviews:** 10 semi-structured interviews were conducted with representatives from client and partner organizations, industry stakeholders and independent experts on the language industry.
- **Internal interviews:** Semi-structured interviews were conducted with 4 internal staff.

Interviews with internal stakeholders were conducted mostly in person, while interviews with external stakeholders were primarily conducted by telephone. Interview guides were developed to align questions with the information requirements as identified in the evaluation framework and were provided to interviewees in advance of the discussion. This process ensured that the information requested would yield relevant information in support of evaluation judgement criteria and indicators. The majority of interviews were individual; however group interviews were conducted with Team Leads, for efficiency and to provide a richer discussion.

Limitations

The calibrated design of this evaluation resulted in a streamlined evaluation that is not as comprehensive as the previous evaluation. The methods and scope of this evaluation were adjusted in accordance with the generally acknowledged low risk of the program and its low materiality (\$2 million per year). This tailored method and scope served the purpose of providing a reasonable assessment of the initiative's relevance and performance within this context. However, it is acknowledged that this strategy was not without its limitations, which are listed below.

Review of selected documents

An extensive review of academic literature was not conducted. Where academic journal articles were reviewed, they were used to establish the relative performance of the initiative in international translation competitions. External documents were reviewed that established the context of the 'language industry' as well as the current and future market landscape for language technologies globally.

Analysis of selected administrative and performance data

An analysis of administrative and performance data was conducted in support of section 4.2 'Resource utilization'. The methodology is limited due to the fact that the initiative is conducted by two separate resource teams who do not code time toward the initiative. This makes a comprehensive analysis of efficiency problematic. However, this lack of administrative data was mitigated through interviews with external experts who were queried as to the effectiveness of NRC's initiative in consideration of associated budget and staff levels.

Key informant interviews

The number of key informants interviewed was limited to 14. This amount cannot be viewed as sufficient to represent the views of the broad array of stakeholders with interests in language technologies. However, the 14 interviewees were specifically selected to represent a cross section of groups, identified by the evaluation team in the planning stages.

APPENDIX B: SELECTION OF DOCUMENTS REVIEWED

Association de l'Industrie de la Langue/ Language Industry Association. (2012). Language Industry Survey, Final Report. Accessed online July 5 at: www.ailia.ca/dl270.

CBC News. (2016). Online translator helps federal workers 'do their job,' say defenders. Access July 7th at: <http://www.cbc.ca/news/canada/ottawa/online-translator-defended-1.3433919>

Choudhury, R., McConnell, B. (2013). TAUS: Translation Technology Landscape Report. Accessed online July 5 at: <https://www.taus.net/think-tank/reports/ja-sugite-xblog2>.

European Commission, Director-General for Translation. (2009). Study on the size of the language industry in the EU. Accessed online July 5 at: http://ec.europa.eu/dgs/translation/publications/studies/index_en.htm.

Government of Canada. (2015), Budget 2015. Accessed online July 5 at: <http://www.budget.gc.ca/2015/docs/plan/toc-tdm-eng.html>

Canadian Heritage. (2008). Summative Evaluation of the Action Plan for Official Languages Coordination Program. Accessed online July 5 at: http://www.pra.ca/resources/pages/files/Reports/Summative%20Evaluation%20of%20the%20Action%20Plan%20for%20Official%20Languages%20Coordination%20Program_E.pdf

Industry Canada. (2008). Final Evaluation of the Language Industry Initiative. Accessed online July 5 at: <https://www.ic.gc.ca/eic/site/ae-ve.nsf/eng/02856.html>

Kuhn, R., Stewart, D., and Islam, R. (2015). BLEU Comparison of Portage vs. Google Translate on Four Domains. Internal Document.

National Research Council. (2015). Machine Translation: Benefits and Advantages of Statistical Machine Translation and NRC's Portage. Accessed online July 5 at: <http://nparc.cisti-icist.nrc-cnrc.gc.ca/fra/voir/objet/?id=72d5a618-0cd3-4b68-b56f-41178020c9d0>

National Research Council. (2013). Multimedia Analytic Tools for Security, ICT Portfolio, Business Plan. Internal Document.

National Research Council. (2015). Corporate Key Performance Indicators. Internal Document.

National Research Council. (2015). ICT Five Year Strategic Plan: Fiscal Year 2015 to 2020. Internal Document.

National Research Council. (2015). Information and Communications Technologies, Annual Operating Plan, Fiscal Year 2015-16. Internal Document.

Price Waterhouse Cooper. (2012). Translation Bureau Benchmarking and Comparative Analysis. Final Report. Accessed online July 5 at: www.ailia.ca/dl359.

Simard, M., Kuhn, R. and Remillard, J. (n.d.). QUAFF: Pilot Experiment. Internal Document.

SLATOR. (2016). Slator State of Translation and Localization Demand 2016. Accessed online July 5 at: <https://slator.com/demand-drivers/show-me-the-business-report-on-the-state-of-translation-demand-201516/>