

Défense nationale

DEFENCE RESEARCH AND DEVELOPMENT CANADA (DRDC)

RECHERCHE ET DÉVELOPPEMENT POUR LA DÉFENSE CANADA (RDDC)



The Defence Research and Development Canada – Centre for Operational Research and Analysis Hackathon 2021

A comprehensive summary

Patrick W. Dooley
Patricia Moorhead
Josée van den Hoogen
DRDC – Centre for Operational Research and Analysis

Terms of Release: This document is approved for public release.

Defence Research and Development Canada

Reference Document
DRDC-RDDC-2022-D111
September 2022



IMPORTANT INFORMATIVE STATEMENTS

This document was reviewed for Controlled Goods by Defence Research and Development Canada (DRDC) using the Schedule to the Defence Production Act.

Disclaimer: This publication was prepared by Defence Research and Development Canada an agency of the Department of National Defence. The information contained in this publication has been derived and determined through best practice and adherence to the highest standards of responsible conduct of scientific research. This information is intended for the use of the Department of National Defence, the Canadian Armed Forces ("Canada") and Public Safety partners and, as permitted, may be shared with academia, industry, Canada's allies, and the public ("Third Parties"). Any use by, or any reliance on or decisions made based on this publication by Third Parties, are done at their own risk and responsibility. Canada does not assume any liability for any damages or losses which may arise from any use of, or reliance on, the publication.

Endorsement statement: This publication has been published by the Editorial Office of Defence Research and Development Canada, an agency of the Department of National Defence of Canada. Inquiries can be sent to: Publications.DRDC-RDDC@drdc-rddc.gc.ca.

[©] Her Majesty the Queen in Right of Canada (Department of National Defence), 2022

[©] Sa Majesté la Reine en droit du Canada (ministère de la Défense nationale), 2022

Abstract

To enhance its analytical culture, capabilities, cohesion, and morale, Defence Research and Development Canada (DRDC) – Centre for Operational Research and Analysis (CORA) conducted its inaugural hackathon from 23–25 November 2021. As the Hackathon Administrative Committee (HAC), we proposed, planned, and conducted the event. Overall, the hackathon provided a flexible, accessible, and fun environment wherein participants could be truly creative, learn, experiment, and collaborate, and yielded multiple products with substantial potential to improve the ways in which CORA conducts its annual Public Service Employee Survey (PSES) analyses for DRDC. Here, we summarize the hackathon's objectives, execution, outcomes, and participants' feedback, and provide recommendations for improvement if a similar event is held in the future.

DRDC-RDDC-2022-D111 i

Résumé

Afin d'améliorer sa culture, ses capacités, sa cohésion et son moral en matière d'analyse, Recherche et développement pour la défense Canada (RDDC) – Centre de recherche et d'analyse opérationnelle (CARO) a tenu un marathon de programmation inaugural du 23 au 25 novembre 2021. À titre de comité d'administration du marathon de programmation (CAMP), nous avons proposé, planifié et dirigé l'activité. Dans l'ensemble, le marathon de programmation a fourni un environnement flexible, accessible et amusant qui a permis aux participants d'être véritablement créatifs, d'apprendre, d'expérimenter et de collaborer, et a donné lieu à plusieurs produits ayant la possibilité considérable d'améliorer la façon dont le CARO effectue ses analyses annuelles du Sondage auprès des fonctionnaires fédéraux (SAFF) pour RDDC. Nous résumons ici les objectifs, l'exécution et les résultats du marathon de programmation, ainsi que les commentaires des participants. Nous formulons également des recommandations pour améliorer une future activité semblable, le cas échéant.

Table of contents

Abstract	. i
Résumé	. ii
Table of contents	iii
List of figures	iv
List of tables	. v
Acknowledgements	vi
1 The DRDC CORA Hackathon 2021	. 1
1.1 Motivation	
1.2 Objectives	. 1
1.3 Execution	. 2
1.4 Outcomes	
1.5 Participant feedback	
1.6 Overall assessment	
1.7 Recommendation	
References	
Annex A Timeline	. 7
Annex B Proposal	. 8
Annex C Invitation	11
Annex D Information document for potential participants	13
Annex E Participants by team	16
Annex F Orientation	18
Annex G Team appellations and awards	27
Annex H Verbal feedback from hackwash session	30
Annex I Satisfaction survey instrument	33
Annex J Satisfaction survey results	36
Annex K Abstracts for the formal presentations of teams' results	41
K.1 DAY 1—Wednesday 19 January 2022	41
K.2 DAY 2—Thursday 20 January 2022	41
K 3 DAY 3—Friday 21 January 2022	42

List of figures

iν

Figure E.1:	Virtual background (hackground) used by HAC members during the hackathon	17
Figure F.1:	Slide 1 of orientation brief	18
Figure F.2:	Slide 2 of orientation brief	18
Figure F.3:	Slide 3 of orientation brief	19
Figure F.4:	Slide 4 of orientation brief	19
Figure F.5:	Slide 5 of orientation brief	20
Figure F.6:	Slide 6 of orientation brief	20
Figure F.7:	Slide 7 of orientation brief	21
Figure F.8:	Slide 8 of orientation brief	21
Figure F.9:	Slide 9 of orientation brief	22
Figure F.10:	Slide 10 of orientation brief.	22
Figure F.11:	Slide 11 of orientation brief	23
Figure F.12:	Slide 12 of orientation brief	23
Figure F.13:	Slide 13 of orientation brief	24
Figure F.14:	Slide 14 of orientation brief	24
Figure F.15:	Slide 15 of orientation brief	25
Figure F.16:	Slide 16 of orientation brief	25
Figure F.17:	Slide 17 of orientation brief	26
Figure F.18:	Slide 18 of orientation brief	26
Figure G.1:	Opening slide for recognition ceremony.	27
Figure G.2:	Example of an award certificate	28
Figure H.1:	Slide displayed to motivate discussion during the hackwash session	30
Figure I.1:	Items 1–5 of the online satisfaction survey instrument	33
Figure I.2:	Items 6–8 of the online satisfaction survey instrument	34
Figure I.3:	Items 9–14 of the online satisfaction survey instrument	35
Figure J.1:	Histograms summarizing participants' responses to survey items 1–5	36
Figure J.2:	Histograms summarizing participants' responses to survey item 6	36
Figure J.3:	Histograms summarizing participants' responses to survey items 9–11	38

List of tables

Table G.1:	Summary of bestowed appellations and awards	29
Table H.1:	Condensed summary of successes and lessons observed from the hackwash	32

DRDC-RDDC-2022-D111 v

Acknowledgements

We gratefully acknowledge the hackathon's 14 participants (Annex E), whose enthusiastic participation and analytical talents underpinned the event's success.

We also extend our thanks to Charles Morrisey (Centre Director, DRDC CORA) for approving what was an unprecedented analytical activity within the centre.

1 The DRDC CORA Hackathon 2021

To enhance its analytical culture, capabilities, cohesion, and morale, Defence Research and Development Canada's Centre for Operational Research and Analysis (DRDC CORA) conducted its inaugural hackathon from 23–25 November 2021 (see Annex A for timeline). As the Hackathon Administrative Committee (HAC), we proposed (Annex B), planned, and conducted the event. Here, we summarize the hackathon's objectives, execution, outcomes, and participants' feedback. We also offer our overall assessment and recommendations for improvement should a similar event be held in the future.

1.1 Motivation

CORA's hackathon concept arose from our belief that a voluntary annual event—that could inspire, professionally develop, and be enjoyed by CORA's diverse analysts—could greatly benefit analysts individually as well as CORA and its clients generally. For instance, we sought to foster creativity and individual engagement by providing a largely unconstrained environment in which each analyst could exercise substantial autonomy, employ/enhance their expertise, and work on a topic of importance to them. By allowing analysts to participate in small teams, we aimed to provide a unique professional development opportunity that was collective, realistic (i.e., "train-as-we-fight"), and potentially multidisciplinary. By teaming analysts with unfamiliar colleagues, we aimed to build individual rapport and Centre cohesion, enable the transfer of knowledge and skills, provide informal mentoring opportunities, and potentially forge enduring linkages between colleagues. Ultimately, we hoped that analysts' hackathon experiences would foster a vibrant analytical culture within CORA and deliver decision support to CORA's clients more collaboratively and effectively.

To provide a degree of cohesion, we specified an overarching hackathon theme that linked teams' efforts, namely, the Public Service Employee Survey (PSES). For several years, CORA has conducted formal, annual analyses of DRDC's PSES results. The PSES-themed hackathon represented excellent means for crowd-sourcing new analytical methods and insights that could be directly exploited in future analyses.

While hackathons are traditionally focused on computer programming, we designed the CORA hackathon to be unique, with the breadth of CORA employee expertise in mind, from strategic analysis to hard-core coding—and everything in between. More specifically, while the PSES theme was chosen for the hackathon, participants defined their own challenge questions and utilized analysis methods and tools of their own choosing.

1.2 Objectives

We designed the hackathon with the following objectives in mind (Annex B):

- to foster unbridled creativity and a vibrant analytical culture within CORA via collaborative and multidisciplinary efforts;
- to promote cohesion and individual morale within CORA via an intellectually stimulating and enjoyable team-building exercise featuring a fun atmosphere and friendly competition;
- to provide a unique, no-cost professional development opportunity to all CORA staff;

-

¹ The term "hackathon" (a portmanteau of "hacking marathon") is "an event in which computer programmers collaborate intensively with one another and sometimes with people in other specialties over a relatively short period of time to create code usually for a new software product or service." [1]

- to crowd source new analytical concepts, methods, and/or tools that might enhance CORA's annual PSES analysis efforts;
- to exploit PSES data sets that CORA's formal analyses have yet to explore; and
- to generate potentially valuable insights regarding DRDC and other organizations.

1.3 Execution

All CORA staff, irrespective of their expertise, were invited to participate (Annex C) in teams of one to three persons (the HAC provided matchmaking services upon request; Annex D). The HAC also individually encouraged dozens of CORA staff to participate, especially new members. Ultimately, in addition to the three-person HAC, 14 CORA staff (including two students) participated as five teams (Annex E). When their hackathon-related duties allowed, two HAC members participated separately as part-time, one-person teams (Annex E). All five of CORA's sections were represented, though in a very unbalanced way (i.e., including the HAC, there were 9 MAS, 3 MEIA, 2 JSA, 2 LOC, and 1 JTS participants). Of the five multi-person teams, three were composed of members from different sections.

Teams were free to work on any PSES-related topic using any data sets, methods, and tools that they wished (Annex D). To spur teams' imaginations, the HAC suggested 12 potential topics. The HAC also provided numerous pre-processed PSES data sets, data-processing scripts, and explanatory documents to make the event more accessible to all.² To that end, teams were required to share their own bespoke data sets with everyone.

Collaboration across CORA was made possible by the use of online collaboration tools, in particular Microsoft Office 365. Online collaboration was the only means by which dispersed CORA analysts could successfully work.³ The HAC provided the following virtual infrastructure for hackathon purposes:

- a persistent, hub-and-spoke Microsoft Teams video teleconference consisting of a virtual plenary room and team-specific breakout rooms plus
- manually mirrored Microsoft OneDrive and Google Drive repositories for file sharing.

Teams were allotted three full days (Annex A and Annex D) to pursue their ideas.⁴ At the outset, the HAC delivered an orientation briefing (Annex F) and fielded questions in the plenary room.⁵ For at least 10 hours each day, one or more HAC members staffed the plenary room to address teams' questions and problems. Each team was free to collaborate in its HAC-provided breakout room or via other means.

2

² All hackathon-related materials are archived on CORA's DWAN shared drive at: O:\New - Defence Sciences\Shared Work\CORA Hackathon - Op ANALYSIS\2021 CORA Hackathon.

³ Potential barriers to participation mitigated by using an online approach included, but were not limited to: time zone differences, travel, meeting room availability, video teleconferencing (VTC) room access, hardware/software access, etc.

⁴ The HAC proposed—and DRDC CORA's centre director approved—the three-day duration believing that participants would be unencumbered by their usual work. In practice, many participants devoted much less than three-person days, due to align the appropriate access and access than three-person days, due to align the appropriate access than three-person days.

would be unencumbered by their usual work. In practice, many participants devoted much less than three-person days, due to client commitments, meetings, etc. Such obligations also precluded many staff from registering at all. A few staff, in anticipation of such disruptions, devoted some time ahead of the event—flexing their hours to enable a full three-day equivalent effort.

⁵ A special orientation was held the day prior, to accommodate British Columbia-based participants.

1.4 Outcomes

Each team briefly presented its results to other participants during a 26 November 2021 plenary session. From 19–21 January 2022, each team was afforded and accepted a longer opportunity to present their results to all interested CORA staff and managers (see Annex K for abstracts). Each briefing was attended by approximately 50 DRDC staff members. The seven teams' efforts spanned a broad range of topics that included:

- automated means for cleaning, analyzing, and visualizing PSES data sets;
- a correlation-based exploration of alternative scales for eliciting PSES-type survey responses;
- additional metrics for PSES analyses;
- methods and means for identifying and visualizing correlated PSES results;
- methods and means for visualizing an organization's overall annual performance "at a glance";
- methods and means for analyzing and visualizing an organization's performance over time;
- analyses based on employee classifications; and
- enhanced online dashboards to enable non-analysts to explore PSES results on their own.

Each hackathon product was archived on CORA's DWAN shared drive and will be assessed for potential future use by CORA's standing PSES analysis team.

The HAC held a light-hearted, informal awards session after teams presented their results (Annex G). The session was meant to provide collegial recognition and reinforce the event's team-building aspect.

Immediately thereafter, the HAC facilitated a hackathon hotwash (i.e., "hackwash") session to elicit participants' views on their individual experiences, the 2021 event itself, and a potential future hackathon. Since the resulting verbal feedback was provided during a plenary session (and was therefore attributable), it might exhibit positive bias. Nevertheless, we provide the consolidated notes taken by the HAC during the hackwash in Annex H.

During the hackwash, the HAC also launched an online participant satisfaction survey (Annex I). The survey was voluntary, anonymous, and collected feedback for three business days. The survey's response rate was relatively high (i.e., 71%; 10 of 14 participants responded). Its results are compiled in Annex J and are less likely to exhibit positive bias than those in Annex H, since the survey was anonymous.

1.5 Participant feedback

Based on the participant feedback received verbally during the hackwash and via the survey, we offer the following observations:

- All survey respondents reported that their overall hackathon experience was of high or highest quality (50% reported having a highest-quality experience). Two participants volunteered to join the HAC, were a second hackathon to be approved.
- The consensus was that the hackathon was most valuable in an interpersonal sense. All participants rated its interpersonal value as high- or highest-value (rated highest-value by 50%).
- The hackathon reportedly yielded substantial value in a personal sense. Its personal value was characterized as high or highest by 80% of respondents (rated highest-value by 30%).
- The hackathon also reportedly provided substantial value in a professional sense. Its professional value was deemed high- or highest-value by 70% of respondents (rated highest-value by 30%).

- The hackathon reportedly yielded substantial value to CORA and DRDC. Its reported value varied from mid-value (40% of respondents) through high-value (30%) to highest-value (30%).
- Respondents reportedly opted to participate primarily on their own initiative (60%) because they desired a break from their usual work (80%) that would enable them to be creative (80%), learn (80%), and explore science or analysis (70%) in a fun setting (70%). Roughly half of respondents reportedly participated in order to work with new collaborators (60%), for team-building purposes (50%), and to make a useful contribution (50%), even though the PSES theme was of minority interest (40%). Relatively few respondents reported being motivated to participate by technical reasons like improving their analytical skills (30%) or testing an idea, method, or tool (40%).
- Respondents' verbal and free-text feedback indicated that the hackathon generally met reported expectations, especially for those seeking an unconventional and fun analytical environment that was conducive to creativity, learning, collaboration, and multidisciplinary opportunities. Key enablers reportedly included:
 - the hackathon's extreme flexibility in terms of problem selection and schedule;
 - the high degree of accessibility afforded by the pre-processed data sets and aids provided by the HAC;
 - the quality of planning, instructions, and infrastructure provided by the HAC; and
 - the HAC's efforts to form teams that were balanced in terms of individuals' general expertise, technical abilities, length of public service, and availability over the three days.
- Multiple respondents' verbal and free-text remarks identified the following aspects for potential improvement.
 - Several respondents' efforts were interrupted by work unrelated to the hackathon. Such interruptions substantially reduced the time and concentration devoted to the hackathon and impeded team collaboration. Respondents suggested that increasing the hackathon's duration from three days to five and/or holding it during a different part of the year might have mitigated such interruptions.
 - Several respondents reported that the 26 November plenary session (that included the initial presentation of results, awards session, and hackwash; Annex A) was too brief and therefore rushed. Multiple respondents stated that each team's initial 10-minute opportunity to present their results to other participants was insufficient.⁶
 - Some respondents wished that more "soft scientists" (presumably, social scientists and strategic analysts) had participated, to increase the number of multidisciplinary experiences afforded by the hackathon.⁷
 - At least one respondent desired that some of the hackathon's products would be integrated into CORA's PSES tool set. One respondent noted that open-ended problems could yield opportunities for post-hackathon collaboration.
- Nearly all respondents (90%) stated that CORA should "definitely" conduct another hackathon and two participants volunteered to join its HAC. Were a second hackathon to occur, nearly all respondents (80%) would reportedly participate (60% "definitely," 20% "probably") though 70% of respondents indicated that they would prefer a different theme (30% "definitely," 40% "probably"). Suggested topics included:
 - those proposed by or relevant to a DND/CAF partner/client (e.g., ship detection, sensor data, etc.);
 - the COVID-19 pandemic;

6

⁶ Such dissatisfaction occurred, at least in part, because the HAC did not reiterate that teams would have a second opportunity to present their results at greater length during a future video teleconference, to which all CORA staff and managers would be invited.

⁷ The HAC's 2021 efforts to recruit social scientists, psychologists, and strategic analysts were largely unsuccessful.

- the environment (e.g., climate change, weather, or natural hazards); and
- open-source data (e.g., geographic information, social media data, published Government of Canada data, etc.).

1.6 Overall assessment

CORA's inaugural hackathon was well attended, ran smoothly, and yielded multiple products that might enhance CORA's future PSES analyses. Strongly positive feedback indicated that the hackathon provided a flexible, accessible, and fun environment wherein participants could be truly creative, learn, experiment, and collaborate (sometimes in cross-disciplinary ways with unfamiliar colleagues). Such outcomes were consistent with both the hackathon's objectives and participants' reported aspirations. The hackathon also yielded useful suggestions for future improvement.

Moreover, the hackathon yielded multiple products (e.g., for data cleaning, metrics, and visualization) with substantial potential to improve the ways in which CORA conducts its annual PSES analyses for DRDC. Indeed, one hackathon product (from Team Moirai) was conceptually consistent with a CORA proposal for a new PSES visualization tool, the PSES Dashboard for Accessible and Transparent Analysis (PSES DATA). Development of PSES DATA was subsequently selected for funding under DRDC's Internal Innovation Program (IIP) in June 2022.

1.7 Recommendation

Given the strong apparent consensus between the HAC and participants that the hackathon's objectives were broadly met, we offer the following recommendation. That is, we recommend that CORA conduct a follow-on hackathon in 2022, whose objectives and delivery would be largely the same, but would involve a different topic, a five-day duration, a time of year that is less busy, and an explicit acknowledgement that part-time participation is welcome.

DRDC-RDDC-2022-D111 5

References

[1] Hackathon (online), Merriam-Webster, Inc., https://www.merriam-webster.com/dictionary/hackathon (Access date: 25 Aug. 2022).

Annex A Timeline

20 September 2021	Proposal submitted to CD CORA for consideration.
29 September 2021	Approval granted by CD CORA.
29 October 2021	Announcement and invitation emailed to all CORA staff.
19 November 2021	Registration deadline.
22 November 2021	Special orientation briefing for British Columbia-based participants.
23 November 2021	Main orientation briefing.
23–25 November 2021	Team collaboration period.
25 November 2021	Informal briefings of teams' results to participants. Appellations and awards ceremony. Hackwash session for collecting verbal feedback.
25–30 November 2021	Data collection period for anonymous online participant satisfaction survey.
07 December 2021	Summary report delivered to CD CORA.
19–21 January 2022	Formal briefings of teams' results in all-CORA fora.

Annex B Proposal

The remainder of this annex contains the 20 September 2021 hackathon proposal that was approved by DRDC CORA's centre director on 29 September 2021.

Proposal for a PSES-Based CORA Hackathon

P. Moorhead, J. van den Hoogen, and P. Dooley 20 September 2021

Context

CORA scientists have analyzed DRDC's PSES data for the past four editions of the survey (2017–2020), with the goals of identifying concerns in need of remediation, and successes worthy of emulation. The analyses have examined DRDC as a whole, and more importantly, its individual L2 and L3 organizations.

Various analysis approaches and tools have been developed over the years, primarily: VOICES for visualization of the PSES data; and OUTER for identification of "extreme" results within DRDC relative to a comparative population (e.g., the broader public service). Exploration of multi-year analysis approaches began in 2020.

The TBS publishes multiple PSES data subsets each year, which are organized by various organizational, geographic, and demographic factors. Since the CORA PSES team's analyses to date have utilized only one of the subsets, there remain unexplored opportunities to glean further insights.

Objectives

The hackathon would have four objectives:

- To crowd source new analytical concepts, methodologies and tools, especially to support multi-year analyses of the PSES data. Analysis methods can be quantitative and/or qualitative, derived from operational research and/or strategic analysis expertise. Both successes and failures are of interest—we can learn as much (if not more) from approaches that did not "work" as from those that did. Valuable new methods and tools would be incorporated into the CORA PSES team's tool-box for application in future years. Participants would be encouraged to consider formal publication of their analysis approaches, where applicable.
- To exploit PSES data subsets that have not been utilized by the CORA PSES team to date.
- To generate new insights into the PSES data and DRDC's results.
- To promote a vibrant analytical culture, collaboration, and creativity within CORA. The hackathon is simultaneously intended to be: an enjoyable team-building exercise that encourages cross-team and interdisciplinary collaboration; a stimulating, no-holds-barred outlet for analytical creativity; a no-cost professional development opportunity; and a friendly competition. The hackathon should boost both morale and cohesion, which are especially important in the current work-from-home context.

8 DRDC-RDDC-2022-D111

Challenge questions

The following challenge questions are examples, meant to stimulate participants' imaginations. Participants are welcome to select one (or more) of these, or develop their own challenge question(s).

- How can we (semi-)automate the mapping of identical questions (or identical for practical purposes) from one PSES edition to another in order to build a structured multi-year data file/database of responses? TBS publishes an annual mapping of the current year's questions to those of the previous edition which could be used as a starting point.
- How can we (semi-)automate the mapping of organizational hierarchies from one PSES edition to another? The organizational hierarchy in each edition of the PSES is taken from Departmental systems of record which may not be up-to-date, and the hierarchies can change from one year to the next due to internal re-organizations.
- What type/example of data structure(s) would be good for storing multi-year PSES data in a format amenable to analysis? The ability to analyze the data over time by individual questions, by themes or sub-themes, as well as across multiple organizational, geographic, and demographic factors is desired.
- What types of errors or inconsistencies exist in PSES data sets, within a given year or from one year to another? Additionally, what techniques or tools to automate data cleaning can be utilized or developed?
- What is the best way to identify and/or visualize correlations in a given PSES edition or across multiple years?
- The CORA PSES team developed the OUTER methodology for comparing DRDC and its subordinate organizations to peer groups in the broader public service. What other approaches could be used to compare DRDC's performance with that of its peers?
- What constitutes a "peer group" for DRDC and its subordinate organizations in the federal public service? What criteria could be used to identify peer organizations against which to compare DRDC's PSES results? Additionally, how would you apply those criteria to the PSES data subsets, given not all organizations in a defined peer group may be identifiable in the PSES datasets?
- The CORA PSES team's analyses do not apply weights to the PSES questions, sub-themes or themes. What weighting schemes could be developed, what is their rationale, and how could they be applied to the PSES data to provide new insights?
- While the PSES questions are organized along fixed themes and sub-themes, these are not the only possible lenses through which the data can be viewed. What other thematic lenses are valuable for interpreting PSES results or for use as indicators (e.g., organizational trust)? Which questions would you include in your proposed theme(s), why, and would you weight them equally or differently?
- What is the best way to visualize the data within a theme containing multiple questions, for one PSES edition or its evolution over multiple years?
- What is the best way to visualize organizational, geographic, and/or demographic differences in results for one PSES edition or their evolution over multiple years?
- Are there reputable bodies of external literature or data that would be beneficial to use in conjunction with PSES data to obtain valuable insights? How should they be used together and what misuses should be avoided?

Data

Each year, TBS publishes multiple data subsets organized by various organizational, geographic, and demographic factors. The PSES datasets from 2017 to 2020 can be downloaded from the TBS PSES website or copied from a CORA shared drive folder (links to be provided); participants are encouraged to obtain their own local copies of the data sets in advance of the hackathon. Participants may use whichever data subset(s) they wish, but are encouraged to utilize as many of the data subsets as they can.

Outputs

Each team would provide a summary of what they develop to the hackathon organizers in a suitable form, e.g., verbal report, slide deck, Word document, code, etc. The outputs would be reviewed by CORA's PSES team for possible incorporation or adaptation to future PSES analyses. Each team would be invited to present their work during a post-hackathon hotwash session open to all CORA staff and, possibly, at a CORA Seminar.

Ground rules

CORA staff can participate as individuals or members of teams.

The CORA PSES team will provide descriptions of the published PSES data subsets to all participants, as the TBS website lacks easy-to-understand descriptions. Some "here be dragons" lessons learned will also be shared.

Participants can use whatever analysis methods and tools they wish. There are no constraints in this regard. The Jupyter (python) scripts and decode tables currently used by the CORA PSES team for data processing will be made available to all participants.

The only organizations that should be identifiable in the analysis results are DRDC and its subordinate organizations (e.g., Level-2 organizations like DGRDSE, Level-3 organizations like CORA). When other Level-0 (i.e., other departments or agencies) or Level-1 organizations (e.g., ADM(HR-Civ), ADM(Pol), etc.) are analyzed, their identities should be anonymized.

Timeline

26–28 Oct 2021	Hackathon announced at the 2021 CORA AGM
12 Nov 2021	Deadline for participant/team registration
15–26 Nov 2021	A 3-day period in these 2 weeks will be selected for the hackathon, followed by a "hotwash" session on a 4 th day
TBD	Participants present their results at a CORA Seminar

Hackathon organizers: Patricia Moorhead, Josée van den Hoogen, and Patrick Dooley

Annex C Invitation

The remainder of this annex contains the hackathon invitation sent to all CORA staff via email on 29 October 2021. Included as an attachment was the hackathon information document contained in the subsequent annex. The invitation and its attachment were sent to all CORA staff again on 15 November 2021, as a reminder.

SENT ON BEHALF OF THE CORA HACKATHON ADMINISTRATIVE COMMITTEE // ENVOYÉ AU NOM DU COMITÉ ADMINISTRATIF DU HACKATHON CARO

(La version française suit)

CORA's first-ever Hackathon will take place 23–25 November, using publicly available Public Service Employee Survey (PSES) data sets. The Hackathon will be a no-holds-barred, friendly competition and outlet for analytical creativity.

The Hackathon is open to all CORA staff, regardless of background or expertise. We want to apply CORA's full skill set to the data. You can use quantitative or qualitative methods from any field. Programming expertise is not required to participate.

You can register alone or as a team of two or three people. If you want to collaborate but don't have a team, we will help you find a teammate or two. Instead of prizes, bragging rights and honourary titles are up for grabs.

The registration deadline is 19 November 2021. To register, send an email to any member of the organizing committee with the names of your team members (if any) and the challenge question(s) you are planning to tackle (if you know beforehand).

Additional information can be found in the attached document. If you have any questions, please do not hesitate to contact any of us.

Hoping to see you at the Hackathon!

Le premier Hackathon du CARO aura lieu du 23 au 25 novembre et utilisera les ensembles de données du Sondage Auprès des Fonctionnaires Fédéraux (SAFF) qui est accessibles au public. Le Hackathon sera une compétition amicale sans restriction et une occasion pour la démonstration de la créativité analytique.

Le Hackathon est ouvert à tout le personnel du CARO, quelle que soit leur expérience ou leur expertise; nous voulons appliquer toutes les compétences de CARO aux données. Vous pouvez utiliser des méthodes quantitatives ou qualitatives de n'importe quel domaine. Une expertise en programmation informatique n'est pas requise pour participer.

Vous pouvez vous inscrire seul ou en équipe de deux ou trois personnes. Si vous souhaitez participer, mais que vous n'avez pas d'équipe, nous vous aiderons à trouver des coéquipiers. Au lieu de prix, des droits de vantardise et des titres honorifiques sont à gagner.

La date limite d'inscription est le 19 novembre 2021. Pour vous inscrire, envoyez un courriel à n'importe quel membre du comité d'organisation avec les noms des membres de votre équipe (s'il en est) et la/les questions du défi que vous envisagez de relever (si vous les connaissez à l'avance).

Informations supplémentaires sont disponibles dans le document attaché (en anglais seulement). Si vous avez des questions, n'hésitez pas à nous contacter.

En espérant vous voir au Hackathon!

Patricia Moorhead, Josée van den Hoogen, and Patrick Dooley

CORA Hackathon Administrative Committee (CORA HAC) / Comité Administratif du Hackathon CARO

Annex D Information document for potential participants

The remainder of this annex was distributed as an attachment to the HAC's 15 November 2021 invitation email to all CORA staff.

The Inaugural CORA Hackathon

23-25 November 2021

What: The CORA Hackathon is a no-holds-barred, friendly competition and outlet for analytical creativity. The subject of the inaugural hackathon is the Public Service Employee Survey (PSES) and will utilize the data sets publicly released by the Treasury Board Secretariat (TBS). The PSES data sets are big and messy, providing lots of opportunity to showcase OR&A at its best. At the end of this document is a list of suggested "Challenge Questions"; however, participants are welcome to develop their own.

The CORA Hackathon is endorsed by CORA management, so all participants will have the approval of their Section Head to set aside their regular program of work for three days to focus on this team building and professional development opportunity.

When:

28 Oct 2021	Hackathon announced at the 2021 CORA AGM
19 Nov 2021	Deadline for participant/team registration
23–25 Nov 2021	Hackathon is on!
26 Nov 2021	Post-hackathon 'hotwash' at 1300-1500 EST (TBC)
TBD	Participants present their results at a CORA Seminar

Where: The hackathon is a virtual (on-line) event.

Why: There are four objectives to the hackathon.

- 1. To promote a vibrant analytical culture, collaboration, and creativity within CORA. The hackathon is simultaneously intended to be: an enjoyable team-building exercise that encourages cross-team and interdisciplinary collaboration; a stimulating, no-holds-barred outlet for analytical creativity; and a friendly competition.
- 2. To crowd source new analytical concepts, methodologies and tools, especially to support multi-year analyses of the PSES data. Analysis methods can be quantitative and/or qualitative, derived from operational research and/or strategic analysis expertise. Both successes and failures are of interest—we can learn as much (if not more) from approaches that did not "work" as from those that did.
 - a. Valuable new methods and tools will be incorporated into the CORA PSES team's toolbox for application in future years.
 - b. Participants will be encouraged to consider formal publication of their analysis approaches, where applicable.
- 3. To generate new insights into the PSES data and DRDC's results.
- 4. For bragging rights—this is a competition after all. Fun and irreverent honorary titles will be bestowed on teams that excel based on yet-to-be determined criteria.

DRDC-RDDC-2022-D111 13

Who: The hackathon is open to all CORA employees, regardless of background or expertise. Participants are encouraged to register in teams of up to three (3) people. You can form your own team, or register individually and the hackathon organizers will facilitate your team membership.

Participants can use whatever analysis methods and tools they wish. There are no constraints in this regard. You do NOT need to have programming expertise (e.g., R, Python) to participate in the hackathon. That is the point of creating multi-disciplinary teams—each person brings something different and important to the table.

Many of the PSES data sets are too large to open in MS Excel, so other software (e.g., R, Python) is needed to extract data subsets from these files. If necessary, the hackathon organizers are available to assist with this process. You may also bribe members of competing teams to assist you if needed.

How:

- **Registration.** Teams/participants can register by e-mailing one of the hackathon organizers (Patricia Moorhead, Josée van den Hoogen, Patrick Dooley) on DWAN or O365. The registration deadline is Friday 19 November 2021.
- Data access. Participants will be granted access to a Google Drive site from which all data sets and related files can be accessed and downloaded. If desired, participants can download the PSES data sets directly from the TBS PSES website (file descriptions and links will be provided). Participants may use whichever data subset(s) they wish but are encouraged to utilize as many of the data subsets as they can. Teams will be required to upload any bespoke data files they extract/create to the Google Drive while the hackathon is in progress.
- Execution. The hackathon runs from 23 to 25 November, inclusive. Teams are welcome to start/end at whatever time of day they wish in their local time zone. Two information/orientation sessions will be held on 23 November at 0900 and 1200 EST (1000 and 1300 AST, 0800 and 1100 CST, 0700 and 1000 MST, 0600 and 0900 PST).
- **Post-hackathon hotwash.** A hotwash session for all participants, and open to all CORA employees, will be held on Friday 26 November, from 1300 to 1500 EST (TBC). Each team will give a short presentation on their work, lessons observed/learned will be shared, and the honorary titles bestowed.
- Outputs. Each team will provide a summary of what they developed to the hackathon organizers in a suitable format(s), e.g., verbal report, slide deck, Word document, code, etc. The outputs will be reviewed by CORA's PSES team for possible incorporation or adaptation to future PSES analyses. Teams may be invited to present their work, in greater detail, at a future CORA Seminar. Participants will be encouraged to consider formal publication of their analysis approaches, where applicable.

Suggested Challenge questions

The following challenge questions are examples, meant to stimulate participants' imaginations. Participants are welcome to select one (or more) of these or develop their own challenge question(s). Upon selecting a challenge question(s), participating teams shall inform the organizers of their selection.

- How can we (semi-)automate the mapping of identical questions (or identical for practical purposes) from
 one PSES edition to another in order to build a structured multi-year data file/database of responses? TBS
 publishes an annual mapping of the current year's questions to those of the previous edition which could be
 used as a starting point.
- How can we (semi-)automate the mapping of organizational hierarchies from one PSES edition to another? The organizational hierarchy in each edition of the PSES is taken from Departmental systems of record which may not be up-to-date, and the hierarchies can change from one year to the next due to internal re-organizations.

14 DRDC-RDDC-2022-D111

- What type/example of data structure(s) would be good for storing multi-year PSES data in a format amenable to analysis? The ability to analyze the data over time by individual questions, by themes or sub-themes, as well as across multiple organizational, geographic, and demographic factors is desired.
- What types of errors or inconsistencies exist in PSES data sets, within a given year or from one year to another? Additionally, what techniques or tools to automate data cleaning can be utilized or developed?
- What is the best way to identify and/or visualize correlations in a given PSES edition or across multiple years?
- The CORA PSES team developed the OUTER methodology for comparing DRDC and its subordinate organizations to peer groups in the broader public service. What other approaches could be used to compare DRDC's performance with that of its peers?
- What constitutes a "peer group" for DRDC and its subordinate organizations in the federal public service?
 What criteria could be used to identify peer organizations against which to compare DRDC's PSES results?
 Additionally, how would you apply those criteria to the PSES data subsets, given not all organizations in a defined peer group may be identifiable in the PSES datasets?
- The CORA PSES team's analyses do not apply weights to the PSES questions, sub-themes or themes. What weighting schemes could be developed, what is their rationale, and how could they be applied to the PSES data to provide new insights?
- While the PSES questions are organized along themes and sub-themes, these are not the only possible lenses through which the data can be viewed. What other thematic lenses are valuable for interpreting PSES results or for use as indicators (e.g., organizational trust)? Which questions would you include in your proposed theme(s), why, and would you weight them equally or differently?
- What is the best way to visualize the data within a theme containing multiple questions, for one PSES edition or its evolution over multiple years?
- What is the best way to visualize organizational, geographic, and/or demographic differences in results for one PSES edition or their evolution over multiple years?
- Are there reputable bodies of external literature or data that would be beneficial to use in conjunction with PSES data to obtain valuable insights? How should they be used together and what misuses should be avoided?

CORA Hackathon Administrative Committee (CORA HAC):

Patricia Moorhead, Josée van den Hoogen, and Patrick Dooley

Annex E Participants by team

Fourteen CORA staff members participated in the hackathon in teams of two or three. Teams were assembled either by (a) the team members themselves or (b) by the HAC from staff members who registered individually and wished to join a team. The team compositions were as follows.

Team APPROACH (self-assembled)

- Michèle Fee (Maritime & Air Section)
- Victor Isaac (Maritime & Air Section)
- Bahaa Khaddaj (Maritime & Air Section)

Team CORA Jones & the Raiders of the Very Messy Survey Data (HAC-assembled)

- Maude Amyot-Bourgeois (Land & Operational Command)
- Lachlan Gray (Materiel, Infrastructure, Economics & Analytics)
- Matthew MacLeod (Land & Operational Command)

Team Hacksta's Paradise (self-assembled)

- Jillian Henderson (Maritime & Air Section)
- Stephen Okazawa (Maritime & Air Section)

Team Medieval Cats (HAC-assembled)

- Sarah Babbitt (Joint Targeting Section)
- Andrea Lane (Joint & Strategic Analysis)
- Stephen Weber (Materiel, Infrastructure, Economics & Analytics)

Team Moirai (self-assembled)

- Lise Arsenault (Materiel, Infrastructure, Economics & Analytics)
- Cheryl Eisler (Maritime & Air Section)
- Mikayla Holmes (Maritime & Air Section)

Two HAC members also participated individually on a part-time basis, when their hackathon-related duties allowed, namely:

Team Charm

• Josée van den Hoogen (HAC member; Maritime & Air Section)

Team Strange

• Patrick Dooley (HAC member; Joint & Strategic Analysis)

Though all five CORA sections were represented, their participation was unbalanced (9 MAS, 3 MEIA, 2 JSA, 2 LOC, and 1 JTS).

During the hackathon, HAC members used a virtual background (i.e., "hackground"; Figure E.1) designed by Team Strange to make themselves instantly recognizable to participants.



Figure E.1: Virtual background (hackground) used by HAC members during the hackathon.

Annex F Orientation

The HAC oriented all participants at the hackathon's outset using the slides depicted in Figure F.1–Figure F.18. The HAC addressed participants' questions during the briefing and afterward.



Figure F.1: Slide 1 of orientation brief.

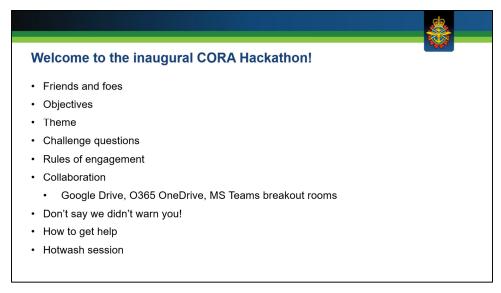


Figure F.2: Slide 2 of orientation brief.

18 DRDC-RDDC-2022-D111

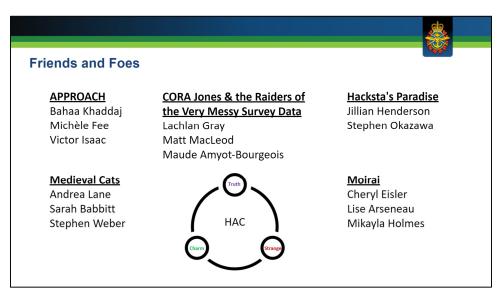


Figure F.3: Slide 3 of orientation brief.



- Encourage cross-team and interdisciplinary collaboration
- · Stimulate analytical creativity
- · Crowd source new analytical concepts, methods and tools (quantitative and/or qualitative)
 - Valuable new methods and tools will be incorporated into the CORA PSES team's toolbox for application in future years.
- · Generate new insights into the PSES data and DRDC's results
- Bragging rights this is a friendly competition after all
- HAVE FUN!!!!

Figure F.4: Slide 4 of orientation brief.



Theme: the Public Service Employee Survey (PSES)

- · Publicly released PSES data from 2017 to 2020
 - 5 subset files per year from 2017 to 2019, and 7 subset files in 2020
 - · Most subset files are too big to open in MS Excel
- We've created descriptions of the subset files for you, and we're sharing the Jupyter Notebook scripts we use for extracting data
- CORA's PSES analysis team has focused on just one data subset which provides breakdowns by organization structure (Subset 5 in 2017-2019, Subset 7 in 2020)
- Uncharted territory: the other four (six) Subsets, which provide breakdowns by various demographic variables at the Department level

Figure F.5: Slide 5 of orientation brief.



Possible challenge questions (1 of 2)

- What types of errors or inconsistencies exist in PSES data sets? Additionally, what techniques or tools to automate data cleaning can be utilized or developed?
- How can we (semi-)automate the mapping of questions from one PSES edition to another in order to build a structured multi-year data file/database of responses?
- How can we (semi-)automate the mapping of organizational hierarchies from one PSES edition to another?
- What type/example of data structure(s) would be good for storing multi-year PSES data in a format amenable to analysis? The ability to analyze the data over time by individual questions, by themes or sub-themes, as well as across multiple organizational, geographic, and demographic factors is desired.
- What approaches, other than OUTER, could be used to compare DRDC's performance with that of its peers?
- What criteria could be used to identify peer organizations against which to compare DRDC's PSES
 results? Additionally, how would you apply those criteria to the PSES data subsets, given not all
 organizations in a defined peer group may be identifiable in the PSES datasets?

Figure F.6: Slide 6 of orientation brief.

20 DRDC-RDDC-2022-D111



Possible challenge questions (2 of 2)

- What weighting schemes could be developed for PSES questions, sub-themes or themes, and how could they be applied to the PSES data to provide new insights?
- What other thematic lenses are valuable for interpreting PSES results or for use as indicators (e.g., organizational trust)? Which questions would you include in your proposed theme(s), why, and how would you weight them?
- What is the best way to identify and/or visualize correlations in a given PSES edition or across multiple years?
- What is the best way to visualize the data within a theme containing multiple questions, for one PSES edition or its evolution over multiple years?
- What is the best way to visualize organizational, geographic, and/or demographic differences in results for one PSES edition or their evolution over multiple years?
- Are there reputable bodies of external literature or data that would be beneficial to use in conjunction with PSES data to obtain valuable insights? How should they be used together and what misuses should be avoided?

Figure F.7: Slide 7 of orientation brief.



Rules of engagement

- 1. Start/end time: 23-25 Nov, your time zone (there is no stopwatch)
- 2. Use whichever PSES data subset(s) you want
- 3. Tackle whatever problem you want
- 4. Share all input files (e.g., data, code) to help others get started
- 5. Retain all outputs for archiving
- 6. In your analysis results, the only organizations that should be identifiable are DRDC and its subordinate organizations*
- 7. Brief your results at the hotwash session (Friday afternoon TBC)
- 8. Help each other. It's about the journey, not the destination.
- 9. Bribery is neither condoned nor condemned

Figure F.8: Slide 8 of orientation brief.

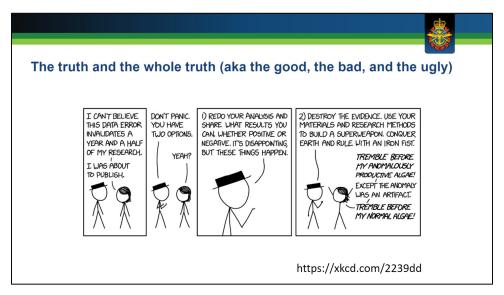


Figure F.9: Slide 9 of orientation brief.

Cartoon reproduced with permission of xkcd.com at https://xkcd.com/about/; the original can be found at https://xkcd.com/comics/data_error.png.

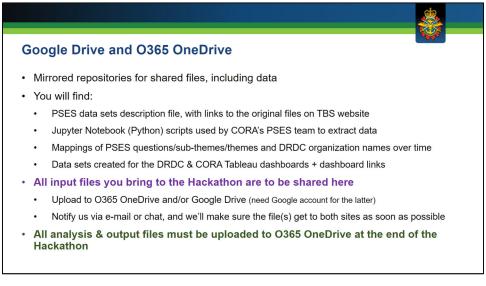


Figure F.10: Slide 10 of orientation brief.



Useful links

- TBS PSES website
 - https://www.canada.ca/en/treasury-board-secretariat/services/innovation/public-service-employee-survey.html
- · Google Drive link
 - https://drive.google.com/drive/folders/1gWnWTa1lj-ePBpUR81gsZ10HUZ3fXIAb?usp=sharing
- · O365 OneDrive link
 - https://018gc-my.sharepoint.com/:f:/g/personal/patrick_dooley2_ecn_forces_gc_ca/EsH_AvhZOthMhVNRyONN1gMBD4Tiht_agZ0d-.l4Q1W4cA?e=Sypmbk
- Tableau dashboards
 - VOICES (bilingual) https://public.tableau.com/app/profile/jvdh/viz/VOICES2020/Title
 - OUTER-NS CORA (english) https://public.tableau.com/app/profile/jvdh/viz/OUTER-NSENCORA/Homepage
 - OUTER-NS CORA (french) https://public.tableau.com/app/profile/jvdh/viz/OUTER-NSFRCORA/Homepage

Figure F.11: Slide 11 of orientation brief.



MS Teams Hackathon rooms

- · Hackathon plenary room is for everyone
- · One break out room for each team
 - Restricted to your team members, but you can invite someone to join you (e.g., one of the HAC; HAC can move hackers from room to room on request)
- Whiteboard app
 - Create a whiteboard for your team (via the "Share Screen" options menu)
 - Accessible from the breakout room and from the O365 Whiteboard app
- · Accessible via both the web and desktop MS Teams apps
- Rooms should be accessible 24/7, although you might be the only one there ©
- · Quick demo

Figure F.12: Slide 12 of orientation brief.



There be dragons ...

- > Data types may not be consistent from year to year
 - e.g., 2020 Subset 7, response percentages are not numeric!
- > Data fields and field names have changed over time
- > Question numbers and organization names change from year to year
- > Organization structures change from year to year (rightly and wrongly) and can be coded incorrectly
 - e.g., DRDC's structure is different in every year of data, and 2020 was a disaster
 - A few organizations appear more than once in the 2020 data (in two different hierarchies!)
- > The PSES organization Level codes are not the same as what DND uses
 - e.g., DRDC is an L1 in DND, but an L2 in the PSES data
- > Different releases of Anaconda (Python) process some data fields differently

Figure F.13: Slide 13 of orientation brief.



... and some dragon slayers

- ✓ In the sample Jupyter Notebook scripts, you will find
 - · notes of where we've found odd data types
 - · code to adjust for different data fields and/or field names in each years' data
 - · notes of where a change in Anaconda version made a difference
- ✓ We've mapped the PSES questions over time for you, and linked them to their themes/sub-themes
- √ We've mapped DRDC's org structure over time for you

> Read the PSES website info and Excel documentation files provided by TBS

- response percentages are weighted data
- what the 9999 code means
- · response categorization definitions

Figure F.14: Slide 14 of orientation brief.

24 DRDC-RDDC-2022-D111



How to get help

- · Who to contact for:
 - Technical issues with MS Teams, O365, Google Drive Patrick
 - Understanding the PSES data Josée, Patricia, Patrick
 - Using the sample data processing scripts Patricia, Josée
 Warning: most of you know more about Python than we do
 - Existential crises Patrick
- · How to contact us
 - MS Teams Hackathon plenary room (verbally or via text chat)
 - O365 e-mail
 - Someone will be available to respond from 0800 to 1800 EST each day

Figure F.15: Slide 15 of orientation brief.

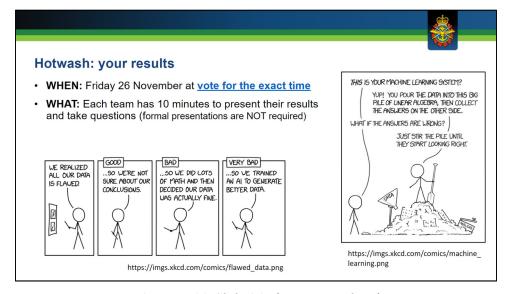


Figure F.16: Slide 16 of orientation brief.

Cartoons reproduced with permission of xkcd.com at https://xkcd.com/about/; the originals can be found at https://imgs.xkcd.com/comics/flawed_data.png and https://imgs.xkcd.com/comics/machine-learning.png.

DRDC-RDDC-2022-D111 25

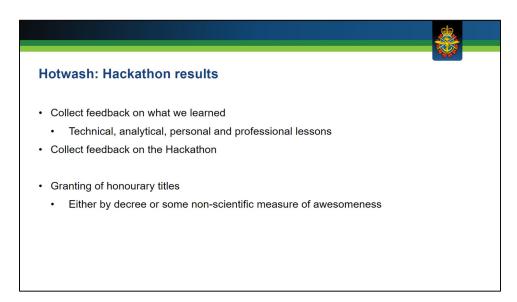


Figure F.17: Slide 17 of orientation brief.

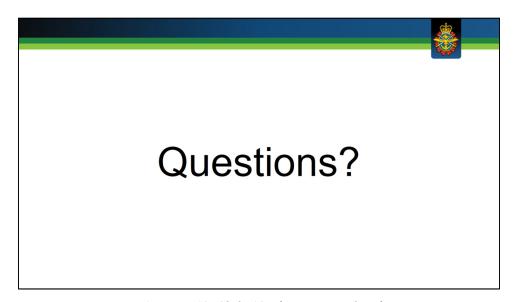


Figure F.18: Slide 18 of orientation brief.

26 DRDC-RDDC-2022-D111

Annex G Team appellations and awards

Immediately following the initial presentation of all teams' results, during an absurdly named ceremony ($S_{10}A_{10}$; Figure G.1), the HAC bestowed various whimsical mock appellations and awards. Such recognition was intended to further a light-hearted atmosphere and good-natured competition. The HAC's decisions were entirely subjective and based on online votes cast by its members during the teams' presentations. Each appellation or award was marked by a display of a virtual certificate that was created immediately prior to its presentation. Figure G.2 depicts one such example.



Figure G.1: Opening slide for recognition ceremony.



Figure G.2: Example of an award certificate.

28 DRDC-RDDC-2022-D111

The lists of awards, citations, and recipients are shown in Table G.1.

Table G.1: Summary of bestowed appellations and awards.

Name of Appellation or Award	ame of Appellation or Award Citation		
A Team by Any Other Name Award For the best team name.		APPROACH	
"The Quarkiest" Appellation	For correctly identifying Charm, Strange, and Truth.	CORA Jones & the Raiders of the Very Messy Survey Data	
The Charm Award	For the most enchanting contribution.	Hacksta's Paradise	
The Strange Award	The Strange Award For the weirdest contribution.		
The Truth Award	For the most revealing contribution.	CORA Jones & the Raiders of the Very Messy Survey Data	
"The Tempters of the Moirai" Appellation			
The Slam Dunk Award	The Slam Dunk Award For the greatest success that wasn't.		
"The Sting" Appellation	For detecting orcs and dragons in data.	Moirai	
The Other GRAPHIC Award	For the most totally awesome visualization.	APPROACH	
The Graphic Violence Award	For the image that made most use of the three-day excuse.	CORA Jones & the Raiders of the Very Messy Survey Data	
The Self-Hactualization Award	Self-Hactualization Award For enabling laziness most effectively via automation.		
The Chocolate in My Peanut Butter Award			
The Hack de Résistance Award	For the most amazing thing.	Medieval Cats	

Annex H Verbal feedback from hackwash session

Following the recognition ceremony, the HAC solicited participants' views on all aspects of the event during the hackathon hotwash (i.e., "hackwash") session (Figure H.1).



Figure H.1: Slide displayed to motivate discussion during the hackwash session.

During the hackwash session, one HAC member facilitated the discussion while two others took notes regarding participants' verbal remarks. A consolidated summary of those notes follows.

Format

Participants commented that the three-day timeframe was not sufficient. Although participants were told they could participate on a full-time or part-time basis, it proved too difficult to balance their everyday work/responsibilities alongside the hackathon. A potential solution would be to extend the hackathon for a week-long period with the understanding that hacking is on a part-time basis (about 50%–60% of participants' time). This would allow participants to attend to pressing client needs while contributing sufficiently during the hackathon. In addition, enabling part-time hacking may entice additional participation from those whose schedule would not permit such an activity otherwise.

November was probably not the best time for the hackathon. It coincided with PDA season and the calendar year wrap-up period. Other suggested timeframes are the spring or summer, although it was noted that summer would overlap with popular vacation times.

Teams found the 10-minute hotwash presentation time limit too short. The HAC's intentions were not sufficiently reiterated to the participants, where the hotwash was planned to be a brief informal presentation followed by an opportunity to present formally and at greater length to all CORA staff.

Recommendation: Based on participant feedback and the HAC's experience, we recommend that a future hackathon would run from Monday to Friday, with a hackwash on Friday afternoon to gather feedback. During the hackwash, each team would provide a three-minute teaser presentation of their results. During the following week, teams would present their results to all CORA staff formally and at greater length.

Value

For many participants, the hackathon was a time to conduct analysis in a fun setting. Participants enjoyed the opportunity to have dedicated time to throw themselves wholeheartedly into a project of their choosing, dubbed a "work field trip." The hackathon provided a unique opportunity to work and learn alongside colleagues who specialized in various fields and promote collaboration, particularly across scientific and analytical disciplines. It was a chance to learn colleagues' strengths and weaknesses, to gain contacts and mentors, to learn about new tools and techniques, to build rapport, and to know whom to reach out in the future.

Hackathon participants proposed sourcing problems to hack from our DND/CAF partners/clients. This would be similar to tiger teaming and could help achieve management buy-in for the event. This approach could have the added benefit of generating real-world impact that could be referenced for career progression or development purposes.

The HAC always intended that team results be provided in formal presentations, and where warranted, publications, but did not convey that clearly to participants. Future hackathons should more clearly promote the potential career progression/ development benefits of participation.

In the vein of the military's "train as you fight" concept, the hackathon could be considered a collective training event for CORA. This could be one means of filling a current capability development gap within the organization. One suggestion was to use a future hackathon to train participants on CORA's cloud computing capability; for example, as a test dataset and problem for learning how to use the analysis tools in the cloud environment.

Immediately apparent value for PSES analysis team

During the hackathon, Lachlan Gray, a member of team CORA Jones & the Raiders of the Very Messy Survey Data, worked on creating a Python package for importing/rationalizing datasets that are structurally similar but not identical. This would reduce the amount of time and oversight that is currently required when preparing the PSES data for analysis. The code was provided to the HAC and Bahaa Khaddaj expressed interest in collaborating to create this data automation package.

Team Moirai developed a prototype PSES dashboard in Power BI and, in the process, facilitated access to Power BI Pro licenses for the PSES analysis team, as well as creation of a shared workspace in D365. The lessons learned and shared by Team Moirai will greatly facilitate exploration of Power BI as an alternative to Tableau Public as a platform for hosting the PSES dashboards.

Overall impressions

CORA should host annual or twice-yearly hackathons. It was recommended that new questions, data and/or topics could be explored in future hackathons. Possibilities include COVID-19, weather, natural hazards, GIS, AIS, or OECD (economic) data. Potential datasets could be sourced from the government of Canada's Open Data site (https://open.canada.ca/en/open-data). Another suggestion was to take inspiration for hackathon themes from strategic analysts within CORA and their support to the SJS Strategic Outlook series.

There was a healthy discussion around the idea of setting up a Hackathon working group, where members would collect new questions/data from partners/clients and advertise and setup a hackathon in response. However, it

was agreed that it would be best to avoid having an overly structured or formal organizational entity, in order to preserve the hackathon's crucial agility and playful ethos.

Additional insights

Participants were predominantly relative newcomers to CORA. That raised the question of how more experienced CORA staff members could be encouraged to join a future hackathon. Changing the hackathon's timing (to the spring or summer) and advertising were suggested. Promoting the hackathon as a week-long part-time activity (versus three days full-time) might elicit more participation from busy staff who feel unable to participate full-time. Furthermore, to counter the common misperception that CORA's hackathons are synonymous with programming, new names for the hackathon were suggested:

- OR&A-thon [Operational Research & Analysis],
- CORA-thon [Centre for Operational Research and Analysis],
- MAR-athon [Modelling, Analysis, Research], or
- Op ANALYSIS.

Future hackathons would ideally include strategic analysts' involvement. Through focused messaging and discussions with strategic analysts, the HAC would hope to garner their participation and solicit new questions/themes to hack. A participant noted that mixed teams/sections and backgrounds promoted creativity. If possible, a strategic analyst on each team would be a great way to keep an eye on the "big picture."

Two participants (Bahaa Khaddaj and Jillian Henderson) expressed interest in helping to organize future hackathons.

Lessons observed

Table H.1 further summarizes the perceived successes and lessons arising from the hackwash session.

Table H.1: Condensed summary of successes and lessons observed from the hackwash.

Successes Lessons MS Teams breakout rooms worked—not Hackwash presentation format—be explicit that used by all teams, but were useful and presentations are brief and informal and that a longer, functional overall formal, CORA-wide opportunity will follow • Late November is not the best hackathon timeframe Team building Learning about the tools and techniques of Three days was too brief as people couldn't dedicate all colleagues, which enables future of their time based on their current workloads More explicit management buy-in may be needed, to contacts/mentors People valued having a break from their ensure dedicated time for the hackathon, hackwash, and usual work; a boost to their mental health all-CORA presentations Professional development/training event • Yearly, if not more frequent

- Flexibility with respect to participation (did not have to be 100% for three days)
- Open-ended problem space allowed for creativity; some pre-defined problems/goals gave those wanting more structure a base from which to proceed
- The term "hackathon" can be a turnoff for some who believe their skills won't apply to strictly data analysis
 Modifications are needed to appeal to experienced
- Modifications are needed to appeal to experienced CORA staff

Future hackathons should be on other topics

- Need to explain the open-ended problem space concept better when advertising the event
- Presentations/publications arising from the Hackathon would further demonstrate its value

Annex I Satisfaction survey instrument

This annex's figures depict the online Microsoft Forms satisfaction survey instrument that the HAC distributed to all participants immediately after the hackathon ended. Survey completion was voluntary and all submissions were anonymous.

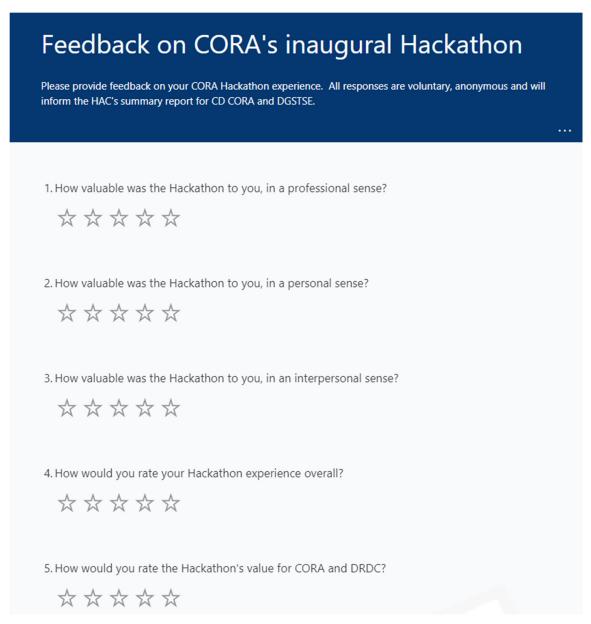


Figure I.1: Items 1-5 *of the online satisfaction survey instrument.*

C.Miles did see and enter in the trade to 2.81	-t - II tht I.
6. Why did you particpate in the Hackathon? Please selec	ct all that apply.
To explore science/analysis	
To be creative	
To test an idea, method, or tool	
To learn things	
To improve my analytical skills	
The topic interested me	
To make a useful contribution	
To work with people I don't usually work with	
For team building	
☐ To work on something different for a few days	
I thought it would be fun	
Someone coaxed, bribed, or guilted me into it	
Other	
7. What aspects of the Hackathon worked well?	
Enter your answer	
8. What aspects of the Hackathon did not work well?	
Enter your answer	

Figure I.2: Items 6–8 of the online satisfaction survey instrument.

	O	O	0	0	O
10. If a second Hackat	hon were held, would	you participate	e?		
	Definitely not	Probably not	Maybe	Probably	Definitely
	0	0	0	0	0
11. If a second Hackat	hon were held, would	you prefer a d	ifferent them	e?	
	Definitely not	Probably not	Maybe	Probably	Definitely
			Waybe	Probably	Delinitely
12. If a second Hackat Enter your answer	hon were held, what o	other themes w	0	0	O
Enter your answer			ould you rec	ommend?	
Enter your answer 13. If a second Hackat			ould you rec	ommend?	
12. If a second Hackat Enter your answer 13. If a second Hackat			ould you rec	ommend?	
Enter your answer 13. If a second Hackat			ould you rec	ommend?	
Enter your answer 13. If a second Hackat			ould you rec	ommend?	
Enter your answer 13. If a second Hackat	hon were held, what v	vould you add,	ould you rec	ommend?	
Enter your answer 13. If a second Hackati Enter your answer	hon were held, what v	vould you add,	ould you rec	ommend?	

Figure I.3: Items 9–14 of the online satisfaction survey instrument.

Annex J Satisfaction survey results

Here, we have compiled the voluntarily and anonymously submitted results of the hackathon's online post-event satisfaction survey. We have binned ordinal and nominal responses and summarized them as histograms. We have grouped free-text responses thematically. Seventy-one percent of participants (i.e., 10 of 14) responded to the survey.

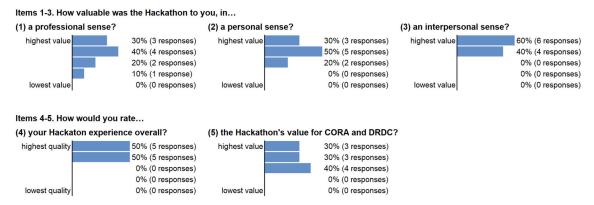
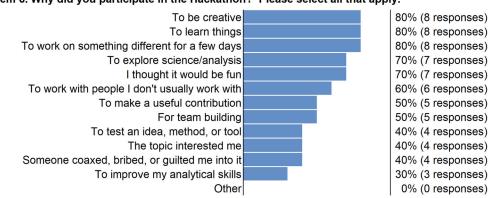


Figure J.1: Histograms summarizing participants' responses to survey items 1–5.



Item 6. Why did you participate in the Hackathon? Please select all that apply.

Figure J.2: Histograms summarizing participants' responses to survey item 6.

Categorized responses to item 7:

What aspects of the Hackathon worked well?

Hackathon format

- The Hackathon was well organized and the variety of problems posed made it very easily accessible for any DS to work with the data and explore/analyze the PSES data.
- Instructions, pre-packaged data, and meetings.
- The flexibility of no specific start times or deadlines.
- A topic that was of general interest/familiarity.

- No limit to the subjects examined.
- The openness of the assignment, so to speak, allowed for maximum creativity.
- The open endedness of the Hackathon allowed for extensive creativity.
- Having the "drop in" chat/Teams room always open made true collaboration possible.

HAC-provided data files

- Getting the files was easy, everything was explained well, the mapping files were very useful.
- The dataset was mature and varied, and allowed for real creativity.
- I liked that there was data available from all steps of the process so we could jump in at any point along the way.

Team composition and collaboration

- Putting different sorts of people together.
- Good use of targeted recruitment.
- The interdisciplinary collaboration was invaluable, especially during the pandemic where in person networking is impossible.
- The spontaneity of the teams that came together.
- Working as a team with a broad [remainder of response not collected].
- Great for networking, actually felt like a team working together for a purpose (which is hard when all remote).
- I also enjoyed being on a team with two people not in my section, and getting to know more of CORA.

Categorized responses to item 8:

What aspects of the Hackathon did not work well?

Hackathon format

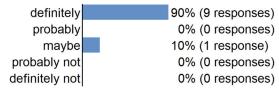
- It was hard to connect with my team, since there were only a handful of hours during the Hackathon when one of us wasn't in a meeting.
- We could have used more time than just 3 days, perhaps a week? As other work obligations prevented us from working consistently on only the hackathon.
- Demands from other projects was frustrating, but I like the suggestion of having this at a different time of the year.
- The time of the year.
- The hackwash felt a bit rushed (understand that people already spent a lot of time on the development, but that's why they want to spend time to see the results).
- The short 10 min briefs at the Hackwash, perhaps more time to present our work.
- Did not understand the names of all of the awards.
- The end result was a bit more clear (see further responses).

Nothing to suggest.

Team composition and collaboration

- I wish we could meet our team members in person; our work was a bit separate.
- I think marketing of the hackathon next year should encourage more soft scientists to participate, and then partner them with data scientists. This was the sake with [my team], and it was really neat to implement [my team member's] creative ideas into a data project.
- This is more of a suggestion, but team sizes between 3–5, rather than a strict group of 3.





Items 10-11. If a second Hackathon were held...

(10) would you participate? (11) would you prefer a different theme? definitely 60% (6 responses) definitely 30% (3 responses) probably 20% (2 responses) probably 40% (4 responses) maybe 20% (2 responses) maybe 30% (3 responses) probably not 0% (0 responses) probably not 0% (0 responses) definitely not 0% (0 responses) definitely not 0% (0 responses)

Figure J.3: Histograms summarizing participants' responses to survey items 9–11.

Categorized responses to item 12:

If a second Hackathon were held, what other themes would you recommend?

Client-focused

- Consulting with the CAF for (potentially) quick-win problems could be interesting.
- Something client-centric would be nice. We could have a CAF partner present an issue or think of an issue facing the CAF as a whole.
- Client posed questions.
- Ship detection.
- Alternatively, doing a hack-a-thon with the Frigate IPMS sensor data could be fun, but feels a bit code-y-er.

Pandemic-related

- COVID-19 related questions using GIS data.
- COVID.
- COVID.

Environment-related

- Climate change.
- Weather.
- Natural hazards.

Open-source data

- Other open data sets, something with a geographic component can be fun and likely something new for many analysts.
- Social media related questions, such as measuring public sentiment on certain topics, also could use GIS
- Maybe looking at the open-data Government of Canada site could provide more inspiration.

Status quo

- I am not sure, but I liked that the PSES is impactful to everyone who participated.
- No idea.

Categorized responses to item 13:

If a second Hackathon were held, what would you add, modify, or remove to improve it?

Announcement messaging

• Slightly more "what is a hackathon?" type guidance at the beginning, which would also have allowed more people who are generally unfamiliar with hackathons to participate maybe. (The solicitation email wasn't very illuminating.)

Hacking time allotted

- Would probably give Monday to Friday timeline but knowing that some individuals may not be able to work full days.
- Change from a few full-days to more half-days to allow people to work around meetings and other commitments. For example, the Hackathon could be a week-long and run in the afternoons.

Presentation of results

- Lengthen the presentation time.
- Given that most of the presentations were looking to produce useful insights on how CORA/DRDC are managed, I expected at least one manager (the Chief Scientist, SHs) to be present to take some interest in the results and note the good work done.8

⁸ This comment was the result of a misunderstanding. The HAC's hackathon plan included two opportunities for teams to present their work. The first opportunity (immediately following the hackathon's conclusion) was brief and meant for participants only. The second opportunity would include all CORA staff and manager who wished to attend. Due to an oversight on its part, the HAC did not remind participants that the second opportunity would be forthcoming. The flagged survey comment was made prior to the second, all-CORA opportunity.

Post-hackathon impact

- Would like to see some follow-on/integration of the results to improve the response to the posed problem.
- Maybe more open ended problems that could provide opportunities for future collaborations that extend past the hackathon.

Status quo

- I don't know.
- Not sure.

Responses to item 14:

Please provide any additional comments that you'd like to share.

- Great job organizing this Hackathon. Bravo! It was a pleasure participating!
- This was excellent fun!
- Thank you for organizing this. I had a ton of fun!!
- Thanks for this guys it was fun!
- Thank you Patrick, Josee and Patricia for organizing the event.

Annex K Abstracts for the formal presentations of teams' results

Each team was offered an opportunity to brief its results formally to all interested CORA staff and managers, from 19–21 January 2022. All teams participated; the abstracts for their presentations follow.

K.1 DAY 1—Wednesday 19 January 2022

12:00–12:20 EST—Hackathon Administrative Committee (Patricia Moorhead, Josée van den Hoogen, and Patrick Dooley)

A brief introduction to the hackathon.

12:20–12:40 EST—Team Medieval Cats (Sarah Babbitt, Andrea Lane, and Stephen Weber)

Combining social science, engineering, and computer science, Team Medieval Cats took the difficulty of
accurately capturing affective/emotional states in the PSES, and explored alternate response methodologies
including colour, emoji, and sliders.

12:40–13:00 EST—Team MOIRAI (Lise Arsenault, Cheryl Eisler, and Mikayla Holmes)

Team MOIRAI tackled the 2020 PSES results from the perspective of accessibility: anyone who wants to view the data, from any department, looking at any question without barriers (be they topic, language, color-blindness, text-to-speech would be the next step...). We also wanted to showcase the ability to utilize approved internal systems, such as the DWAN, to develop modern, shareable, flexible, extensible, reusable tools in Microsoft Power BI. This allows users to work up to Protected B in the Desktop version and Protected A in D365, as well as integrating R and Python scripts. Use of Power Bi Pro enables people to collaborate in workspaces, develop reports that can be shared with specific user groups, and produce dashboards that can be published internally or publicly (assuming unclassified data).

K.2 DAY 2—Thursday 20 January 2022

12:00–12:20 EST—Team APPROACH (Michèle Fee, Victor Isaac, and Bahaa Khaddaj)

The APPROACH team sought to automate PSES data extraction and processing, as well as provide interactive visualizations of the data. This resulted in a script that automates data processing, a shiny application displaying DRDC/DND results as heat maps, and an interactive tool for displaying DRDC/DND results in trees.

12:20–12:40 EST—Team Hacksta's Paradise (Jillian Henderson and Stephen Okazawa)

The Public Service Employment Survey (PSES) is about making the public service a better place to work. But... it's also interesting to know where each organization ranks. So, if we treat PSES data as the results of a competition to determine how "good" each organization is, how can we give each organization a score and a final ranking? Also, some questions are more informative than others. How can we quantify the utility of each question in terms of distinguishing good from bad? Can this be used to weight each question? In this work, we show a simple and elegant way to rank organizations using a single score calculated from the PSES Likert scale results. We further refine the ranking scheme by determining question utility to the overall ranking by considering response concentration, randomness, and question redundancy.

12:40–13:00 EST—Team Charm (Josée van den Hoogen)

Due to the size and nature of the PSES datasets, a lot of manual effort is required to prep and clean the data for analysis. Linking multiple datasets and decoder tables is possible within Tableau. This greatly reduces the amount of manual effort required to prepare the data and allows otherwise too large datasets to be visualized. A visualization dashboard was created using the raw data and PSES decoder tables to showcase net sentiment per question, including filters by organizational unit or by demographic. In addition, Tableau Public has a visualization creator that was used to experiment in order create a new summary visualization for the relative change in PSES results year to year. Merging a tornado plot and a treemap plot captures both the relative size and colour differences. Dubbed a Tornado Treemap, this new visualization summarizes the yearly change in extreme net sentiment in an organization.

K.3 DAY 3—Friday 21 January 2022

12:00–12:20 EST—Team CORA Jones & the Raiders of the Very Messy Survey Data (Maude Amyot-Bourgeois, Lachlan Gray, and Matthew MacLeod)

- We present a comparison of the responses given by the "Defence Scientific Service" (DS) Research classification within DND, with the responses given by other departmental/agency cohorts of Research Scientist classifications within the Canadian Government, in departments or agencies. First, the sentiment towards senior management and overall satisfaction with one's organization are explored over a period extending beyond the realm of the Hackathon (2011–2020). Then, the 2017–2020 data survey are probed to compare DS in DND to three other departmental cadres of researchers regarding six sub-themes of interest to the presenters (Discrimination; Harassment; Diversity and Inclusion; Employee Engagement; Senior Management; and Use of Official Language) during this 4-year period.

12:20–12:40 EST—Team Strange (Patrick Dooley)

Senior managers generally want clear, concise, insightful, and accessible characterizations of their organizations' overall PSES results. Satisfying those requirements when distilling a hierarchical organization's responses to hundreds of survey items into a single big-picture graphic can be challenging. How might one develop such a comprehensive visualization—or go further to reveal how an organization's results changed over time? Alternatively, how might one compare two organizations' overall results at a glance? Might such visualizations apply to surveys beyond the PSES? Can one create them using 1980s-era technology and techniques? What might such graphics reveal about how CORA's PSES performances varied over time—or how DRDC's results changed during SCInergy 2020? What might they reveal about DRDC's approach to remediation? Might they provide support for a different approach? For answers to these questions, kindly attend and judge for yourself!

	*Security markings for the title, authors, abstract and keywords must be entered when the document is sensitive							
1.	 ORIGINATOR (Name and address of the organization preparing the document. A DRDC Centre sponsoring a contractor's report, or tasking agency, is entered in Section 8.) DRDC – Centre for Operational Research and Analysis Defence Research and Development Canada Carling Campus, 60 Moodie Drive, Building 7S.2 Ottawa, Ontario K1A 0K2 Canada 				2a. SECURITY MARKING (Overall security marking of the document including special supplemental markings if applicable.)			
			alysis	CAN UNCLASSIFIED				
				2b. CONTROLLED GOODS				
			NON-CONTROLLED GOODS DMC A					
3.	TITLE (The document title and sub-title as indicated on the title page.)							
	The Defence Research and Development Canada – Centre for Operational Research and Analysis Hackathon 2021: A comprehensive summary							
4.	AUTHORS (Last name, followed by initials – ranks, titles, etc., not	to be	used)					
	Dooley, P. W.; Moorhead, P.; van den Hoogen	, J.						
5.	DATE OF PUBLICATION (Month and year of publication of document.)	6a. NO. OF PAGES (Total pages, including Annexes, excluding DCD, covering and verso pages.)		s, including cluding DCD,	6b. NO. OF REFS (Total references cited.)			
	September 2022			4	8	1		
7.	DOCUMENT CATEGORY (e.g., Scientific Report, Contract Report	t, Scie	entific Let	ter.)				
	Reference Document							
8.	SPONSORING CENTRE (The name and address of the departme	nt pro	oject offic	e or	laboratory sponso	ring the research and development.)		
	DRDC – Centre for Operational Research and Analysis Defence Research and Development Canada Carling Campus, 60 Moodie Drive, Building 7S.2 Ottawa, Ontario K1A 0K2 Canada							
9a.	PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)						
Operational Research and Analysis								
10a.	DRDC PUBLICATION NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.)	OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)						
	DRDC-RDDC-2022-D111							
11a.	1a. FUTURE DISTRIBUTION WITHIN CANADA (Approval for further dissemination of the document. Security classification must also be considered.)							
	Public release							
11b.	FUTURE DISTRIBUTION OUTSIDE CANADA (Approval for furthe considered.)	r diss	seminatio	n of	the document. Sec	curity classification must also be		
12.	KEYWORDS, DESCRIPTORS or IDENTIFIERS (Use semi-colon a	as a c	delimiter.)					
	DRDC CORA; hackathon; PSES; Public Service Employee Survey; survey analysis; operational research and analysis; operational research methodologies; operational research; data analysis; data analytics; data visualization; training; collective training; team building							

13. ABSTRACT (When available in the document, the French version of the abstract must be included here.)

To enhance its analytical culture, capabilities, cohesion, and morale, Defence Research and Development Canada (DRDC) – Centre for Operational Research and Analysis (CORA) conducted its inaugural hackathon from 23–25 November 2021. As the Hackathon Administrative Committee (HAC), we proposed, planned, and conducted the event. Overall, the hackathon provided a flexible, accessible, and fun environment wherein participants could be truly creative, learn, experiment, and collaborate, and yielded multiple products with substantial potential to improve the ways in which CORA conducts its annual Public Service Employee Survey (PSES) analyses for DRDC. Here, we summarize the hackathon's objectives, execution, outcomes, and participants' feedback, and provide recommendations for improvement if a similar event is held in the future.

Afin d'améliorer sa culture, ses capacités, sa cohésion et son moral en matière d'analyse, Recherche et développement pour la défense Canada (RDDC) – Centre de recherche et d'analyse opérationnelle (CARO) a tenu un marathon de programmation inaugural du 23 au 25 novembre 2021. À titre de comité d'administration du marathon de programmation (CAMP), nous avons proposé, planifié et dirigé l'activité. Dans l'ensemble, le marathon de programmation a fourni un environnement flexible, accessible et amusant qui a permis aux participants d'être véritablement créatifs, d'apprendre, d'expérimenter et de collaborer, et a donné lieu à plusieurs produits ayant la possibilité considérable d'améliorer la façon dont le CARO effectue ses analyses annuelles du Sondage auprès des fonctionnaires fédéraux (SAFF) pour RDDC. Nous résumons ici les objectifs, l'exécution et les résultats du marathon de programmation, ainsi que les commentaires des participants. Nous formulons également des recommandations pour améliorer une future activité semblable, le cas échéant.