Towards “No Wrong Door” on the web

Virtualizing in-person referrals and wayfinding assistance

Presented to:	Youth Employment and Skills Strategy Employment and Social Development Canada
Presented by:	Magnet & MENTOR Canada
Date:	07 May, 2021
Version:	2.0

Contents

Executive Summary

We are pleased to present ESDC with this exploration for implementing an intelligent technology to match Canadian youth to relevant resources, including face-to-face services, and opportunities from a national network of service providers and sources connected. These services can be identified and accessed from anywhere across that network, regardless of the point of entry (No Wrong Door approach), using a common tool that alerts users where they are and directs them to other opportunities and resources located elsewhere that may match their needs based on who they are and what they are presently looking at.

This report was commissioned by the Youth Employment and Skills (YESS) program to advance ESDC's understanding of digital infrastructure (both for the Youth Program Navigator pilot, as well as for other system-wide supports), and supported by input from a technology working group that included representation from:

- Magnet at Ryerson University
- MaRS
- O4AY
- LinkedIn
- MENTOR Canada
- 211 Central Region / Findhelp Information Services
- Employment and Social Development Canada (YESS, Youth Digital Gateway)

This outline is further informed by research and a series of interviews with service providers, taxonomy architects within the employment services space, youth job seekers, and technologists serving actors in the nonprofit sector (Appendix B).

This tool, known as CORDS, will be supported by algorithmic taxonomy, using natural language processing and machine learning. It will consider users, their interaction with participating web properties, and tags applied to the resources and services presented on those web properties. A feedback mechanism will allow users to indicate if they found the information useful. This feedback will be used to refine the wayfinding over time. Additional research considered how existing taxonomies and datasets could be reconciled to better serve users.

This report includes:

- An overview of the challenges in the youth services ecosystem and objectives of CORDS in addressing those challenges.
- The supporting research and background information that informed recommendations.

- A series of recommendations related to infrastructure, technology pathways, and product development to ensure project success.
- An outline of the next possible stages of the project and stakeholder requirements.

We expect that this document be treated as a working model that can be adjusted based on ongoing discussions with ESDC.

Terminology

- **Algorithmic Taxonomy**

  An algorithmic taxonomy is using an algorithm to create a taxonomy. In this application, we are using a natural language processing model to turn descriptions of opportunities and resources into numbers (vectors) and then using those numbers to easily compare the corresponding descriptions as a taxonomy. This will also allow us to compare across taxonomies.

- **Data Pool**

  A data pool is a central repository of data that needs to be accessed by an end-user. Mechanisms like vectorization or algorithms can be used to feed relevant data to users based on their needs.

- **Expert System**

  An expert system mimics human knowledge, skills, or behaviour. Expert systems are founded on a pool of knowledge or data and an inference mechanism that applies rules and facts to the pool of data. The expert system, therefore, determines how a user interacts with data.

- **No Wrong Door**

  No Wrong Door is an approach to connecting users with services across multiple entry points. The objective of No Wrong Door is to direct users to the information they need even if they didn't initially find themselves there. Applied to human-facing services, No Wrong Door entails developing a referral mechanism that assesses a user's needs and points them in the direction they need. For example, a user who needs access to a food bank can land on a directory for housing services through an online search, but can be referred to the page they were looking for and potentially other services they may require.

- **Opportunities and Resources**

  The data addressed in this report fall broadly into two categories as follows:

Opportunities	Resources
Any service related to job training, wage subsidies, job opportunities, and learning opportunities such as workshops. These are specifically accessed by the user on an individual basis. Examples: ● Academic bridging programs ● Trade certification workshops ● Student wage subsidies ● Virtual job fairs ● Youth entrepreneurship programs / small business startup loans (e.g., Futurpreneur)	Informational and instructional pieces that help youth navigate skills training ecosystems and that allow them to navigate career and educational pathways, make informed decisions, and set goals. Examples: ● Foodbank and shelter directories ● Instructional guides for securing and updating relevant documentation ● Referrals to providers of career counselling ● Mental health services ● Interview preparation tools

- **Taxonomy**

  A Taxonomy is a system that names and classifies entities (things or concepts) based on shared characteristics. The practice of developing a taxonomy is useful for grouping a wide range of things and directing observers to one of those things based on what the observer is looking for. Within the social impact sector, taxonomies have been implemented to classify occupations, health services, and skills.

- **Web Property**

  Something intended to present to people (as opposed to computers) digital information accessible from an IP address. These include simple informational websites like Big Brother Big Sisters of Central Vancouver Island, chat-based apps like Chalmers, information aggregator websites like 211 Alberta or Emplois Québec's Online Placement system.

Summary of CORDS

Vision

> **Create a digital No Wrong Door ecosystem in which youth across Canada can land on any publicly funded web property related to employment opportunities and resources, and automatically receive relevant referrals to other supports that might assist them on their journey.**

Across Canada, a network of organizations, including all publicly funded employment and skills development projects directed at youth, will be connected through a common technology that can push these services to users across any web property connected with that national network.

The system will be supported by an automated approach by which youth job seekers can be alerted to other relevant resources and opportunities to support their employment journey.

Service providers and community organizations across Canada can have greater visibility among youth without taking on significant costs for advertising or search engine optimization.

The Problem

Youth have to rely on search engines to identify services and resources. In instances where they land on a page and the results are not relevant, their first option is to search again.

There is no tool to discern what they may be looking for and to point them in the right direction the way an experienced youth services professional might in an in-person encounter or any tool that can translate queries and the language youth use to search for services into relevant recommendations. There is also no tool to recommend additional services and resources based on a user's initial needs, such as access to food banks, CV and interview preparation courses, or health-related services.

The Solution

Two technologies in cooperation can solve this problem well. We can leverage machine learning to accelerate and make multilingual and synonym-accepting search possible. By transforming the descriptions into numbers, we can easily and quickly relate queries, opportunities, and resources. Subsequently, after gathering usage data, we can implement an improved

recommendation system that does not rely on taxonomy. It instead relies on the relationship between those quantified descriptions and data about prior users' navigation. This latter approach will connect users to any opportunity or resources that have undergone a community uptake process. This process will create the network that connects them as one CORDS network.

The Challenge

There are a multitude of resources and opportunities available for youth across Canada, but uptake and utilization might be slow and challenging due to a series of structural and functional factors.

Structural Challenges

Service Fragmentation

Human facing and virtual services related to youth employment and counselling are delivered through a variety of channels. If users need to access services across networks, there is no provided path from the website of one provider to a related website or to services that cross sectors or jurisdictions.

Lack of common language

There are individual search taxonomies developed by governments and service providers across Canada. In each case, youth can drill down based on broadly defined categories for services. The results of these searches are determined only by the user's inputs and limited by their understanding of their current situation. Furthermore, results are typically confined to services available through that provider. In many instances, users will need additional services that are available from other providers or professionals.

In face-to-face encounters, experienced front-line staff can make inferences about additional services clients may need. Current taxonomies have also captured data on what services are needed based on initial requests. As the need for greater digitization increases, new tools should capture this knowledge to serve youth clients on a national scale.

Linkages not defined between related services

Based on a user's inquiries, they may need access to additional resources that are typically considered outside the purview of career or employment services, including mental health and legal services. This knowledge exists, though largely institutionally, in the decades of knowledge that front-line providers have accumulated and in data collected by active taxonomies like 211 LA. While preserving people's privacy, it's possible to capture linkages between the kinds of support needed beyond what's captured in the initial matching taxonomy.

The fragmentation and lack of awareness among individuals accessing service providers entail an arduous and time-consuming process for users to connect with the services they need, thereby increasing the chances that individuals will "drop off" in their search, abandoning their goal.

Functional Challenges

Lack of Awareness

Youth job seekers may not know what services are available to them or which services they need based on their circumstances or where to start looking. Users will benefit from a No Wrong Door approach in which they can land on any service provider's page and be redirected.

Navigational "Landlock"

Each web property a visitor might land on manages and maintains different information organized in different ways. The effect is a "landlock". Users can land on a specific page through an online search, an "island" of a site, but are completely unaware of other relevant resources that are available through other websites or portals. If the user lands on a page, but it's not what they need, their recourse is to return to the search engine. However, the results they will find by repeating their search will not be based on the user's needs, but:

- a content producer's ability to properly describe their content, potentially employ SEO techniques, and leverage tools like Google Analytics to continuously refactor content to optimize the acquisition of an intended audience delivered by search engines,
- the searcher's ability to know and describe their wants/ needs,
- the search engine's ability to, or intent in, connecting the two things, and potentially
- the content producer's ability to pay the search engine for placement in results for specific searched terms.

The following scenarios are thought experiments of online experiences to demonstrate how both youth and front-line providers may experience the current ecosystem followed by the implications for developing a more robust and connected experience.

Scenario 1

A youth in Northern Ontario heard about an arborist training program being delivered at a local college. The user searches “arborist training” and the results are largely for the home pages of forestry and arborist services. The user also tries searching for “arborist training YMCA” because they recall hearing that the YMCA was involved in the program. The second search yields somewhat more relevant results, including pre-apprenticeship programs from the Career Foundation and the provincial Second Careers program.

In either case, the results don’t bring the user any closer to the page they wanted to find and the result they wanted, which was to confirm eligibility and apply. The student would also have benefitted from additional information on stipends to cover tuition as well as clothing and equipment allowance supports.

The user can continue searching, but the results they find are significantly influenced by paid ads and SEO, not necessarily an understanding of the user’s needs. There is also no mechanism to take them from the pages where they land to either a different web property where their needs will be better served or a page within the website they are currently on.

In the case of the second search, where they land on the page of a publicly funded employment or skills training program, there is an opportunity to provide a tool that allows for “hand-holding” or redirecting, rather than having to jump over to searching across a different taxonomy. At a very basic level, if we know that a student is looking for an arborist training program, we can offer suggestions to view the skills taxonomy of that occupation, additional training programs, and financial aid options.

Scenario 2

A non-profit in the Niagara Region has received funding from the Federal government to deliver a small program for youth between the ages of 18-24 starting careers in specific sectors in the region to access a subsidy to cover the costs of training. Due to shortages in staff, technical know-how, and lack of funding to support large-scale promotional efforts, the organization may face a challenge in driving uptake of the initiative.

In a system where search engines are relied upon to create visibility, service providers will have to invest in a combination of online ad campaigns as well as Search Engine Optimization efforts, both of which can be expensive and not necessarily ensure that content is visible.

A networked approach, that can understand when youth job seekers might need to connect with this type of content and navigate them to it, allows the heavy lifting to be shifted from individual organizations to an automated and machine-learning supported system.

Breaking the “Landlock”

To break the cycle in which users and services can’t find each other we would need a way to emulate what happens when contact is made face-to-face between an opportunity and resource-seeking youth and staff or a volunteer.

CORDS resolves all of the challenges:

- Service fragmentation, lack of awareness, and navigational landlock are removed by connecting web properties with a ubiquitous interface that refers across sites.
- Lack of common language is alleviated by multilingual, synonym-friendly search.
- Undefined linkages between services are recorded and surfaced to users.

> **Scenario**
>
> An individual is looking for the services of a foodbank and knows there is one being run out of a local community church, which also manages around 40 shelter beds. They are focused primarily on their next meal because they have also experienced homelessness and addiction for the past year. When they arrive at the church, they inquire about the shelter beds after seeing a notice about available beds upon walking in. The shelter is a women’s only space, so they are not eligible, but the shelter manager mentions to them that they can be served at another shelter facility across town, which also manages a detox program. Based on their entry point into the first shelter, the manager knew that someone walking in the door might require additional services associated with precarious housing and addiction and could make recommendations accordingly, even if it wasn’t explicitly stated by the person walking in.

To execute scenarios like this digitally, we would need to implement an Expert System to actively (e.g., “Please tell me a bit about yourself…”) or passively (e.g., by inference based on the user’s navigation) interact with visitors and present recommendations or referrals. Such a system would encapsulate the knowledge of “front-line workers” who are capable of understanding the circumstances of a client and assisting them with referrals to other supports.

CORDS Fundamental Requirements

To reproduce the functions of a front-line worker, an automated system would need to:

- Interact with the user where they are (i.e., be available on the web properties to which the user travels).
- Gain an understanding of the user and their circumstances.

- Know about a wide range of reliable and reputable opportunities and resources.
- Given the user's circumstances, relate the user to appropriate opportunities and resources.
- Adapt its recommendations based on experience (i.e., learn on the job).
- Report on its activities and share its insights.

Key Considerations

What is the range of options we might consider to implement the Fundamental Requirements? We have framed exploration of these options as answers to a series of questions that need to be addressed together to realize our objective:

- What like CORDS is already out there?
- In what different ways are opportunities and resources described?
- How can we make recommendations across unrelated taxonomies?
- What could be the common way to represent opportunities and resources?
- How can we access data?
- What do we need to know about a user to help us to help them?
- How could we acquire data about the user?
- Is it a good idea to leverage user profiling to provide recommendations?
- How do we make good recommendations to users?
- How can we best track users while respecting their privacy?
- How should the system be hosted?
- How do we ensure the quality of the content in CORDS referrals?

Based on our exploration of these questions, a series of recommendations is presented.

What like CORDS is already out there?

We have examined initiatives that allow users to explore and access services that are available locally and that are relevant to their needs. Our objective was to learn more about the kind of services people were able to access via structured data pools and how the services select recommendations for users from those data pools (e.g. did the system provide additional recommendations based on the user's activity or prompt additional requests from the user to direct them more strategically?).

The following are examples of services that have implemented a single entry point by which users can access services across multiple networks. We were not able to find anything that implements the Fundamental Requirements exactly, but we found some examples of initiatives tackling similar challenges.

HelpSeeker.org	HelpSeeker allows users to privately browse thousands of community, provincial and federal health and social services, programs, resources, helplines, benefits for mental health, counselling, parenting, education, training, addictions, domestic violence, affordable housing, shelters, food support services, recreation, and more (via website). HelpSeeker also offers services to community partners that take up the technology, including Integrated Needs Assessment to explore trends, gaps, and strengths among community networks of providers. The assessment can open the gateway to cross-partner referrals.
FindHelp.ca	Operates a 24/7 call-in service to support referrals to food, settlement, housing, employment, and victim support services, among others. Referrals are driven by trained operators who assess a caller's needs, connect those needs to relevant programs, and guide users through application and registration requirements.
American public sector examples	In 2015, the Administration of Community Living awarded a total of $5 million to agencies across 13 states to develop a No Wrong Door approach that connected counselling services. Examples: Colorado Access: Collects long-term health providers from

	across the state. Meant to support access for citizens over the age of 65. Does not offer additional recommendations based on user searches. 2-1-1 Virginia: Similar to 211 LA, users can search through a statewide database of resources and programs under general headings such as housing, mental health, clothing, personal needs, and income support. Users can drill down within these categories by zip code to find resources available in their jurisdiction. DC Support Link: Centralized database and search tool for veteran and active service member supports. Includes interactive search tool for support services as well as news updates, and a learning centre of informational material that can also be accessed within the search function. Results turn up not just service agencies, but relevant documents, blog posts, and news articles. Network of Care: Supported by Trilogy Integrated Resources. Collects services in select areas across 30 states. In addition to a service directory, NoC also includes e-learning services, a knowledge database, and public health data to support policy development. These examples represent what can be called a networked or "one-stop shop" approach, where one entry point can take users to a variety of endpoints that fall under the umbrella of human services. The search tool or taxonomy does not follow the user across properties, but they can return to the centralized page to continue finding resources.
FutureGov	FutureGov has implemented a range of tools across the UK to enable digital transformation in the social services sector. One relevant example includes the "Digital Front Door" project in North Tyneside. The project is currently in the process of prototyping a digital advice tool that allows residents to input their own care needs and link to relevant services. The Essex County Council is also collaborating with FutureGov on a project that maps service journeys and patterns. The project will identify how services are linked, with research supporting new user journeys that reflect those patterns.

In what different ways are opportunities and resources described?

To offer recommendations of opportunities and resources from data pools that might not be managed by the same organization, we need to understand how the different services youth might need to access are currently made available, and how they relate information to users.

Job training programs, networking, learning opportunities, social services, health services, and a host of additional resources that youth may need to access are currently organized, without overarching standardization, in structured data pools. The more inclusive and extensive data pools rely on sophisticated taxonomies managed by distinct actors. For example:

Canada Job Bank	Employers across Canada can post available jobs, which job seekers can search based on categories like salary, location, industry, hours, and more. Jobs are classified by NOC code, which in turn encompasses associated skills, duties, and requirements.
NOC	The National Occupational Classification (NOC) is Canada's national system for describing occupations. You can search the NOC to find where an occupation is classified or to learn about its main duties, educational requirements, or other useful information.
AIRS Taxonomy for 211	211 is a guide to services available across several regions. This taxonomy is now maintained by AIRS.
Quebec Vocational Groups	Quebec Vocational Groups is promoted as a one-stop-shop for all the resources and information an individual needs to access and complete vocational training in Quebec. Includes background information on bursaries, but no linkages to services available outside of its purview.
Canadian Skills Taxonomy	The Taxonomy serves to streamline terminology across several competency domains and concepts (e.g. skills, personal abilities and attributes, knowledge, interests), occupational work context, work activities, and tools and technology information, while aiming to improve the comparability of their incidence and application throughout occupations and sectors. The Taxonomy also complements the development of a range of labour market information products within ESDC, such as the Career Handbook, which details the competency requirements for specific occupations and provides other skills utilization

	indicators (e.g. the importance of a skill within a particular occupation, and/or the frequency of its use).
OnetoOnline Skills Taxonomy Browsing	Drill down based on occupation and job level to discern knowledge, skills, tasks, etc. associated with that occupation. Can drill down further and find training and certification programs for those occupations.

No one taxonomy exists that includes the breadth of descriptive power needed to, itself, represent the range of opportunities and resources to which CORDS would want to be able to refer users[1]. For CORDS to work, we require a common way to treat opportunities and resources that are represented differently in different data sources.

[1] The Institute for Research on Public Policy has published a paper titled, "Mapping Canada's Training Ecosystem: Much Needed and Long Overdue".This paper is a call to action for employers, training providers and government agencies of all levels to work together to lay the foundation of a robust pan-Canadian mapping of training and employment opportunities.

How can we make recommendations across unrelated taxonomies?

To leverage existing data pools with opportunities and resources already structured using different taxonomies, CORDS would need a way to relate information in one pool to information in another in order to give cross-taxonomy recommendations (e.g., someone looking for a vocational program, may also be looking for trades equipment donations).

Approach	Details
Map inferences derived from one taxonomy to another	**Description** We have a mechanism to be able to relate opportunities and resources together to bootstrap cross-taxonomy recommendations. With known linkages between records in one taxonomy, we can use natural language comparisons between the records in that taxonomy and those in a new taxonomy to map those linkages to the new taxonomy. 211 has provided us with one such linkage model. **Advantages & Limitations** This is untested. Theoretically, it would allow us to completely bootstrap new data into a recommendation engine without needing additional data collection about trajectories.
Recommend from collected trajectory data	**Description** By collecting data about the sequence of opportunities and resources a user navigates, we can develop a model of which elements we should recommend based on where a user is. As an example, collaborative filtering is an algorithm that can be implemented without much difficulty. Its input is data about previous pairings, then from new data, it can make recommendations. **Advantages & Limitations** CORDS would need to collect track user navigation which presents both technical challenges as well as security and privacy challenges. There is a high likelihood that this method will deliver better results than the bootstrapped approach.

Another approach would be to represent opportunities and resources in a single way.

What could be the common way to represent opportunities and resources?

With a standard way of representing opportunities and resources, we wouldn't have to make recommendations across data pools.

Approach	Details
Establish taxonomy	**Description** The taxonomy would be used in the classification, storage and retrieval of opportunity and resource records. Computations would be implemented based on the fields of the taxonomy to allow records to be located (e.g., searched for certain things like keywords, or related terms), sorted, and compared. This taxonomy could be created as a mapping between different taxonomies, an entirely new taxonomy which acts as a super set of existing taxonomies, or it could be an extension of an existing taxonomy. **Advantages & Limitations** Regardless of the implementation, this approach requires lengthy manual effort either for mapping or reclassification. Each taxonomy presently has users and a process (e.g., method of governing) for modifying the taxonomy. A new taxonomy would require the same elements, and possibly have interconnections to the governance of pre-existing taxonomies upon which it is based. This presents a range of challenges, and regardless of approach to overcoming those challenges, will require ongoing manual effort to accommodate change both in terms of governance and technically. The main advantages are that taxonomic approaches to data structures and algorithms are well established and easily conceptualized and understood (i.e., with little extra effort, the layperson can identify with the methods used to organize and work with the records).

Generalized quantization and analysis	**Description** In this approach, we would find descriptive content that exists in all opportunity and resource records, and apply generic data analysis tools to perform computations. Regardless of the system for representing and storing opportunities and resources, they are always represented with some accompanying descriptive natural language elements. These elements can be included as the path through the taxonomy used for the item's classification, or directly in the item's content (e.g., title, description, meta data). Natural language elements can be searched, compared and sorted. Text search is facilitated either by key term matching or vectorization (word embeddings). **Advantages & Limitations** The generality of these approaches requires minimal human intervention to determine relationships between opportunities and resources using equally generalized data analysis tools. These techniques can be more computationally intensive and require specialized skills (i.e., beyond those associated with current standard data structures and algorithms) to implement and maintain, but they do not require complex governance to manage how and what is classified. These techniques also offer the greatest flexibility in how data is accessed. Finally,

How can we access data?

Opportunities and resources are represented on the Internet in two main ways: directly via independent web properties (e.g., going straight to the website of your local food bank), and indirectly via listing services (e.g., Google, 211). The former may not be represented, or not well represented by listing services (e.g, listing services may have out of date information, or not prioritize presentation of a valuable listing). The latter typically provides links out to the independent web properties the listing represents, but in some cases the listing has no associated independent web property (e.g., the local food bank only has a location and telephone number). Therefore, CORDS needs to be fed with opportunities and resources from both listing services as well as independent web properties.

Approach	Details
Data partners	**Description** CORDS could rely on relationships with data partners which provide access to their data pools. Using any number of techniques, opportunities and records would pass into CORDS using a well defined interface that allows CORDS to translate the representation of records in the data pool into the CORDS equivalent representation. **Advantages & Limitations** CORDS would have to remain abreast of and adapt to changes in the way records are represented and accessed with each data partner. In both approaches to uniform data representation set out in the previous section, this approach to data access requires the least work - once the interface between a data partner and CORDS is established, CORDS has access to large pools of data. Data partners can be selected for participation in CORDS based on their subject matter expertise (e.g., 211 for human facing services, Magnet for job postings, provincial governments for accredited higher education programs), and requirements on data validation (i.e., procedures for establishing the legitimacy of opportunities and resources, as well as the appropriateness of records) could be among criteria for admission to CORDS that establishes end user trust in the content served by CORDS. Independent web properties not listed by a data partner, would not have any

	visibility. To make this options optimal for the end user, when an opportunity or resource has an associated independent web property in the data partner's listing, it would be best for CORDS to link directly thereto which could pose problems for the operating model of the data partner (e.g., if performance metrics or funding requirements require counting hits) - there are ways to manage this, but any of those would add some complexity to the implementation.
Independent web properties	**Description** CORDS could either, like a search engine, crawl the Internet and scrape web properties for relevant opportunities and resources to include in its recommendations, or a range of tools could be produced that allow independent web properties to interface with CORDS voluntarily. **Advantages & Limitations** In the case where the approach to opportunity and resource representation (see previous section) follows the application of a taxonomy, sourcing data directly from independent web properties is more complex. Where the web property owner is expected to apply some tool to their web property for participation in CORDS, they would have to manually apply the taxonomy to each URL (e.g., page) to classify its contents. In the case that web property content needs to be mapped to a taxonomy in a process of crawling and scraping, sophisticated automation would be required because of the scale of work. Although in both cases, the generalized quantization and analysis approach makes things easier, but there are other challenges. For instance, in both the case of crawling or volunteer participation, data validation (e.g., ensuring that content from independent web properties remains appropriate and accurate) poses technical, governance, and operational challenges. Sourcing data from independent web properties, would allow CORDS to ensure that opportunities and resources not captured or not well maintained by indirect listing services are properly represented, but would suffer the limitation of not capturing those opportunities and resources not associated with a web property.

What do we need to know about a user to help us to help them?

With the ability to uniformly represent opportunities and resources, we can develop a model to automatically compute their relevance to a user if, in part, we can quantify the user (i.e., establish and apply a "visitor taxonomy").

Although mass data aggregators ("big data") such as search engines and social network operators undoubtedly have extensive models for parameterizing visitors, these are not made publicly available. The most commonly used methods for collecting demographic information for analysis come from sociological studies such as a national census. These are effectively taxonomies and there is no gold standard. The following, for instance, are the topics of inquiry for which respondents to MENTOR Canada's national surveys (administered by Social Research and Demonstration Corporation) can select from several options:

- Age
- Postal
- Living Location
- Indigenous
- Ethnocultural/Racialized Population
- Newcomer
- Gender Identity
- Trans Umbrella
- Sexual Orientation
- Disability
- Disability Diagnosis
- In Care
- Relationship Status
- Dependents/Household
- Household Income
- High School or GED
- Further Education/Training
- Highest Education
- Employment Status
- Industry/Sector
- Specific Roles

Because the topics have several fixed options from which the respondent may choose, the responses can be parameterized, quantized, and compared. Therefore, when we establish the relevance of an opportunity or resource to a given set of user parameter values, we can assess the likely relevance of that opportunity or resource to another user by comparing their parameter values.

How could we acquire data about the user?

Below are a few examples of how we can gather information about a user. Each strategy differs based on the amount of information it requires us to collect and how that information can be used in the wayfinding process.

Approach	Details
Understand the user based on their needs or their initial entry point into the CORDS network	**Description** There is no universal taxonomy that classifies visitors that is currently used within a social impact context. Experienced call agents each have their mental models around providing referrals based on their experience. With this approach, it's less relevant to develop a taxonomy for a visitor, but more so to understand that someone who lands at x, usually needs to go to y. A navigation model would point users to different services, allowing them to indicate along the way if the recommendations are useful. Based on where the user lands, however, automated inferences can be made about other potential needs. There is no predefined model of what individuals may need based on who they are, but we can understand what people who landed in a specific spot also needed. **Advantages & Limitations** Without having to gather too much information about the user, this approach allows for broader recommendations but ones that are nonetheless informed by historical records and which can be augmented by machine learning to continually make them more relevant. Furthermore, the navigation journey can be refined through interactive prompts as results become differentiable. For instance, as the system discovers that the user needs food, if local options are differentiable along vegan / non-vegan lines, that differentiator can be presented as a prompt to refine results and inform further inference about the user's needs. This without explicitly asking the user whether they are vegan.

Ask the user each session	**Description** CORDS can deploy an automated dialogue that prompts the user with a series of questions that determines what the user is looking for and whether or not recommendations from the system are helpful. This information would be requested at each time a user begins a navigation journey starting at a CORDS-enabled web property. The benefit of this approach is a reduction in risk associated with privacy breaches, as well as reduced storage. The tradeoff is diminished ability to capture longitudinal data about a user's journey through the opportunities and resources. **Advantages & Limitations** This will have to be informed by further engagement with experts to glean what needs to be discerned about a client to make useful references and map the possible responses to services within the CORDS data pool.
Purchase data	**Description** Examples of data providers include Facebook and Google. We would be required to place javascript tracking on digital assets which would make web requests to third-party, non-CORDS network servers, and those responses would inform us about the habits of users exploring social services. **Advantages & Limitations** Purchasing data is a cost and requires us to tell our data suppliers about what URLs the current user is on. This could be a privacy breach and there has been litigation for similar practices in the UK.
Ask the user once, keep an account for them	**Description** Provide a service that relies on an account. This could be storing their history of searches, selections, viewed results, self-identification, etc. It would likely be personally identifiable information and so we would need to ensure the highest

	standards in security. This can be accomplished simply without requiring passwords but using “Single-Sign-On” technology AKA OAuth2, which lets users maintain an account through, for example, their Microsoft, Gmail, GitHub, Facebook, etc. The minimum information that we would be required to provide these 3rd-party providers would be when sign-on/up occurs. This service encourages keeping the account over a long period, which allows us to collect data about the longer-term trajectory of youth. **Advantages & Limitations** A big advantage is that we would be able to keep track of users over a longer period. A challenge with this is delivering enough value in having an account to get the user to maintain one. This also exposes a range of data risks related to security and privacy that would have to be mitigated.

Is it a good idea to leverage user profiling to provide recommendations?

There are scenarios in which information about an individual improves the efficiency and accuracy of their referral to opportunities and resources; based on where a user is, we can imagine how we might better predict whether they are in the right place or where they need to ingo if we know a bit about them.

> **Example**
>
> Similar to the scenario above, a youth hears that a local agency is registering individuals in the area for an arborist apprenticeship program. A staff member managing program intake completes their registration but also asks them
>
> The front line staff can infer, of course, that the youth is interested in the program because they showed up to the office, but also verifies their age group and asks them if they identify as an underrepresented group (e.g. an Indigeneous person, a newcomer, a person with a disability). The staff also knows that there are prerequisites for confirming the registration, such as the right clothing and safety equipment.
>
> Based on the responses given by the applicant, the staff refers them to a provincial program that gives youth a clothing allowance for starting a new job or training program as well as an additional stipend for Indigenous youth.
>
> By being able to discover other attributes about the person who walked into the office and to ask questions, the staff relies on their experience of other services that were useful to anyone who previously applied for the program.
>
> If we're not able to safely and securely ascertain key bits of information the user's journey is likely to end where they start, in this case by finding and applying for the arborist apprenticeship program, which leaves additional benefits aside.

However, user profiling is prone to reinforcing prejudice built into the data upon which the recommendations are modelled.

> **Example**
>
> A person from a marginalized group, let's say a refugee from a country without English or French, might already speak English or French perfectly. If a system maps the needs of a user strictly to their identity, there is a possibility that recommendations become too narrow. If a system assumes that the primary needs of the user are based on the fact that they are a refugee rather than their activity (i.e. the fact that they have landed on a job search site and

are actively interacting with the site in English), we risk leading the user to irrelevant resources.

With careful engineering, mechanics for identifying and removing prejudices built-in by modelling data, such as the application of relevant user feedback about the accuracy, appropriateness, or validity of the predictions, can, over time, remove those biases. However, those corrective measures can suffer from bias as well, and where they rely on user feedback, the process can be slow going as the system alienates those users who are the subject of the prejudice and from whom post-alienation feedback is required (i.e., we have driven away many of the people we need feedback from).

Where CORDS is intended to improve access to resources and opportunities by marginalized populations, establishing and using an explicit visitor taxonomy runs the risk of propelling or furthering systemic disparities or injustice inherent in the data that would relate classified visitors to the content.

How do we make good recommendations to users?

In the pursuit of a given objective, a user may need to access opportunities and resources from entirely different "categories". They may register for a program in one place, but need to be prompted to secure financial aid available in an entirely different place. Based on where a user is, how can CORDS provide recommendations for where they may want or need to be?

Approach	Details
Manually coded recommendations	**Description** We could create a database of "If This Then That" for relationships between parameterized opportunities and resources. When a user is observing one record, we would then be able to present other records following pre-established relationships between records. **Advantages & Limitations** This concept implies the use of a taxonomy and manually coding relationship between data, which both limits the options for data representation, and engenders a great deal of manual work. This is the approach taken by most operators of data pools (e.g., those entities that might become data partners) and has a high degree of reliance on the quality of the logic applied to the inference. It also requires periodic manual updates to ensure the relevance of the applied logic.
User feedback	**Description** When users are presented with CORDS recommendations, they can be offered an opportunity to indicate whether or not a recommendation was relevant or useful. The feedback would inform such things as ranking algorithms or machine learning processes that modulate or revise the type of recommendations offered based on that feedback. **Advantages & Limitations** This is good for keeping service providers accountable and could help uncover when links have gone stale but could be abused with fake feedback, either positive or negative.

Leverage taxonomy-user historical traffic	**Description** Data Partners like 211 can provide metadata that describes what services and resources users accessed in conjunction. This data can serve as a foundation for recommendations that can be presented to users based on their initial search. **Advantages & Limitations** By leveraging known user interactions, we will be able to help users find likely-related services based on common interactions.
Quantitative techniques, including machine learning	**Description** By using bayesian recommendation, collaborative filtering, or reinforcement learning, we can program the system to improve itself over time. This is a process we can automate rather than having to periodically take on a huge manual amount of work. With quantitative techniques, we can leverage user traffic to predict where users will subsequently end up. Some kind of loss in precision would be ideal to group services together so that the system doesn't create too much momentum with one provider. Truncation techniques include shortening the precision after the vectorization process, clustering ahead of time, principal component analysis and then reducing the precision on the vectors. Reducing the precision would group entries. With a little work, this can all happen programmatically, without relying on any person's labour. **Advantages & Limitations** This approach requires less manual intervention and is data-driven. Although it doesn't suffer from overt human bias in the design of the algorithm, it may inadvertently capture biases in data (e.g., there are more people of type 1 and they tend to take path A to B whereas people of type 2 who are fewer in number take path A to C, therefore the algorithm is more likely to recommend the less relevant first path to type 2users).There

	are ways to mitigate this by blending these techniques with user feedback.

How can we best track users while respecting their privacy?

CORDS' value is really in its ability to accurately recommend other opportunities and resources than the one a user is currently exploring. As these opportunities and resources evolve, and as the character and nature of user wayfinding changes in response to the global environment (e.g., what we look for during a pandemic may be different than what we look for in more "normal" times), CORDS will need to adopt techniques to constantly improve the model of recommendations so that the automated process offers up accurate outputs. To do this without profiling the user as the basis for prediction, we would need to track user journeys across sessions to gather data that can inform and update the recommendation engine.

Privacy is well-legislated in Canada and so we have some guidelines about storing tracking information on a user's computer. We have options concerning how to track users to remember what their selections are, which we need to improve our future recommendations.

Approach	Details
Expiring cookies	**Description** Cookies are data that is stored by the webserver on the user's computer. They can have an expiry set to preserve privacy. The server sends a response to the browser from a request that includes a session token that is stored in a cookie with an expiry. That session token would be used **Advantages and Limitations** Every few years there is a vulnerability discovered with cookies that presents a security or privacy risk.
Expiring JavaScript application memory session token	**Description** Storing data in temporary memory and not on the hard drive. The browser stores memory separated by tabs. When a tab is closed, that memory is deallocated and disappears. The server can respond with values to store in the relevant tab. Those values can be clear by the browser manually or in response to a reaction from a response from the server.

	Advantages and Limitations This method is the most secure and still performs well.

How should the system be hosted?

Data can be hosted and processed in a centralized server, it could be hosted and processed by data providers, or it could be outsourced to the blockchain. Today, most memory and processing are integrated.

Approach	Details
Decentralized	**Description** Blockchain technology is an example of decentralized storage and computation. There are other, similarly architecturally complicated, decentralized options. More simply, each data provider could store and process data for us (e.g., the taxonomic classification of a record, or its vectorized representation). **Advantages and Limitations** Depending on the manner of decentralization, the overall operating costs could be very high (e.g., blockchain is actually very expensive to operate). The complexity and cost of decentralized approaches increases with measures that need to be taken to overcome the challenge of ensuring data validity. Where storage and computation are offloaded to the data source, what exactly those sources are required to do will have to be very carefully considered. For example, most websites for community organizations are hosted on low cost shared servers that could not accommodate major increases to computational load or data storage. Another drawback is that this approach relies on the data source operator to perform regular updates to software, not to mention the implications for the software developers in supporting a wide range of target operating environments. An advantage would be the collective nature of the network which may engender improved resiliency.
Centralized	**Description** CORDS' server would receive necessary updates from data sources. This would be stored on a centralized server, where it would allow for an immediate edit of applicable services.

	Advantages and Limitations It's a very simple data architecture but requires data updates whenever data partner inventory is changed. This can be automated. Ongoing monitoring and maintenance of the system is required.

How do we ensure the quality of the content in CORDS referrals?

CORDS needs a mechanism to also ensure that the data being shared with users is safe and appropriate, and serves their needs. Every piece of content served by CORDS needs to be run through some kind of validation framework. But who performs the validation?

Approach	Details
Validation at the source	**Description** This is a more "decentralized" approach. CORDS could require data sources to implement a validation framework. This could be in the form of confirmation of compliance with certain criteria at the time of data creation or update, in order for the record to be admissible to CORDS. With data partners, it could be confirmation that all records in a data pool have undergone some form of validation. **Advantages and Limitations** As a decentralized process, the validity of CORDS outputs to end users would rely on the integrity of third-parties to perform record validation which injects a range of risks including compromise of the third-party as a result of ill-intentioned internal action, crime, or negligence. CORDS could rely on the expertise and resources of third-parties to establish subject matter appropriate validation frameworks and the means to apply the framework. With data partners, this approach decentralizes the governance of subject matter expertise therefore leveraging and supporting existing, sector specific initiatives to surface opportunities and resources. Technically and operational a decentralized approach is much more simple to implement than a centralized approach as the data provider is responsible ensuring validity of content not only on creation, but following update or change.
Centralized validation	**Description** CORDS would implement the data validation framework centrally and all pieces of data flowing through CORDS would have to be subject to data validation in a manner that ensured

	any changes to contents at the data source result in re-validation. **Advantages and Limitations** If performed by CORDS directly, both expertise and criteria would need to be established and maintained in the subject matter of the various areas in which opportunities and resources exist (e.g., education, employment, mentoring). Additionally, this would require significant work and resources on both the CORDS and data source ends as each time data was to be sent out from CORDS, in order not to have to rely on the third-party to notify CORDS of a change (which would obviate the benefits of centralized control), CORDS central would have to re-fetch the data from the source and re-validate it. Although this process could potentially be automated, eventually someone will figure out how to circumvent the automation if they are so inclined.

Recommendations

Based on the considerations above, what follows sets our recommendations for the development of CORDS. In brief, we are recommending using modern quantitative techniques to integrate data from various taxonomies, to implement an interface that can be present on all data partners' sites, and for a server to facilitate wayfinding by recording users' passages through the web properties.

Build something new that integrates across sites

There is a special opportunity to build something that unites the efforts of well-meaning "co-opetition" (cooperative and competitive) agencies. By creating a platform that integrates data partners' sites and other co-operative agencies, we can improve experiences.

Data partners to identify and validate opportunities and resources

There are active tools that have used taxonomies to link users to resources and services across multiple networks. Users can search for services based on a variety of categories, including location, language, and target groups.

These existing initiatives can act as data partners. CORDS can leverage their ongoing work as subject matter experts who curate data pools of quality opportunities and resources to ensure the validity, safety and security of CORDS outputs, while not having to rely on direct uptake of technology by independent web properties for CORDS to work.

> *Leveraging data partners allows for strategic investment in the development of subject matter expert initiatives that, within their domain, set the quality standard for opportunities and resources and have a vested interest in surfacing them.*

Trust in CORDS outputs via data partner requirements

In order to establish trust in recommendations output by CORDS, data partners would be required to demonstrate implementation of defined procedures for regularly verifying the validity, safety and security of opportunities and resources which it feeds into CORDS.

Trust in CORDS outputs is established in the chain of trust in data partners applying defined procedures for validating and ensuring the quality of the opportunities and resources they curate for their sector.

Move beyond taxonomies with generalized data techniques

CORDS doesn't need to create a new taxonomy, or merge, map or expand existing ones. Instead it will use modern quantitative data representation and analysis techniques to extract information about opportunities and resources and apply a generalized method for "classifying" them.

Google has open-sourced their 16-language vectorization model and code. Using this, we can produce consistent vectors from any of 16 languages. We store one vector per opportunity or resource. Incoming queries (searches or site descriptions) are vectorized using the same engine, and then used to search closest vectors and categories.

Although CORDS will need to perform work to establish interfaces for getting at opportunities and resources managed by data partners, the flexibility of this approach allows CORDS to easily adapt to the different ways that opportunities and resources are represented by data partners, and to relate those records in very powerful ways.

The application of quantitative techniques to develop an "algorithmic taxonomy" relieves the need for potentially complex governance of data standardization that is required in managing taxonomies, as well as removing inherent biases in the way traditional taxonomies are used to establish relationships in the data they represent.

"Bootstrap" recommendation engine from known user journeys

Using known relationships between user journeys through the opportunities and resources in existing data pools (see Appendix B for proof of concept), CORDS can establish an initial "recommendation engine" to propose other opportunities and resources based on where a user presently is.

Evolve recommendations through machine learning and feedback

Along their journey, explicit feedback about the utility of CORDS recommendations as well as their choices along the journey can be used to refine recommendations.

Each user journey will allow CORDS to learn how to better recommendations on the next journey CORDS facilitates.

Track the journey not the user

A zero-user-knowledge referral engine allows CORDS to make recommendations that are informed by data and historical trends, which can also be augmented and improved over time, without leveraging specific demographic or private information that raises legal or privacy concerns, or that is vulnerable to systemic biases that may be present in initial modelling data and a visitor taxonomy. In addition, while the recommendations that we can make based on what the user selects and their interaction with a zero-user-knowledge navigator will not be as specific as the ones made based on a user's demographic profile, it prevents CORDS from discriminating.

Instead of using cookies, we will use a session token stored in the JavaScript memory of the site, which will be randomly generated text, mapped from an IP address in an expiring cache. This way, the users will be able to navigate within the CORDS network and while their IP address stays the same, their session will be temporarily stored so that they can see their saved services. As well, the user can navigate away from the CORDS website and there will be no trace on other websites.

> *Tracking user journeys is very important to both gaining insights that allow algorithms to improve and better serve the user, and this can be accomplished in ways that anonymize data minimize privacy concerns.*

Centralize data operations only after collection and validation

Data collection and validation is effectively decentralized by passing that role onto data partners. All other data operations will be centralized to

- ensure quantitative and analytic techniques are properly implemented and maintained,
- maintain control over the implementation of security and privacy requirements (especially as they relate to user data even if anonymized), and
- facilitate scaling.

> *Decentralized data selection, input, and validation, removes a host of organizational, operational, and governance issues CORDS might have faced (e.g, governance of taxonomies, funding competition, establishing subject matter*

expertise for data validation). Centralizing post-validation data operations allows the initiative to focus on its core objectives in a well-controlled manner.

Leverage data partner paradigm to catalyse community uptake

CORDS will not require the rebuilding or reconfiguration of current services or the development of an entirely new platform or system. We recognize that providers across the country are operating with varying levels of capacity, staff time, and technical skills. CORDS will be built to ensure that a wide range of publicly funded employment service providers can implement and adopt the tool through their website or existing platform.

Uptake of CORDS will be driven by data partners across Canada. Data partners curate and maintain data pools that CORDS can leverage. This has the advantage of not interfering with established modes of data quality assurance and ensures that the content provided by CORDS can be trusted. CORDS, by driving traffic equally to and from various data partners, will not be a competitor to these existing services, but a tool that augments the experience of their users and increases the visibility of the opportunities and resources they manage.

CORDS will leverage data partners as the quality assurance engine confers the trust they place in their data to CORDS. That trust, in turn, can be presumed by the consumers of CORDS output whether other data partners, youth navigating CORDS enabled web properties, or web properties headlessly (i.e., without the interactive search components) displaying CORDS results (e.g., the Youth Digital Gateway).

This solution enables the validation of direct community (i.e., non-data partner) web properties to implement CORDS without maintaining an independent trust registry: a community web property endpoint led to by CORDS will only be a destination if a data partner has already validated records from that web property.

Although the development and distribution of a range of tools to facilitate the implementation of CORDS on independent web properties will be necessary, the use of generalized quantitative techniques in representing content alleviates the need for operators of those systems to do work beyond a simple installation. In order for the tool to work, however, the web property will have to be admitted as a listing in the relevant data partner's data pool. This mutually reinforces the work of data partners in surfacing opportunities and resources relevant to their sector.

To further motivate or catalyze the broad implementation of CORDS, funding for an initiative could include a required deliverable that a data partner interface with CORDS, or that independent web properties be admitted to the listing service of a data partner. This has the added advantage of implicitly applying an implicit data standardization regime across funded initiatives.

Stay open source to avoid unnecessary competition and conflict

Keeping the project open source[2], limits the potential for conflict and competition arising from ownership or control of intellectual property. It also encourages participation in the project based on its intrinsic value, and participant interest in the availability of the product.

Run CORDS through a collaboration of key data partners

The key element of the CORDS system in terms of governance, therefore, is the question of what entity might oversee the system's production and operation, and determine which data partners may be admitted. Built and operated as described herein, CORDS doesn't require enormous financing to operate, nor a large staff, nor a lot of infrastructure. As such, a well defined collaboration of key data partners interested in CORDS as a product that improves their ability to fulfill their missions, should be sufficient for governing and operating CORDS.

Operationally, staffing can be accomplished through collaborating partners. Financing for staff, infrastructure (rented servers), and any materials can be achieved through joint or independently obtained grants, in-kind contributions, or direct voluntary financing.

Governance and operations mechanics are beyond the scope of this report, but one would presume that the more open the system and inviting of sharing and scrutiny we are, the more robust the product will be.

[2] The source code for the current CORDS prototype at https://cordsconnect.ca is publicly available on GitHub at https://github.com/mentoringcanada/cords

Conclusion

Based on our exploration of current systems, and the possibilities offered by appropriate investments in technology, CORDS can be built and made to serve youth and service providers across Canada fairly quickly, fairly cheaply, and likely with positive impact. An example of how such a project could unfold in stages is presented in Appendix D.

We look forward to continued discussions with ESDC on achieving our common goal of better serving youth job seekers across Canada.

Appendix A: Youth Experience and Referral Data

A series of questions has been crafted for youth and call agents to gather intelligence and inform the function of CORDS.

Youth Perceptions of Navigating Services Ecosystem

Youth were asked about how they discovered support services, how they decided what training to pursue, whether they were working, how they found jobs, and whether resource/opportunity workers were referring them to others.

The following summarises questions that were posed to youth:

1. How did you find the support service that you interacted with?
2. Did you take a long time to consider which resource to reach out to? If yes, approximately how long?
3. When you found the service, were you referred to any other services?
4. Looking back, were there any other services that you could have used?
5. Were you referred to those other services in one way or another?
6. Have you ever heard of 211?
7. In your opinion, is there anything else that could be done to improve the reach of support services? If so, what?
8. Is there anything I missed that you would like to let me know about?
9. Do you have any questions for me?

Youth, whose information will be kept private, were identified and contacted for interviews through a variety of organizations including:

- Roots Community Services
- First Book Canada
- Unity Charity
- Desta
- Keep6ix
- ACOMI
- YouthRex
- Dream Legacy
- Pathways to Education
- Boys & Girls Club

- NPower Canada
- Canadian Council for Youth Prosperity

In some communities, it takes years before a youth trusts a service enough to reach out. In some neighbourhoods, outreach programs to 13-year-olds are too late. In other communities, they reach out and call hotlines often because they trust but feel underserved. Referrals were rare, which indicates minimal deadweight for a CORDS.

Some youth feel that jobs with small responsibilities are easy to find. Others couldn't find gainful employment and that they were unaware of their rights or scared to speak up for them in case they lost their job. Grassroots clinics were described as community-specific and not publicized sufficiently. Secular and non-secular food banks don't always communicate. Scholarship resources are sometimes based on competitions that can be seen as time-intensive.

Service Providers

Service provider staff and volunteers were asked about their service model, the audience they serve, and how they receive and give referrals. There was a large divide in answers. Much is covered in the above section. Their responses have shaped this document. Most notably, Kim Cooper of Youth Culture recommended a social network for youth as a feeder network to LinkedIn.

Appendix B: Prototypes and Preliminary Results

Non-Semantically Related Recommendations

We can leverage past data to deliver augmented recommendations to users today.

Findhelp Information Services provided us with 211 Central data that links an anonymous identifier representing a caller, with referrals made by a human phone agent to service listings within the 211 database.

The necessary data was just 2 columns: ID and Taxonomy. Using an input of a single taxonomy category, we can suggest the most likely another taxonomy category, using statistics. Below are some examples.

The results below can be interpreted as follows:

> For a given **input**, we output:
> [# of coincident services, 'Recommended Service']...

Results

If a caller needed **Food**, we output:
[190, 'Housing/Shelter'], [151, 'Material Goods'], [141, 'Individual and Family Support Services'], [104, 'Utilities'], [82, 'Temporary Financial Assistance']

If a caller needed **Housing/Shelter**:
[248, 'Food'], [151, 'Information Services'], [110, 'Individual and Family Support Services'], [106, 'Utilities'], [96, 'Temporary Financial Assistance']

If a caller needed **Employment**:
[33, 'Housing/Shelter'], [30, 'Food'], [23, 'Individual and Family Support Services'], [18, 'Information Services'], [16, 'Mental Health Assessment and Treatment']

If a caller needed **Community Recognition**:
[62, 'Social Insurance Programs'], [49, 'Information Services'], [29, 'Public Health'], [28, 'Legal Services'], [20, 'Community Economic Development and Finance']

If a caller needed **Judicial Services**:
[13, 'Legal Services'], [9, 'Legal Assistance Modalities'], [6, 'Law Enforcement Services'], [6, 'Information Services'], [4, 'Legal Services Organizations']

Vectorization Prototype for Text Search

We have a text search prototype using vectors instead of substring search. This allows for “semantic search” which is more flexible and faster than key term search.

Vectors in the mathematical sense refer to magnitudes and direction. They can be represented also as points, just like latitude and longitude coordinates except in higher dimensions. There now exist vector representations of most words in over a dozen languages. These vector representations are semantically significant: “dog” and “cat” are close to each other but far from “house.” We implemented an open-source program that turns 211 Central’s “Agency Descriptions” into vectors that we can conveniently compare. The following shows the input prompt and associated output of this system:

Prompt → **Please briefly describe your circumstance and request.**

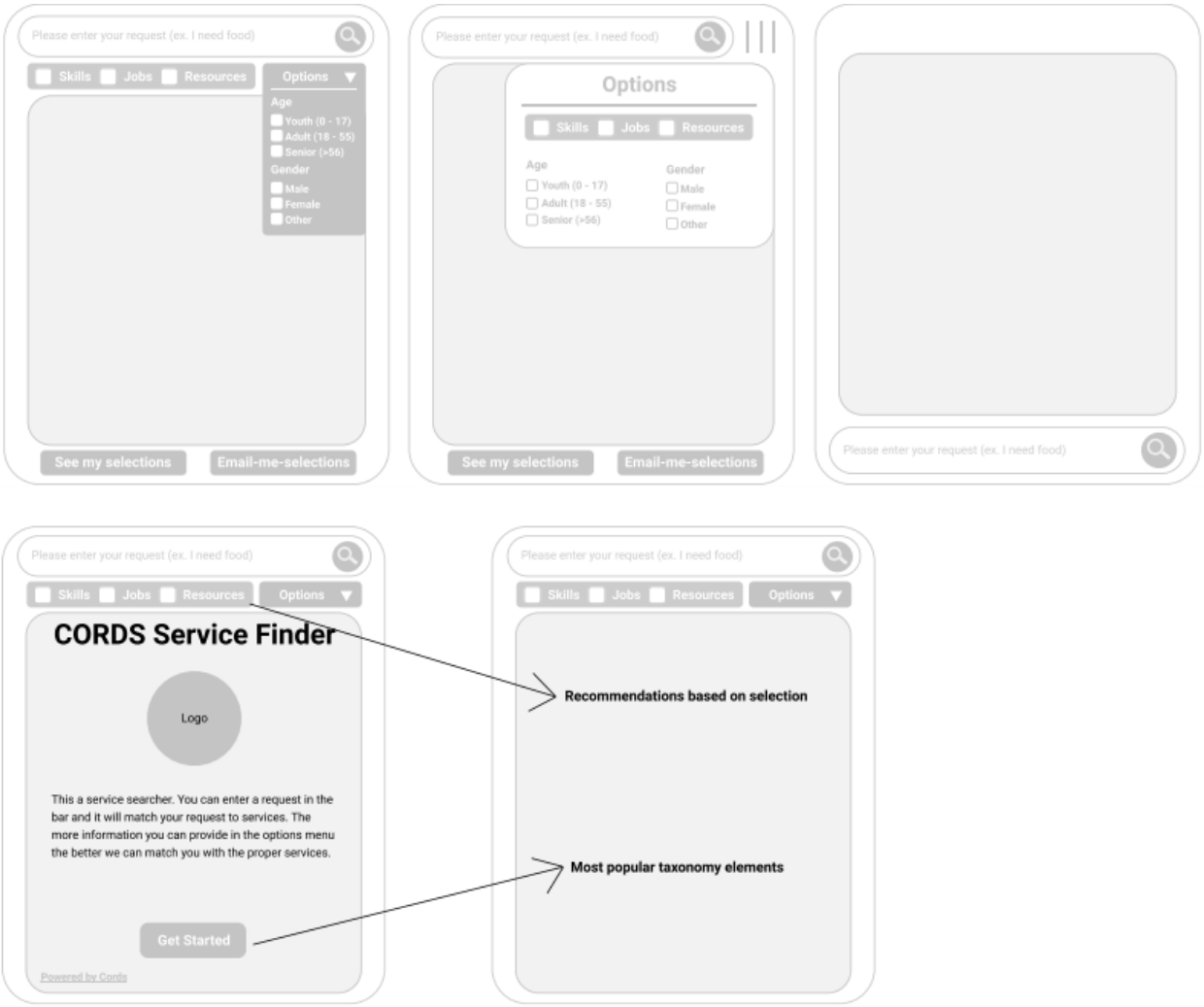
User input → ***I need clothes for a job interview.***

Output → **Dress for Success Toronto**
Women's employment suiting program provides professional work attire to women seeking employment. Virtual suiting, virtual styling and in-person appointments...

→ **New Circles Community Services**
Gently used and new clothing by appointment only...

→ **JobStart - Capability program**
One-on-one employment and job search assistance including customized workshops on enhancing interview skills and increasing self-confidence...

Appendix C: Preliminary User Interface Mockups



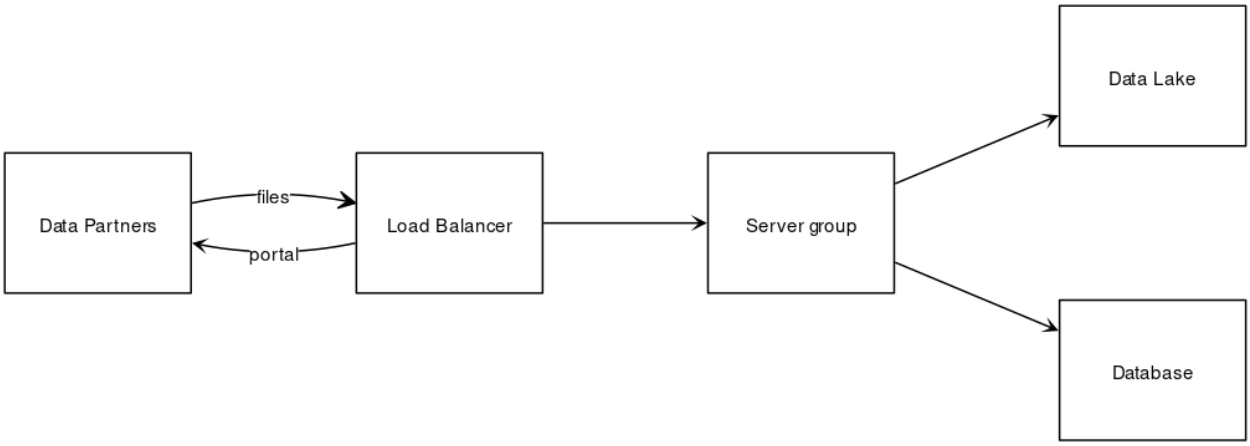
Appendix D: Preliminary Project Roadmap

The following sets out a sequence of project stages through which the recommended implementation of CORDS can be achieved.

Stage 1 - Searchable data lake

- Accept data exports from the current service providers, streamlining the process across services
- Success
    - Receiving data from relevant service providers in a format deemed suitable to proceed to stage 2.
- Risks
    - Duplication of data (can be mitigated)
    - Stale data (include timestamps - e.g.,.Findhelp Ontario/211).
- On the way
    - Categorize their systems by frameworks to facilitate Stage 4.

Infrastructure



Steps

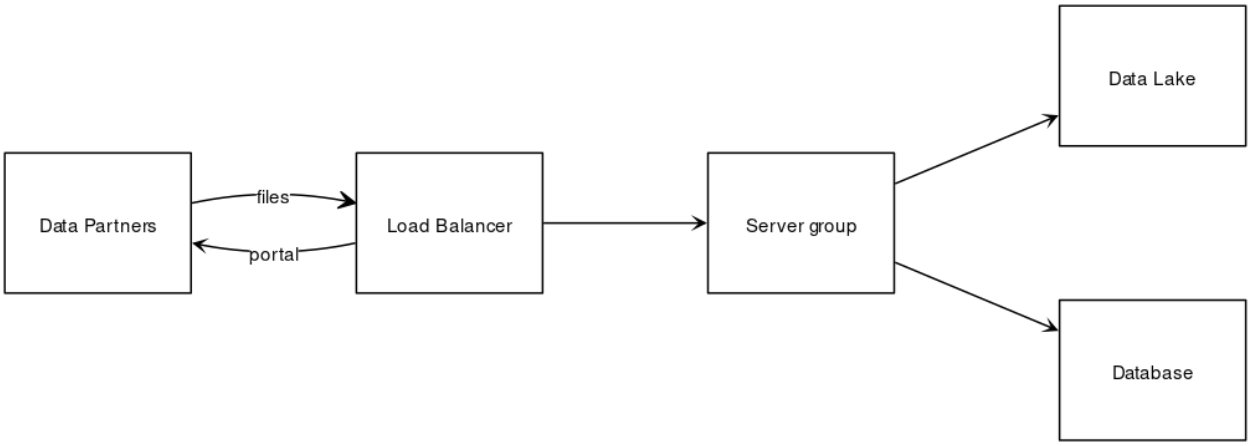
1. Working group to negotiate data format and frequency.
2. Secure API server to accept regular exports from various sources.
3. Data lake to store the regular exports, their attribution, and timestamp.
4. Process to remove duplicate and stale data.
5. Reporting portal or export to the data contributors for Quality Assurance and value-add in subsequent stages.

Stakeholder	Input	Activity	Output	Outcome	Impact
System Maintainers	Skills, time	Build and maintain	# of data contributors; # of data distinct points (resources/opportunities)	A data process to retrieve data from contributing organizations; A portal for them to see what they have contributed	A portal and process for data partners; a foundation for later stages.
Data Partners	Time, data	Contribution to a data process; Quality assurance on their data	A well-maintained resource to search and review their services		A portal and process for data partners; a foundation for later stages.
Project sponsor	Funding				A portal and process for data partners; a foundation for later stages.

Stage 2 - Searchable records and data navigation for data partners

- Offer recommendations of resources and opportunities (the two categories unlinked)
- Success
    - System to assess circumstances, needs, and matching opportunities/resources for youth.
- Risks
    - Poor recommendations (can be improved), inappropriate data
- Show multiple icons/logos per service on duplicated

Infrastructure



Steps

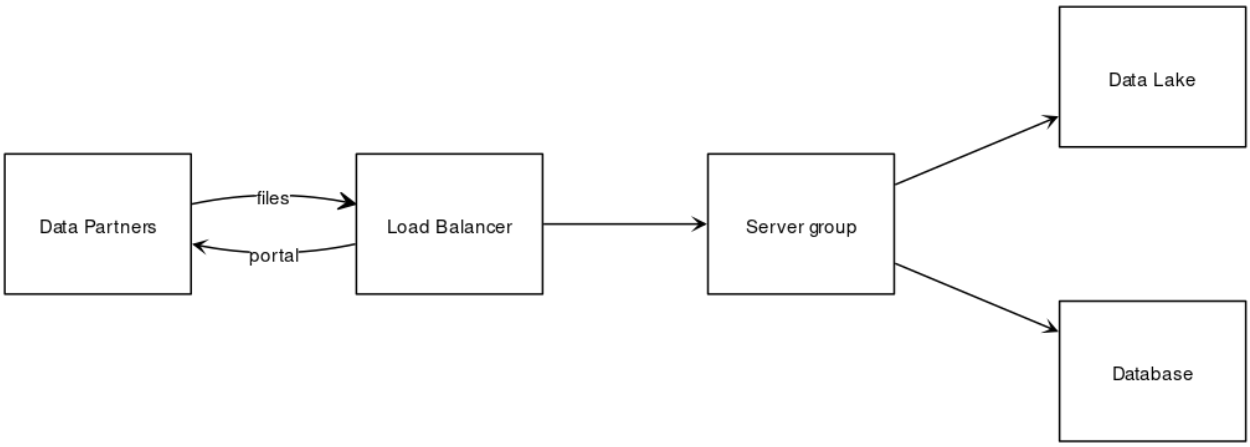
1. Searchable records through a secure API
2. Data analysis on how resource and opportunity usage
3. Analysis implemented into value-adding wayfinding algorithm
4. Expose the wayfinding algorithm through an interface for data contributors alongside their reporting portal

Stakeholder	Input	Activity	Output	Outcome	Impact
Youth	Calling a data partner	Talking to the service provider	Being referred to service	Improved referrals	
System Maintainers	Skills, time	Build, maintain, and grow system	# of participating data partners; metrics around their satisfaction	A more sustainable social system; more skilled and employed youth; improved life of youth	
Data partners	Skills, time	Same as today	Same as today	A platform with features at par across all service providers	

Stage 3 - Wayfinding for data partners

- Collect session data to identify and improve search trajectories
- Success
  - System can short-circuit trajectories to improve recommendations
- Risks
  - Privacy breach
  - MITM
  - Poor anonymization

Infrastructure



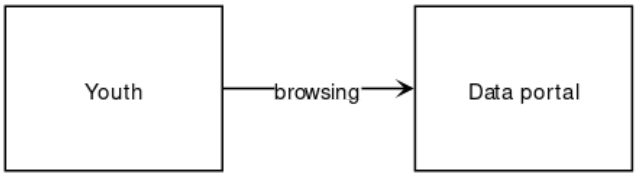
Steps

1. Collect data from an interface on how data contributors are using service
2. Analyze utility of the service
3. Implement analysis into improved wayfinding algorithm
4. Set up data engineering to automatically improve the algorithm based on data partner usage

Stakeholder	Input	Activity	Output	Outcome	Impact

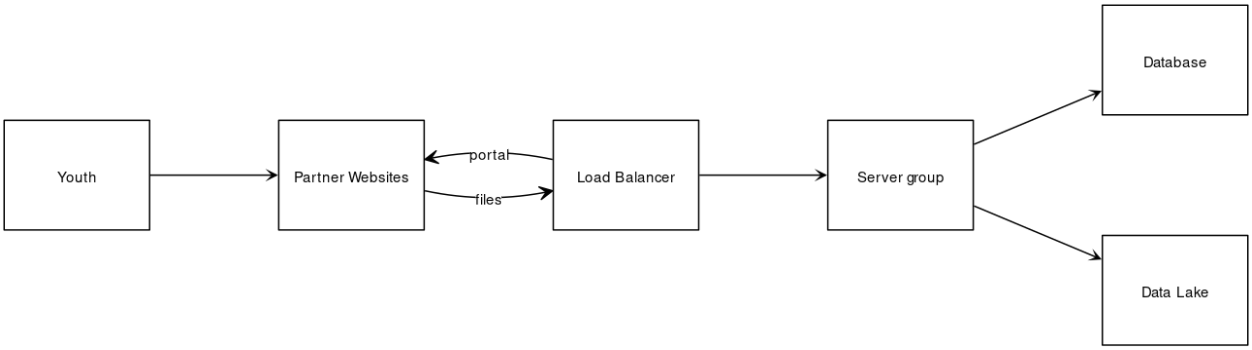
Youth	Calling a data partner	Talking to the service provider	Being referred to service	More complete referrals	
System Maintainers	Skills, time	Build, maintain, and grow system	# of participating data partners; metrics around their satisfaction	A more sustainable social system; more skilled and employed youth; improved life of youth	
Data partners	Skills, time	Machine-assisted wayfinding for service referrals	Improved ability to offer referrals for youth		

Stage 4 - Public wayfinding



- Expose an API/SDK to allow a site plugin to private-label the wayfinding with an onboarding process
- Success
  - Developing API for plugin with API-key
  - SDKs developer-assisted integration
  - Plugin for WordPress and additional platforms
- Risks
  - Poor UX design leading to poor adoption

Infrastructure

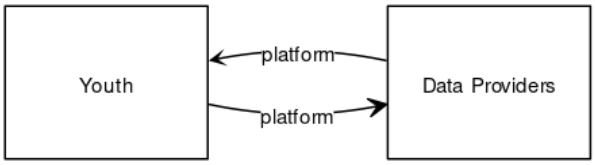


Steps

1. Prepare infrastructure for a heavier load
2. Create secure, open API and site-plug-in for sharing the wayfinding API with service providers for their public-facing infrastructure
3. Oversee endpoint access to assess success

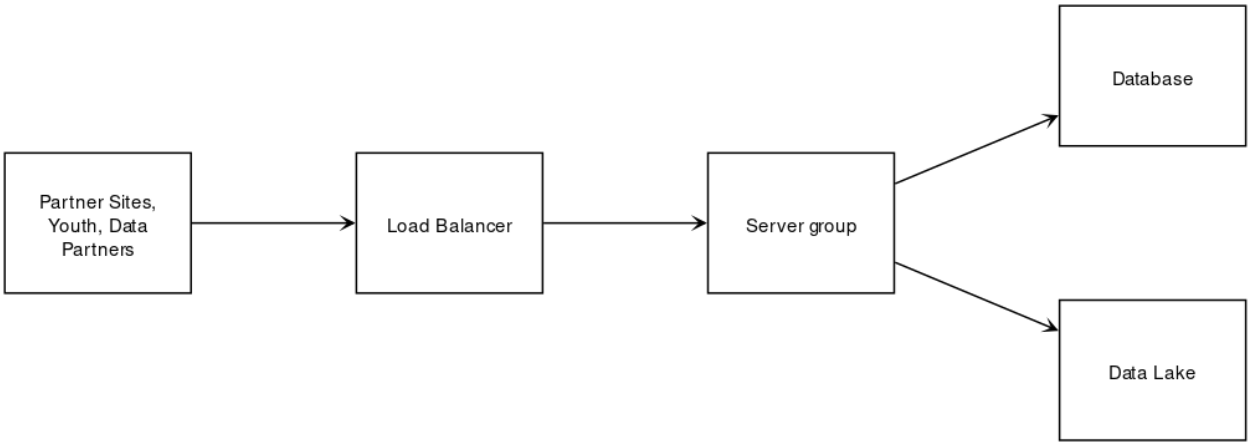
Stakeholder	Input	Activity	Output	Outcome	Impact
Youth	Visiting any data partner's website	Browse statistically-related services	Greater awareness of resources and opportunities	Increased likelihood of receiving appropriate services; Improved quality of information	
System Maintainers	Skills, time	Build, maintain, and grow system	# of youth participating; # of resources and opportunities available		
Data partners	Time, skills	Integration onto their public-facing site	Ability to service youth without interaction while still recording data		

Stage 5 - Social accounts



- Create a social network linking the sessions together in an account, to record longitudinal data
- Success
    - Engaging, long-term platform for youth and stakeholders
- Risks
    - Privacy breach

Infrastructure



Steps

1. Create a portal for new service providers to take on the same benefits and responsibilities as initial data partners/contributors that are hosted by CORDS (THIS IS WHERE TRUST COMES IN TO DECENTRALIZE)
2. Enhance portal to allow public profiles, SEO (searchability)
3. Allow resource/opportunity seekers to create accounts to track and improve understanding and ability to wayfind
4. Make it social (this is a big step that should probably get its own stage) to get seekers connected with opportunities via board and chat

Stakeholder	Input	Activity	Output	Outcome	Impact
Youth	Using a social less corporate social network to understand which skills and time should be acquired	Discovering various education/skills facilitate which employment opportunities and which supports they might need, given their queries and trajectory status updates	Youth becomes skilled, educated, and supported	Increased self-confidence; improved stability; ability to support self instead of relying on services; improved life	
System Maintainers	Skills, time	Build, maintain, and grow system	# of youth participating; # of resources and opportunities available	A more sustainable social system; more skilled and employed youth; improved life of youth	
Employers	Time	Posting opportunities and engaging with youth in a forum; potentially some overlap in employers and youth	Employees		
Support Service Providers		Enter contact information and services into the CORDS database.	Services	Services are shared with end-users across the CORDS network.	Quicker and more accurate recommendations for youth users.