DEVELOPING INTERNET DISSEMINATION IN STEP WITH CLIENT EXPECTATIONS

Pamela Best
Statistics Canada

Abstract

Effective Internet sites constantly evolve and change to meet client needs. This paper examines the driving forces behind the evolution of the structure and content of Statistics Canada's website and how they have affected content development, searching, design, navigation, and electronic commerce. The paper also touches on the impact that the Canadian Government On-Line initiative has had on the electronic dissemination of information by Canada's statistical agency.

Keywords: Internet, dissemination, market research, electronic commerce, usability testing, Government On-Line,

From Canoe To Flagship

Statistics Canada had an online presence even before the advent of the Internet. A small group of clients had access to file transfer protocols (FTP) for large files, and *The Daily* (the agency's official release bulletin) was sent out through a text-based gopher service and listserves.

When www.statcan.ca was launched in 1995—without any permanent funding—it was one of the very first Canadian government websites. Its major content was *The Daily*.

Six years later, the website has over 60 000 HTML pages and is one of the public's most important points of access to Statistics Canada, serving more than 14 000 clients each day. The website serves the general public interest by providing free data on Canada and Canadian society as well as information about the agency's products, services, surveys and statistical methods. It also provides an e-commerce gateway for paying clients to obtain selected products and services.

Statistics Canada's website was nominated a "Select site" by the editors of the Dow Jones Business Directory and was named the "Best Institutional Online Product" by the Canadian Online Products Awards. More than 10,000 other sites link to it.

However, client expectations increase and like other websites, www.statcan.ca is constantly evolving to meet its users' needs and expectations.

Discovering Users' Needs

The site's evolution has been driven by users of all kinds, including the National Statistical Council (the Chief Statistician of Canada's advisory body), our partners in the Data Liberation Initiative[1], the Depository Service Program[2] and the areas within Statistics Canada that produce the data and author the publications—and, most importantly, the general public.

[1] The Data Liberation Initiative consists of a community of researchers in Canadian colleges and universities that access public use microdata files and macrodatabases through the Internet site.
[2] The Depository Service Program consists of a network of over 700 academic and public libraries across Canada and throughout the world that receive Government publications in order to ensure that they are accessible by all Canadians.

We have an extensive market research program to assess the nature of our clients, their interactions with the site and their expectations. We do periodic market research using pop-up questionnaires on the entire site as well as specific segments of it. We conduct usability testing and observational research, and monitor e-mail and traffic logs. We work closely with the staff in our regional offices to make the website work for them and for the clients to whom they provide services.

In the beginning, we sought our clients' feedback to test their level of tolerance for change in this emerging environment; now, we use their feedback to evaluate how *we* have to change to meet *their* needs.

Necessity Was The Mother Of Automation

Statistics Canada releases new data and statistical products every day through *The Daily*, our online database services and publications. Because *The Daily* was the website's first content when it was launched, we began our Internet venture with a website that had to be updated at 8:30 a.m., without fail, every working day. *The Daily's* users—journalists, news agencies, policymakers, bankers, industrialists, consultants and the general public—needed to be assured that they would receive *The Daily* reliably and predictably.

The problem was that not everyone was receiving *The Daily* by Internet. We were still delivering it in print, by fax, by listserve and in a voice-synthesized format. We were providing it in two versions on the Internet as well: in hypertext mark-up language (HTML) and in portable document format (PDF). And we were doing all of this in two languages.

It soon became obvious that we would have to automate as much as possible if we were going to be able to keep the website current and *The Daily* delivered in all its forms, with a minimum of intervention. We solved the problem by developing a standard generalized mark-up language (SGML) application that allowed a single source to be used to feed the variety of formats and meet user needs. We learned very early on that automation would be essential. Today, *The Daily* accounts for 6% of the total traffic on our site. The listserve has more than 2,000 subscribers.

Electronic Publishing: From In The Mail To On The Web

Statistics Canada has always been concerned with delivering quality service in a cost-effective manner. The Internet provided us with the opportunity to curb the costs of our publishing program while at the same time enabling us to deliver our publications to clients in a timelier fashion. One of the more expensive aspects of our operations was providing paper copies of all our publication titles to the Depository Service Program libraries. If we could send some of these publications over the Internet rather than through the mail, we would save not only the printing costs but also the shipping and handling fees. The clients would not have to wait for mail delivery of the publication, they could get the publication on-line at the time of release. The first titles to be converted were those with fewer than 50 paying clients.

To be successful, we needed to know how comfortable the library community would be with getting these low selling publications on-line. We undertook a six-month pilot project that evaluated this preparedness. The results were reassuring: 83% of the participating libraries said that they were ready. Within two years, all Depository libraries were receiving Statistics Canada publications electronically via the Internet.

An Electronic Commerce Gateway

When most companies started their e-commerce sites, they sold hard-goods—CDs, books, cars—using an online interface. Statistics Canada had hard-goods (publications) but it also had data, specifically data from the CANSIM[3]

[3] The Canadian Socio-economic Information System now CANSIM II. It is the Agency's corporate data warehouse and is being expanded to become a repository for all publishable data in any output format.

and the International Merchandise Trade databases. So rather than setting up a shop to do e-commerce for hard-goods, we focused on the future—delivering data electronically.

In 1996 we set up another pilot project to see if our clients were ready to conduct electronic commerce transactions on the site in a "self-serve" environment for online databases. In the first full year of operation, we generated C$70 000 in revenue. By the next year this had doubled to C$155 000.

Realizing that the pilot was a success, we added a facility in 1997 that allowed clients to purchase publications online in PDF format. The following year, we expanded this service to allow clients to subscribe to these online publications. Now, total revenues from electronic commerce are over C$350 000.

Statistics Canada wavered in its stance on the visibility of the products and services that we sell on the site. At one time, we buried everything we sold deep within the site. However, the doubling of e-commerce revenues each year on the site from 1997 to 2000, caused us to re-think the importance of this aspect of the site in the eyes of our clients.

Clients told us in focus groups that they wanted to know at the outset what was available for free on the site and what was for sale. They also said that we should not be ashamed to use the site to promote our products and services. They recommended that we develop a "best seller" area on the site and suggested that we use the home page to more aggressively promote popular products.

High Visibility For High-Level Data

One of the missions of www.statcan.ca is to provide significant free or "public good" content and it was in this context that, in 1995, we developed a module called Canadian Dimensions (now Canadian Statistics). Initially, the module held 40 HTML tables containing national- and provincial-level statistics about Canada and Canadian society. The content was designed to be eminently usable: the data were structured and labelled so a general user could understand them easily and the tables were built so that they could be viewed in their entirety without scrolling (which imposed a limit on the number of columns and rows).

Canadian Statistics attracts roughly one in five visitors to the Statistics Canada website. The content of this area is always expanding; over 700 tables are now available in this module. It would be impossible to keep this many tables up-to-date manually, so here too we are applying the "automate" principle: new tables, with few exceptions, must draw their data from the CANSIM database so that they can be updated automatically at the moment the data are officially released. More than half the tables in Canadian Statistics are now automated.

We have kept a firm hand on the content's public-good value, however, by performing log file analyses on the tables to ensure that they are relevant to our users, and we have rigorously maintained its screen-friendly look and feel.

Thousands Of Hits On Canada's Thousands Of Communities

The Internet was the ideal medium to get the 1996 Census results out to the Canadian population. At first, hand-coded HTML tables containing summary information for Canada and the provinces were put on the site. This was fine for a few tables, but coding several tables for each of the 6,000 communities would have taken until shortly after the *2001* Census was completed.

We needed an application that would dynamically generate tables of information on population characteristics, work, families and dwellings for each community across the country. This application became Community Profiles and was added to the site in November 1998. All a user had to do was type in the name of their hometown and seven tables of data would be generated for that city, town or community.

Community Profiles was an instant success, attracting roughly 20% of traffic on the site within two weeks. The number of visitors to the entire site jumped from roughly 4,000 to 7,000 visitors a day.

The site log files indicated that this application was heavily trafficked but did not tell us if the Community Profiles module was meeting users' needs and expectations. We hired an external consulting company to evaluate users' overall satisfaction with the module and determine the extent of any navigation and interface issues.

The research showed that overall, clients were very satisfied. The language used in the module was easily understood and the data were easy to interpret. They found the information useful and the site easy to navigate. But they wanted more. Clients suggested that we add more data on other population characteristics such as cultural and aboriginal groups.

The Users: A Profile Over Time

While the foundations of Statistics Canada's Internet presence were laid between 1995 and 1998, by 1999 it became evident that the site needed to be more than just a collection of documents placed on the Internet in order to meet the Agency's needs. Some benchmarking market research had been conducted in 1997 that profiled our clientele. We conducted a second round two years later to update our users' profile, assess our clients' needs and behaviours and to find out how satisfied they were with the site.

The findings from these quantitative studies revealed that there had been a significant increase in the proportion of students visiting the site, up to 40% from 22% in 1997. The proportion of frequent users to the site had also increased in this two-year period by 10% to 34%. In fact, many respondents (29%) had bookmarked the site.

A Satisfactory Experience

Surfing Statistics Canada's website was overall an agreeable experience for our visitors. More than two-thirds of the respondents indicated that they were very satisfied or satisfied with the site. Most said that they would visit the site again (87%) or recommend it to others (82%). They rated the site well on the consistency of design, said the language was clear and easy to understand, and that there were no delays in using the site.

However, when we looked in more detail, we found that although we performed well in these areas, they were not necessarily the most important aspects of the site for our clients. Quadrant analysis revealed that the site actually had a low rating in a number of areas that were very important to our users, notably

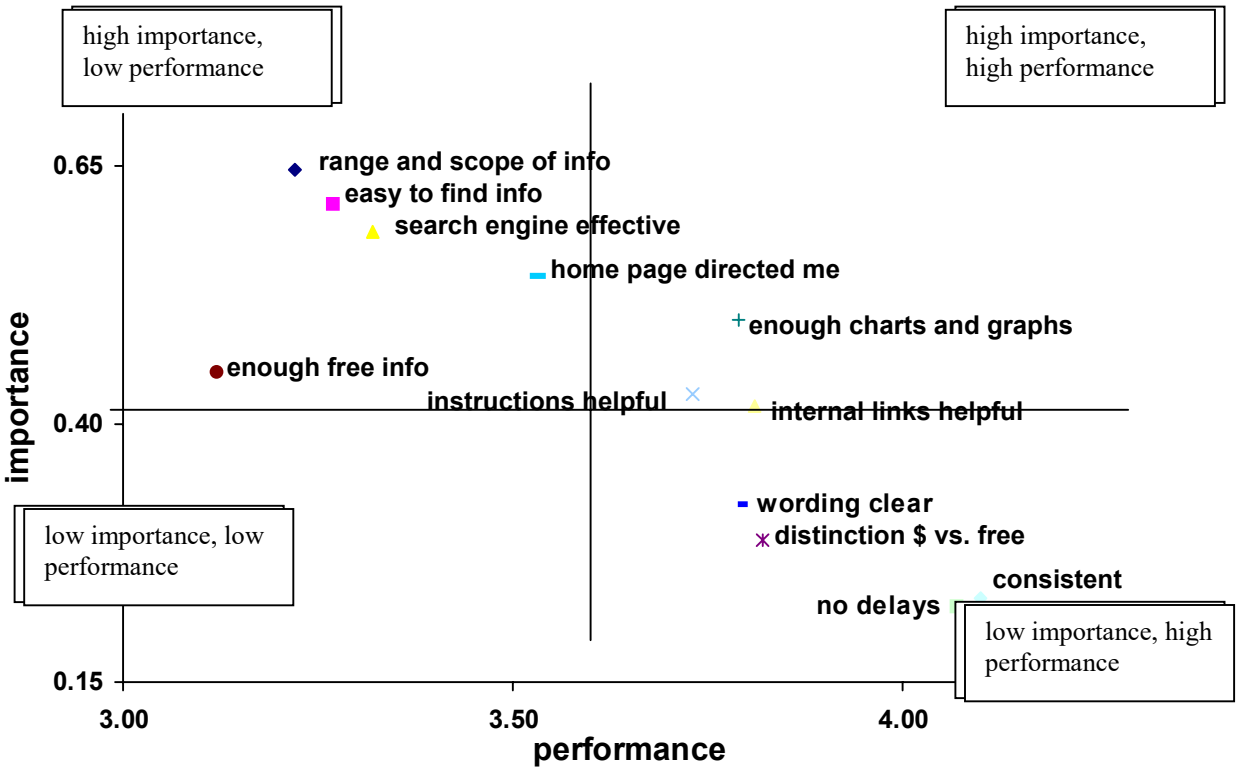
- the range and scope of information was too narrow;
- the information was not easy to find;
- the search engine did not return effective results;
- there was not enough free information; and
- the home page did not direct users to the information that they needed.

The Action Plan

Users had told us what they thought; now we had to make the improvements. An action plan was drawn up that addressed the recommendations from the research. The plan focused on

- content development—increase range and scope of free information;
- searching;
- navigation and design.

Figure 1: Importance-Performance Quadrant Analysis of Site Features



Source: Phase V Consulting, 2000

Increasing Public Good Content

The project to increase the range and scope of free information on the site was tackled on three fronts.

- We encouraged authors to put their national- and provincial-level data into Canadian Statistics module of the site.
- Community Profiles were expanded to include more than just census data. The number of births and deaths in each community was added along with data for over 130 health regions across the country. Eventually we hope to incorporate education and justice data in this module.
- There was also a move to make older versions of publications available on the site for free.

Of the three initiatives above, the idea of making older versions of publications available for free was perhaps the easiest to grasp and the hardest to implement. Author areas were very supportive of the initiative—they wanted their publications to be widely accessible—but there a was a need to establish a policy that determined how "old" a publication had to be before it could be free. We did not want the publication sales to decline and we needed a standard that could be applied in all cases. There was also a technical limitation that is now being addressed.

Searching Successfully

With upwards of 60 000 pages on the site, searching and retrieving accurate results is critical. When the market research was conducted in 1999, searching on the site was fragmented. *The Daily*, Canadian Statistics and the databases (the Online Catalogue, CANSIM, International Merchandise Trade database) were being searched using one search engine. Community Profiles used another. In fact, this application was running on a different server and

could not be found using a site search. The remaining areas of the site (Education Resources, Concepts, definitions and methods, Are you in a Statistics Canada survey etc.) were searched using a free search engine that was a legacy from the site's early days.

One of the first challenges was to integrate searching for all free content on the site—including the Community Profiles. We also needed a search engine that could search metatags in the HTML code. A new search engine was installed on the site in February 2001 that did all this. However, we had to exclude CANSIM and the International Merchandise Trade databases which use separate search engines.

Solving Users' Search Problems

Though we have made great strides in improving the search features on our site, clients are still having difficulty finding information. Last summer, we conducted research on our users' information needs and their success rate. Only 42% of the clients found all or most of the information they sought.

Although this survey was conducted prior to the installation of the search engine, it revealed that the difficulty that clients had finding information on the site went beyond search engines and search results. There were three facets to this problem:

- the information was gathered by Statistics Canada but was not available on the site
- the data was not gathered by Statistics Canada
- the data was on the site but could not be found using the client's terminology.

Managing Expectations

People will come to the Statistics Canada website to find out how to get a copy of their birth certificate or to get a weather forecast. We needed to develop a strategy to manage the expectations of our site visitors—this is the scope of the information that we *do* have and this is what we *don't* have.

We have added two modules on the site that outline the scope of information that we *do* have. A "First visit" section, which describes in narrative format the type of information on the site, and an A to Z index—essentially an alphabetical site map with links to the top three layers of information on the site.

We added links to information that is frequently sought on our site but is the responsibility of other arms of government. For example, we have added a link to the portion of the Government of Canada website where people can find out how to get a copy of their birth certificate—a service under the jurisdiction of each province.

The Terminology Quandary

Most large organizations have their own lexicon and Statistics Canada is no exception. The challenge is to bridge the gap between the way our clients describe the world and the way we describe the world. Statistics Canada is addressing this issue in a number of ways.

- We are developing a synonym list for our search engine so when a client enters the term "cost of living" or "inflation" they will get results for "Consumer Price Index" instead of zero hits.
- We are encouraging authors, wherever possible, to use "lay" terms. Now a client gets search results for "poverty" although Statistics Canada gathers statistics on "low income."
- We have developed a controlled vocabulary of 27 high-level terms and 222 secondary terms to describe Statistics Canada's information holdings. This subject list was tested with clients to ensure that the terms were clearly understood.
- Clients can now "browse" the site using the subject terms from our controlled vocabulary

Integrating Content

One of the problems that both clients and Statistics Canada staff had with the site was that the information was fragmented. There would be a news release in *The Daily*, a publication under Products and services, data in

Canadian Statistics and metadata in Statistical methods. This was frustrating for both authors and users because they often found only one small piece of the puzzle.

We are working on integrating the various elements of the site so that the puzzle is virtually complete: the user has only to find one piece to be linked to the rest. For example, a major release has a short summary on the home page that is linked to the full article in *The Daily*. *The Daily* article contains links to the publication it describes, which contains a full analysis of the data and a full set of charts and tables. *The Daily* article has its own tables and charts, but it also links to summary tables in Canadian Statistics, which are linked to the metadata for the survey. Analysts who want to analyse the full data set can follow the links to the relevant CANSIM tables and matrices given at the bottom of *The Daily* article or the tables in Canadian Statistics. Our goal is to automate as many of these links as possible so that little manual intervention is required to create and maintain the inter-relationships.

Redesigning The Site To Meet Users' Needs And Government Standards

Two things drove us to redesign our site: the results of the market research and the requirement that we implement new Government of Canada standards for the *Common Look and Feel for Internet*. The standards were developed to ensure that information is consistently presented on the government's many departments' and agencies' websites.

All Government of Canada Web pages must have the "signature" of the department or agency, the Canada wordmark and a black Government of Canada common menu bar at the top of every page. The menu bar contains five buttons: Language (English/Français), Contact Us, Help, Search and Canada Site (which links back to the Canada site at www.canada.gc.ca). Each department or agency must then include an institutional menu bar, similar in design and placement to the common menu bar, that describes the organization's programs and services. Statistics Canada analysed the site traffic logs and placed the site's most frequently accessed areas in the institutional menu bar.

Figure 2: Government of Canada menu bar and Statistics Canada institutional menu bar

The new Government of Canada menu bar changed our site's navigational paradigm, shifting the primary navigation tool from the sidebar to the top menu bar. It was a shift that enabled us to move some information up one or more levels—and respond much better to our users' needs.

For instance, when the site was evaluated in 1999, *The Daily* was buried three levels down in the site. When we redesigned *The Daily's* main page (which was now just one click away from the top menu bar and on the new home page), we were able to incorporate *The Daily* navigational page elements into the sidebar while freeing up the centre of the page for *The Daily* itself.

The users we tested had found our old home page cluttered, with too many choices and weak navigational value. We analysed the log files and listened to the recommendations from focus groups in order to determine which of the navigational choices on the home page were the most important to our users.

Our first objective in redesigning our home page was to make it easier to locate information, so we modified it to include direct access to data and the latest releases without forcing the client to click through several levels of navigational pages. It was the users who suggested that we include an article or a feature from *The Daily*. Now brief,

hyperlinked, summaries of the major releases in *The Daily* and a list of all the other releases and new publications appear on the home page. So in fact, Statistics Canada's website has a new home page every day.

We found in the market research that many of our users were either new or infrequent visitors to the site. As a result, we added the "First visit" and the "Latest indicators" sections to make the site more appealing to less frequent users and encourage repeat visitors.

Our clients found that the site was lacking in colour and graphics—not surprising, given that 40% of our visitors are students. The new toolbar, with its bolder blue, increased the site's visual appeal, and we redesigned the product banner ads to be more attractive. We also added an icon to direct clients to the 2001 Census information—only to find during usability testing that *all* of the participants used the Census button in the top menu bar when they went looking for 2001 Census information during a usability test. This is where we discovered that graphics might well enhance visual appeal, are limited as a navigational tool. The speed of displaying pages remains an important criteria so we do avoid unnecessary and time-consuming graphics.

Finally, we had a prototype home page that met with the agency's approval and conducted usability testing during which we watched the interactions of 10 site users (5 English, 5 French) as they completed a series of information retrieval exercises. We also consulted the staff in the Regional offices because they are on the front lines and often assist clients over the phone who are having difficulty finding information on the site. The feedback from both these exercises resulted in a few small changes to the prototype—and on April 19, 2001, we launched it.

Making The Hyperleap: Internet Challenges For Statistics Canada

Writing for the Web

Statistics Canada's employees have been authoring paper publications for decades, but now that the Internet is becoming the agency's primary dissemination vehicle, they are having to learn how to write effectively for users who will no longer flip pages, but scan, scroll and click. It is not an easy change to make, because it requires all our authors, agency-wide, to think about and structure information in a completely different way. It is a major undertaking, but one which the agency is supporting through a course on Writing for the Web.

The Marketing and Information Services Branch began its "Writing for the Web" workshops in 1999. The course shows Statistics Canada authors how users interact with the Web and how to write for them while maximizing the opportunities the Web offers and working within the limitations the Web imposes. So far, nine workshops have been conducted with a total of 94 participants.

The 94 authors that have taken the course so far are just the tip of the iceberg. We have a unit that provides consulting services to the other authors who still need help making their publications "web-friendly". In addition to this, we are finding that our authors are increasingly Internet users themselves and so are learning through their own experiences on the Web.

Government On-Line

Another driving force behind the evolution of the site has been the Government On-Line Initiative, which was launched in October 1999 in the Speech from the Throne. "Governments can, and should, be at the leading edge of the information revolution. A model user of the new technologies to improve services to Canadians. By 2004, our goal is to be the most electronically connected government in the world to its citizens. So that Canadians can access all government information and services on-line at the time and place of their choosing."[4]

For many government departments and agencies, the Government On-Line strategy meant a foray into the previously uncharted area of the Internet. For Statistics Canada the focus is not so much on establishing our online presence but rather tailoring this presence in such a way that Government-wide web crawlers can access the

[4] Prime Minister's Response to the Speech from the Throne, October 13, 1999

information contained on our website. Government On-Line poses challenges for content development because the site has to be conceived in a framework of governmental information architecture instead of the departmental information architecture that has been our mould until this point.

Online Data Collection

One of the biggest challenges for Statistics Canada in the next years is developing methods that allow clients to respond to our statistical surveys online. These applications must be intuitive and flexible enough to suit the "lowest-common denominator" of our respondents' computer systems, and ensure that the confidentiality of the information provided to the Agency is secured through leading-edge encryption technologies. This is an area with a steep learning curve and that will be supported by user testing.

New Technologies

Government departments and agencies will also have to be prepared to deliver information that is accessible in multiple electronic formats. Canadians are now using wireless devices and televisions to access online information. Providing content for these devices with as little re-working as possible (undoubtedly applying the "automate!" principle) will be a burgeoning area of activity in the coming years. The challenge of the Government On-Line strategy is ensuring that all development is done in consultation with clients to ensure that it responds to their needs for online service delivery.

The Hardest Lesson

We built our site on our own assumptions about what users wanted, and we modified it on the basis of their reactions—in other words, when they complained, we fixed it. We believed we would never get it right the first time, so we might as well put the draft on line. What we were actually doing was user-testing live.

Our early market research taught us a hard lesson: we cannot trust our assumptions about what users want and our predictions about how they will interact with our website.

Now we test at the earliest stages of our projects. We listen to users' recommendations, build prototypes, conduct usability testing or observational research, re-build the prototypes and consult with our users again. We sit with our staff in the regional offices to find out what kind of problems they face on the website every day as they help their clients. It all helps us get it closer to right before we go online, and helps us get better as we go along.

We have to remember that pre-launch user-testing still only gets it *almost* right. In the final analysis, our users are our greatest resource and their feedback through e-mail and formal market research is invaluable.

Pamela Best
Manager, Internet Content Development
Dissemination Division
Statistics Canada

8 'C' R.H. Coats Building
Ottawa, Ontario
K1A 0T6, Canada

Telephone: (613) 951-8201
Fax: (613) 951-4513
Email: bestade@statcan.ca