A SYSTEMS APPROACH TO STREAMLINING THE CREATION OF WEB-BASED CONTENT

Jed C. Jones
MindEcology, Inc.
4202 Fairmeadow Drive
Round Rock, TX 78665
jed@mindecology.com

ABSTRACT

In a wired world, the information one already knows is becoming less important than how adept one is at conducting effective searches for the information one desires. In the case of publicly-available, indexed information resources such as those made possible by the World Wide Web, content that cannot be found may as well not even exist in terms of its usefulness for human consumption.

The need for content to be findable on the Internet presents an important challenge for creators of content intended for consumption on the Web. Specifically, the content one creates must not only be valuable (i.e., useful and relevant within the context of a particular need) to human consumers, but it also must be properly indexed by search engine agents so that it can be made accessible to those consumers in the first place. Given the complexity of this dual requirement, content developers today lack a framework for guiding them in creating content that consistently satisfies both of these requirements.

In order to assist the creators of online content to do so in a way that is both findable and valuable to human consumers, the current paper proposes a systems approach to modelling the complex relationship between Web-based content, the immediate content needs of its intended human consumers, and the technology agents that index that content for human consumption. The intended outcome will be a Content Consumer Profile which future content creators can leverage to help them create content effectively and efficiently.

Keywords: information search, Web-based content, systems approaches, content consumption, search engines

Information Search in a Wired World

The Increasing Ubiquity of Information
As the people and organizations of the world become increasingly interconnected via the Internet, information of almost all sorts is ever-more readily accessible to the average citizen of many countries around the world. As information technologies such as processors and storage devices continue to become more affordable from year to year, this trend should continue.

For the first time in history, almost any type of information can be accessed for free by anyone who is adept at searching for it. While much of the information available on the World Wide Web remains part of the *invisible Web* (Schlein, 2006) - content that stipulates fee-based access or that

is not indexed by search engines at all - there is an unprecedented amount of information available to us that was unimaginable even 50 years ago.

The New Game: Being able to Find the Right Information Quickly
While having knowledge and information "in one's head" are still very important for many business, academic, governmental and personal contexts, the sheer amount of information in the world today is changing the rules of the information game. And, as the world speeds up, some of the most important information itself is changing more quickly than ever, making it even more important for people to constantly gain access to new information as it becomes available. Nowadays, for many people it is as or more important to be able to effectively access the information one needs than it is to be able to memorize large amounts of data. Search, rather than memorization, is the new name of the game in most areas of information-intensive human endeavour.

Advances in Search Technologies
Commensurate with this marked increase in the ubiquity of readily-available information has been rapid advances in search technologies. Since the introduction of the first search engine, Archie, by a McGill University student in 1990 (Battelle, 2005), search technologies have come a very long way in terms of their ability to connect relevant content to would-be consumers of that content. And the infrastructure that enables search has been bolstered significantly as well: it is estimated that most popular search engine, Google, employs upwards of 200,000 servers in indexing and delivering content to hungry Internet searchers.

Increased Pressure to Provide Relevant Content
The playing field for search companies has also significantly narrowed over the past decade, with the top three search engines (Google, Yahoo! and MSN) comprising about 88% of all daily searches on the Internet. Since there are significant revenue opportunities in the area of search for these companies, this narrowing of the field means that the top search firms are under continual pressure to provide increasingly relevant, valuable content to consumers. Only high levels of relevance for search results can guarantee that the top search engines will retain their positions - and their revenue streams derived from their search-related ventures.

Content: The Primal "Stuff" of the Internet

The World Wide Web and the industry of Internet search would not be possible without the existence of *content*. Content refers to any type of text or media-based item that people want to consume (viz., read, watch, play with, manipulate or listen to).

The amount of content on the Internet is unimaginably huge and growing constantly. This means that it is impossible for any individual or organization to meaningfully consume more than a tiny fraction of the content that exists on the Internet. At the same time, the large amount of content available provides an enormous array of opportunities for individuals to find content that suits their particular needs. Searchers for content are only limited by their imaginations and their competence in executing proper searches for the type of content they desire.

Content Creation and Delivery Landscape

The primary stakeholders involved in the creation, storage, and delivery of content on the Internet can be broken into three primary categories: *creators, indexers* and *consumers*. Each of these has their own unique goals, motivations and challenges.

Content creators have the goal of creating content that some target group will be able to find, consume and find valuable. Most creators of content have the key motivation of doing so to meet certain personal or organizational goals such as earning revenues, raising awareness about a certain

topic, or eliciting inbound contacts. Content creators face the challenge of competing with thousands or millions of others who are creating content that either actively targets the same audiences, shows up among the same search results as their own content, or both.

Content indexers are defined here as any company, individual or its electronic agents that scans (or crawls) as many Web sites on the Internet as possible, indexes the content that is found there, and then provides that content among search results for potential access by content consumers. The goal of content indexers is to provide content that best matches the inferred intentions of the keyword-based searches conducted by content consumers. The key motivation of indexers is usually to generate revenues through the selling of advertising or related, Web-based services such as Web hosting. The challenge for content indexers is the perennial difficulty in accurately inferring the intentions of searcher (i.e., consumers) based upon the keywords the searchers use when conducting searches. Another challenge is the indexers' ability to appropriately rank some content over other content in the search results list based upon perceived relevance.

Meanwhile, content consumers have the goal of accessing and consuming the highest-value content they can find. In this case, "value" is a function of the *relevance* of the content to the user's search keywords and the *usefulness* of the content in terms of its intended area of application or means of consumption. The key motivation of content consumers involves fulfilling a need or desire (e.g., answers a question, provides a sense of community, offers entertainment, provides revenue opportunities, etc.). The challenge that content consumers face is that of sifting through the preponderance of content available on the Internet to find the content one wants (or believes one wants).

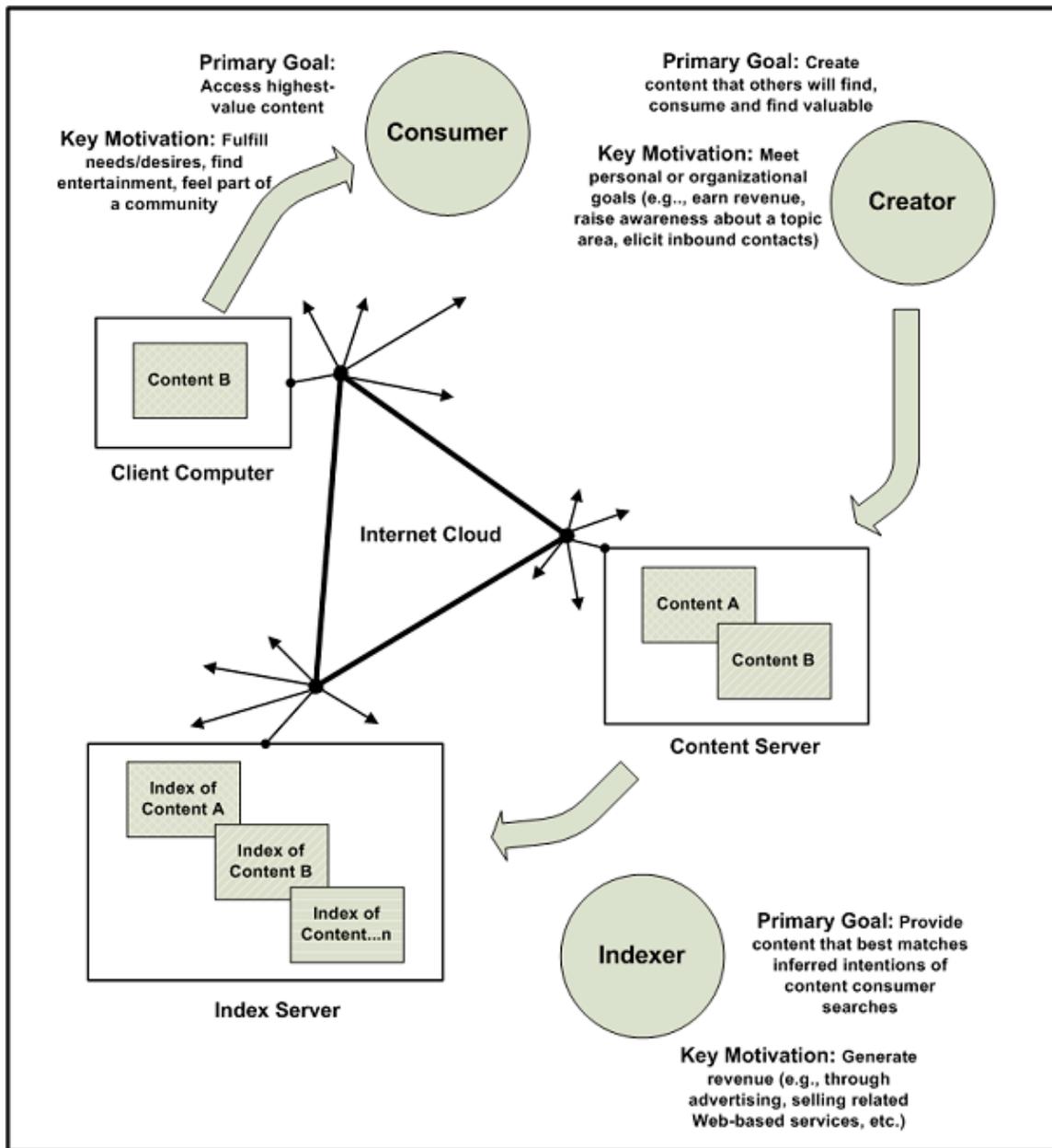# A Systems Approach to Streamlining the Creation of Web-Based Content

Primary Goal: Access highest-value content

Key Motivation: Fulfill needs/desires, find entertainment, feel part of a community

Consumer

Primary Goal: Create content that others will find, consume and find valuable

Key Motivation: Meet personal or organizational goals (e.g.., earn revenue, raise awareness about a topic area, elicit inbound contacts)

Creator

Content B

Client Computer

Internet Cloud

Content A
Content B

Content Server

Index of Content A

Index of Content B

Index of Content...n

Index Server

Indexer

Primary Goal: Provide content that best matches inferred intentions of content consumer searches

Key Motivation: Generate revenue (e.g., through advertising, selling related Web-based services, etc.)

Figure 1. Structural Model of the Internet Content Delivery Landscape: Creators, Indexers, and Consumers

A Systems Approach to Streamlining the Creation of Web-Based Content

CATWOE Analysis of Internet Content Creation and Consumption

The Internet landscape can be viewed as a system of content creation and consumption. The CATWOE approach of Checkland and Scholes (1990) offers a useful means for gaining insight into the key elements – or root definition - of any system. CATWOE stands for Customers, Actors, Transformative Process, Worldview, Owners, and Environmental Constraints. It is applied here in order to provide a root definition for Internet-as-system.

Customers

The *customers* of a given system are those who directly benefit from the *transformative process* (see below). In the case of the Internet landscape, the primary beneficiaries of the creation of content are the content consumers. In sense, the indexers and the content creators are also potential beneficiaries of the Internet-as-system. However, in the context of this paper, we will view the content consumers as the primary customers of the Internet-as-system.

Actors

The *actors* of a system are those who perform the transformative process for the benefit of the customers. For our model, the actors are both the content creators and the content indexers.

Transformative Process

The *transformative process* for a given system is that which transforms inputs to outputs. In the case of our model of Internet-as-system, the transformative process entails the creation, discovery, indexing and delivery of content to consumers.

Worldview

The *worldview* for a given system is those set of beliefs, assumptions and values that underpin the existence of the said system in the mind's of the majority of its stakeholders. The worldview that supports the Internet as an idea is that the Internet is a universally-accessible medium for the dissemination of user-relevant content to anyone, anywhere and in an unbiased manner.

Owners

The *owners* of a system are those people or entities who could put an end to the existence of the system. In the case of the Internet, many would argue that there are no true owners since it is by definition a decentralized collection of multiple independent networks. If pressed, we can assert that the owners of the Internet are the Internet Service Providers (ISPs) who provide individuals and organizations with access to the Internet, as well as the large multinationals like AT&T and WorldCom who own the large Internet backbones over which the majority of Internet traffic travels.

Environmental Constraints

The *environmental constraints* of a system relate to those items that place limitations on its operation or use. The environmental constraints of the Internet include: 1. limitations to the indexers' ability to infer consumer intentions from their search keywords, and 2. lack of a universal standard for content quality and relevance.

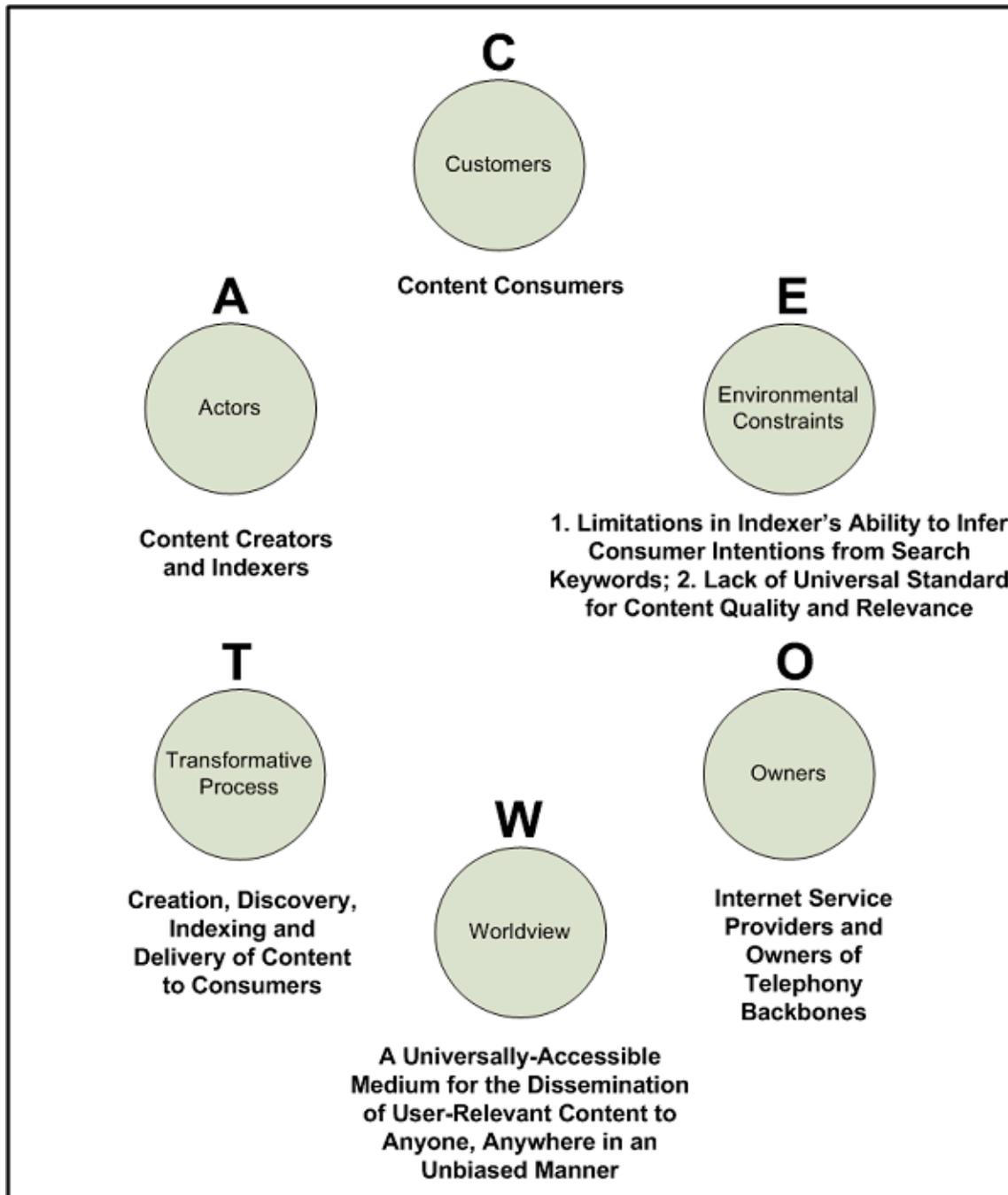# A Systems Approach to Streamlining the Creation of Web-Based Content

Figure 2. CATWOE Analysis of Internet Content Creation and Consumption

A Two-Fold Challenge for Content Creators

A number of how-to guides have been written for content consumers concerning ways to improve the efficiency and effectiveness of conducting Internet searches. In a similar vein, search engine firms continually conduct extensive research into finding ways to improve their effectiveness in delivering the most relevant content to users in the timeliest fashion. However, there seems to be missing in the literature sufficient guidelines for content creators in a way that helps them to

maximize their goals of creating content that consumers will be able to find and that they will value when they consume it. This paper makes an attempt to address that gap in the literature.

Findable Content

Content first and foremost must be able to be found by the target content consumer group for which it was intended. The ability for content to offer value to people is null if the said content cannot be found in the first place by those who might benefit from it.

Findable content is content that can be found through the reasonable efforts of people who are either actively searching for the type of information the content offers or by those who might stand to benefit from the content by merely "surfing" or browsing the Internet with only a vague sense of the purpose of their search. The key elements of findable content the following:

- Keyword-optimized

- Streamlined technical elements

- Hosted on a search engine-trusted site

- Host site has multiple backlinks

- Host site is listed in Internet directories

- Content creator has a budget for paid listings

Many creators of content fail to take the appropriate steps toward ensuring that the content they create will be found by the target consumer group.

Valuable Content

Of course, even if content one creates can be found effectively, it must also be perceived as valuable by its target consumer group if it is to be leveraged effectively.

Valuable content is content is both perceived as useful by a particular content consumer group and is at the same time relevant for use in a given context. The key elements of valuable content are as follows:

- Fulfils consumer needs/desires

- Packaged in an Internet-friendly format

- Keyword-optimized

- Appropriately-targeted

- Implies a credible and trustworthy source

A Systems Approach to Streamlining the Creation of Web-Based Content

Valuable

Fulfills Consumer Needs/Desires

Gets them closer to what they need or want
Answers a question
Provides actionable information
Yields a competitive advantage
Increases pleasure
Decreases pain
Offers a sense of community
Presents revenue-generation opportunity
Saves money
Generates professional or personal contacts
Alleviates a worry or concern
Provides hope or inspiration

Internet-Friendly Format

Easily-digestible format
Presented in a way that is easy to store
  for later consumption
Accompanied by audiovisual aids

Keyword-Optimized Content

Contains instances of search keyword
Key elements (e.g., title, anchor text) are
  topic-specific

Appropriately-Targeted Content

Tailored to psychographic profile of target
  consumers
Catchy and relevant title
Appears made for people "just like them"

Credible & Trustworthy Source

Few grammatical errors
Few syntactical errors
Few typographical errors
Well-written copy
Well-structured flow
Trusted or trustworthy-seeming source

Findable

Keyword-Optimized Content

Key elements (e.g., title, anchor text) are
  topic-specific
Content contains 3-5% of target keyword(s)
Optimized for keywords that are often-
  searched
Optimized for keywords that are under-
  represented on other sites

Streamlined Technical Elements

Code that underlies online content vehicle
  (e.g., Web site, blog) needs to be
  streamlined so that search engine crawlers
  can find and index the content

Hosted on Search Engine-Trusted Site

Content like articles and blogs that is hosted
  on a site trusted by search engines ranks
  better in natural search results
Content hosted on sites with high Google
  PageRank™ ranks better on Google

Host Site has Multiple Backlinks

Having multiple other contextually-relevant
  sites "point" links to host site contributes
  better rankings

Host Site Listed on Internet Directories
Listed on top 2-3 directories
Listed on multiple second-tier directories

Budget for Paid Listings

Content creator pays for favorable search
  engine placement
Content creator pays for ads on cost-per-
  click or cost-per-impression basis

Figure 3. Criteria for Valuable and Findable Content

Content Consumer Profile

Creating content that is both findable and valuable is something that most content creators take for granted, resulting in the content they create not being found by the intended consumer group and/or not being valued. The result is that the goals of these content creators are not being met sufficiently, with potential adverse impacts for indexers and consumers, as well.

A Systems Approach to Streamlining the Creation of Web-Based Content

The majority of the factors in creating findable and valuable content are rooted in an understanding of the "who, what, when, why and how" of the target consumer group will be using the content. This paper proposes the creation and use by content creators of a Content Consumer Profile (CCP). The CCP would be created and filled out by the content creator before the content for a particular Web site, blog, online article, news item or other item is created. The purposes of the CCP are:

- To orient the content creator to the nature of the consumer or consumers of the content

- To serve as a guideline to which the content creator can refer repeatedly during the creation of the content such that the content created is both valuable and findable by the target consumer group

The exact format and contents of the CCP employed by a given content creator would vary. However, the CCP should consist of at least the following components: demographic profile, psychographic profile, technology profile, motivations for search campaign, details of search campaign, and content search history. (Note: *search campaign* is defined here as a series of one or more searches by a single content consumer that are tied together by a certain intention or set of intentions vis-à-vis the outcome of the search).

Demographic Profile

The demographic profile for a given group may include details such as location, age range, gender, occupation, income, average education level, nationality, marital status, children/dependents, and other relevant demographics.

Psychographic Profile

The psychographic profile can include items like personality characteristics, likes, dislikes, hobbies, overarching opinions or attitudes, values, concerns, life goals, decision-making style, who or what influences their opinions, what impresses them, how conservative are they in terms of being open to new ideas, and what it takes to earn their trust.

Technology Profile

The technology profiles includes items such as how tech-savvy they are, frequency and duration of Web use, locations from which they access the Internet (e.g., work, home, etc.), most often-visited sites, general attitudes toward technology, fears about technology, propensity to belong to online groups or communities, and the degree to which they express emotions online.

Motivations for a Given Search Campaign

The profile on motivations for a given search campaign include their reasons for searching for the target content, the pain they are trying to overcome, and their expectations for the interaction.

Details of a Given Search Campaign

Details of a given search campaign include whether they were deliberately or only passively searching for this content, what problems they encountered, and what are their search or content-related annoyances and pet peeves.

Content Search History

Finally, the content search history profile includes items such as whether they have searched for this type of content in the past, the types of sites (e.g., corporate sites, blogs, forums, etc.) they normally use to search for the content, and the types of sites they avoid or find least useful.

Methodology for Filling out the CCP

One or more CCPs will be created for each subject area with which the content creator is involved. At the outset of a given project, it may not be clear exactly how many CCPs will be required for a given content area; this will become clearer as the initial CCP is created. Creating the CCP is thus an iterative process, requiring that the content creator go back and forth between information that he or she finds about the content consumers and the evolving CCP itself. During this process, it may become apparent that more than one CCP is needed in order to represent two or more distinct content consumer groups, each with different reasons or motives for potentially consuming the content that will be created.

In gathering data for the CCP, the content creator may choose to visit multiple online resources such as forums, discussion groups, blogs, and group pages. Content that describes people's thoughts and experiences in the first-person can yield very important insights into that particular consumer group. At the same time, objective, second-person or third-person accounts or studies about the group may also yield very valuable information that can be used to complete the CCP.

The content creator filling out the CCP may also choose to glean relevant information from resources such as interviews, surveys, focus groups, periodicals, books, and direct observation - or even from one's own knowledge and experience.

It is important to note that not every section of the CCP need be filled out. This will likely be true for sections whereby the item is either not relevant for gaining insights into this particular content consumer group or whereby the information is not readily available via the information resources at the disposal of the content creator.

A continual challenge will be how to know when a particular CCP is complete and accurate. For example, it will be difficult to know whether the information one has gathered about a particular content consumer group is merely anecdotal or whether it actually represents tangible evidence about the nature of the consumer group in question. However, it is expected that experienced researchers will come to know when a particular profile is complete as they reach a point of information saturation. This point has likely been reached if the content creator notices an increased incidence of repeat or "known" information about a given group after extensive searching has been completed, while at the same time there is very little new information being revealed about the target group.

Conclusion

The increased availability of information that the Internet and its supporting technologies afford people today have shifted the importance from the ability to memorize information to the ability to access relevant information, or content. The primary stakeholder groups in the creation and consumption of content on are content creators, indexers, and consumers. Each of these stakeholder groups has a unique set of goals, motivations, and challenges pertaining to their role in the creation, discovery, indexing, and delivery of content to content consumers on the Internet. The Internet content and creation landscape can be seen as a system, with well-defined customers, actors, transformative processes, a supporting worldview, owners and environmental constraints that constitute the root definition of the Internet-as-system.

A Systems Approach to Streamlining the Creation of Web-Based Content

There is ample literature on how content consumers can conduct more effective searches, and content indexers (i.e., search engine firms) are under continual pressure to make search more efficient and effective such that the right content is delivered to the right consumers. However, there is a gap in the literature that informs content creators on how to create content that is both findable and valuable to a particular content consumer group.

The present paper proposes that content creators would be able to create content that is more findable and valuable if they are guided and informed by a clear understanding of the target content consumer group for which the content is ultimately being created. To that end, the paper proposes that content creators create and fill out a Content Consumer Profile (CCP) that details the primary characteristics and search-related motivations of a given group vis-à-vis the content in question. This will allow them to create much more findable, valuable content the further the goals of all key stakeholders in the Internet-as-system.

REFERENCES

Battelle, J. (2005). *Search*., Penguin Books, Ltd., London.

Checkland, P and Scholes, J. (1990). *Soft Systems Methodology in Action*., John Wiley & Sons, Ltd., West Sussex, England.

Schlein, A.M. (2006). *Find it Online: Fully Revised 4th Edition.,* Facts on Demand Press, Tempe, AZ.