Exhibit B-12

ACC - 12

Invalidity Chart Braden in view of Kurtzman, II and Additional Prior Art References

Invalidity Chart Braden in view of Kurtzman, II and Additional Prior Art References

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
45. A data processing method		Kurtzman, II 1:54-56 "[A]n	Salton 1968, p. 414
for generating a user data		object of the invention is to	Incoming items and documents to be stored Technical personnel and system users
profile representative of a		provide a more sophisticated	
user's social, cultural,		profiling technique for	Micrafilming and Indexing and abstract- Preparation of interest profiles for users
educational, economic		generating a more useful user	
background and of the user's psychological profile, the		profile."	Document profiles User profiles
method being implemented in a		Kurtzman, II 3:21-23 "User	Microfilm readers viewing Automatic search and
computer system having a		profiling uses content stream	retrieval system
storage system, comprising the		analysis, as well as	Document depot
steps of:		demographic, geographic,	Copies Selective Abstract Search requests
		psychographic, digital	dissemina- tion secondary retrieval journals files
		identification, and HTTP	
		information."	Fig. 10-4 Typical technical information center.
			Salton 1968, p. 93 "There are many ways in
			which higher level terms, corresponding in the
			natural language to phrases or to word
			combinations, might be assigned to documents
			as content identifiers. These include, for
			example, statistical procedures measuring the
			strength of association between text words, and
			syntactic analysis methods that detect syntactic relationships between words. A third
			possibility, called the statistical phrase process,
			incorporated into the Smart system is based on
			a pre constructed phrase dictionary, and
			phrases are detected by a look up procedure
			similar to that previously described for the
			regular word stem thesaurus."
			Chislenko 4:15-18 "For example, the system
			may assume that Web sites for which the user
			has created "bookmarks" are liked by that user

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
			and may use those sites as initial entries in the user's profile."
			Belkin p. 399 "In the general information seeking interaction, the IR system needs to have (see Table 1 for a brief listing of the ten functions and their acronyms): a model of the user himself, including goals, intentions and experience (UM)."
			Belkin p. 402 "I ³ R (Intelligent Interface for Information Retrieval) is a system designed to help overcome the difficulties of using text retrieval systems. As an interface system, it is responsive to a wide variety of users, who have varying levels of ability in computer use and varying levels of knowledge about the topic being investigated."
			Herz 27:62-66 "In a variation, each user's user profile is subdivided into a set of long-term attributes, such as demographic characteristics, and a set of shortterm attributes that help to identify the user's temporary desires and emotional state."
			Herz 20:35-37; 11:31-38 "User profiles may make use of any attributes that are useful in characterizing humans written response[s] to Rorschach inkblot test multiple-choice responses by [the person] to self-image questions their literary tastes and psychological peculiarities."
			Herz <i>See also</i> Abstract; 1:18-43; 4:49–8:8; 28:41–55:42; 55:44–56:14; 56:15-30; 58:57–60:9; Figures 1-16.

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
(a) retrieving, by the computer		Kurtzman, II 3:45-49 "The user	See Salton 1968, p. 414 (Fig. 10-4)
system, user linguistic data		instructs the website server to	
previously provided by the		access the website corpus and	Salton 1968, p. 95 "If it is also desired to use
user, said user linguistic data		retrieve and transmit specific	the syntactic option, those sentences containing
comprising at least one text		website files. These specific	statistical phrases are separated from the
item, each said at least one text		files selected and viewed by the	remainder of the text in order to be used later
item comprising at least one		user are recorded by the	as input for the syntactic analysis programs.
sentence;		affinity server."	These programs, to be described in Chap. 5,
			are designed to eliminate statistical phrases that
			do not pass the syntactic screens; they need not
			be applied to sentences in which no statistical
			phrases were ever detected."
			Salton 1968, p. 96, Fig. 3-19.

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
			Read in one segment of the phrase dictionary Set up an array containing up to six component concepts for a given phrase Set ** = 1 Does the **I component in the array match any of the concepts attached to the present text word? No Set ** equal to **+1 No Is the input array empty? Is the next text word in the same sentence? Yes No Calculate frequency of phrase in present sentence and add if to frequency count of this phrase on output list Have we reached the end of the phrase found in this document? Yes Move phrase found in this document Yes Move phrase to output list Are there more phrases in core? Yes Are there more phrases on tape? Yes Are there more phrases on tape? Yes Statistical phrase detection.
			Chislenko 4:15-18 "For example, the system may assume that Web sites for which the user has created "bookmarks" are liked by that user and may use those sites as initial entries in the user's profile."
			Chislenko 4:40-50 "Ratings can be inferred by the system from the user's usage pattern. For example, the system may monitor how long the user views a particular Web page and store in that user's profile an indication that the user

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
			likes the page, assuming that the longer the
			user views the page, the more the user likes the
			page. Alternatively, a system may monitor the
			user's actions to determine a rating of a
			particular item for the user. For example, the system may infer that a user likes an item
			which the user mails to many people and enter
			in the user's profile and indication that the user
			likes that item."
			Dedrick See, e.g., 3:37–4:9, 5:34–6:3, 6:53–
			8:19, 14:65–15:10, Abstract, Figures 1–8.
			Dollsin m. 402 "The executed attended to the
			Belkin p. 402 "The overall structure of the system is based on blackboard architecture, a
			collection of independent cooperating experts
			which communicate indirectly using a shared
			global data structure. The I ³ R system can be
			compared to the Hearsay-II system. Hearsay-II
			is a speech understanding system that
			synthesizes the partial interpretations of several
			diverse knowledge sources into a coherent understanding of a spoken sentence. [21]
			Knowledge sources communicate by reading
			and writing on a blackboard. The blackboard
			has several distinct levels which hold different
			representations of the problem space."
			[Similar to the speech understanding system,
			I ³ R takes data provided by the user, which can
			include sentences, and later uses the
			information.]
			Belkin p. 403 "For I ³ R to be adaptable, it must
			be able to assess the user's abilities so it can
			adjust the interface to match them.[22]."
			Salton 1989, p. 405-6 "To help furnish

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
			semantic interpretations outside specialized or
			restricted environments, the existence of a
			knowledge base is often postulated. Such a
			knowledge base classifies the principal entities
			or concepts of interest and specifies certain
			relationships between the entities. [43-45]
			The literature includes a wide variety of
			different knowledge representations [one of
			the] best-known knowledge-representation
			techniques [is] the <i>semantic-net</i> In
			generating a semantic network, it is necessary
			to decide on a method of representation for
			each entity, and to relate or characterize the
			entities. The following types of knowledge
			representations are recognized: [46-48] A
			linguistic level in which the elements are
			language specific and the links represent
			arbitrary relationships between concepts that
			exist in the area under consideration."
			Herz 27:62-67 "In a variation, each user's user
			profile is subdivided into a set of long-term
			attributes, such as demographic characteristics,
			and a set of short-term attributes such as
			the user's textual and multiple-choice answers
			to questions."
			-
			Herz 56:20-28 "As in any application
			involving search profiles, they can be initially
			determined for a new user (or explicitly altered
			by an existing user) by any of a number of
			procedures, including the following preferred
			methods: (2) using copies of the profiles of
			target objects or target clusters that the user
			indicates are representative of his or her
			interest."

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
			Herz See also Abstract; 1:18-43; 4:49–8:8; 28:41–55:42; 55:44–56:14; 56:15-30; 58:57–60:9; Figures 1-16. Culliss 3:46-48 "Inferring Personal Data. Users can explicitly specify their own personal data, or it can be inferred from a history of their search requests or article viewing habits. In this respect, certain key words or terms, such as those relating to sports (i.e. "football" and "soccer"), can be detected within search requests and used to classify the user as someone interested in sports."
(b) generating, by the computer system, an empty user data profile;		Kurtzman, II Abstract "Content stream analysis is a user profiling technique that generates a user profile based on the content files selected and viewed by a user. This user profile can then used to help select an advertisement or other media presentation to be shown to the user." Kurtzman, II See, e.g., 3:45-50.	See Salton 1968 p. 414 (Fig. 10-4). Chislenko 3:38-39 "Each user profile associates items with the ratings given to those items by the user. Each user profile may also store information in addition to the user's ratings." Chislenko 4:15-18 "For example, the system may assume that Web sites for which the user has created "bookmarks" are liked by that user and may use those sites as initial entries in the user's profile." Chislenko 4:40-50 "Ratings can be inferred by the system from the user's usage pattern. For example, the system may monitor how long the user views a particular Web page and store in that user's profile an indication that the user likes the page, assuming that the longer the user views the page, the more the user likes the page. Alternatively, a system may monitor the

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
			user's actions to determine a rating of a
			particular item for the user. For example, the
			system may infer that a user likes an item which the user mails to many people and enter
			in the user's profile and indication that the user
			likes that item."
			Dedrick See, e.g., 3:37–4:9, 5:34–6:3, 6:53–
			8:19, 14:65–15:10, Abstract, Figures 1–8.
			Belkin p. 403 "For I ³ R to be adaptable. it must
			be able to assess the user's abilities so it can
			adjust the interface to match them.[22] This
			requires a user model builder. As each user
			may have his own view of the subject area
			being searched, it would be valuable to capture this information and remember it from session
			to session in a domain knowledge expert."
			to session in a domain into wiedge emperu
			Herz 56:20-31 teaches that user profiles should
			be created for "new users," 27:49-51, and
			specifies how user search profiles can be "initially determined."
			initially determined.
			Herz See also Abstract; 1:18-43; 27:47-49;
			27:62-67; 4:49–8:8; 28:41–55:42; 55:44–
			56:14; 56:15-30; 58:57–60:9; Figures 1-16.
			Culliss 3:46-48 "Inferring Personal Data.
			Users can explicitly specify their own personal
			data, or it can be inferred from a history of
			their search requests or article viewing habits.
			In this respect, certain key words or terms,
			such as those relating to sports (i.e. "football" and "soccer"), can be detected within search
			requests and used to classify the user as
			someone interested in sports."

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
(c) retrieving, by the computer	Braden 7:47-49 "each of the	Kurtzman, II 3:49-50 "The	Salton teaches retrieving locating multiple text
system, a text item from said	documents in the set is	content stream to be analyzed	items. See Salton 1968, p. 96 (Fig. 3-19) ("Are
user linguistic data;	subjected to natural language	includes the specific files	there more phrases on [magnetic storage]
	processing, specifically	selected and viewed by the	tape"), above.
	morphological, syntactic and	user."	C1:1 1 415 10 4F
	logical form, to produce logical	и с 7 10	Chislenko 4:15-18 "For example, the system
	forms for each sentence in that document."	Kurtzman, II, Figs. 6, 7, and 9.	may assume that Web sites for which the user
	document.		has created "bookmarks" are liked by that user and may use those sites as initial entries in the
	Braden Abstract "Apparatus		user's profile."
	and accompanying methods for		user's proffic.
	an information retrieval system		Chislenko 4:40-50 "Ratings can be inferred by
	that utilizes natural language		the system from the user's usage pattern. For
	processing to process results		example, the system may monitor how long the
	retrieved by, for example, an		user views a particular Web page and store in
	information retrieval engine		that user's profile an indication that the user
	such as a conventional		likes the page, assuming that the longer the
	statistical-based search engine,		user views the page, the more the user likes the
	in order to improve overall		page. Alternatively, a system may monitor the
	precision. Specifically, such a search ultimately yields a set of		user's actions to determine a rating of a particular item for the user. For example, the
	retrieved documents. Each		system may infer that a user likes an item
	such document is then		which the user mails to many people and enter
	subjected to natural language		in the user's profile and indication that the user
	processing to produce a set of		likes that item."
	logical forms."		
			Dedrick See, e.g., 3:37–4:9, 5:34–6:3, 6:53–
	Braden See, e.g., 11:62-14:61.		8:19, 14:65–15:10, Abstract, Figures 1–8.
			2
			Belkin p. 402 "I ³ R (Intelligent Interface for
			Information Retrieval) is a system designed to
			help overcome the difficulties of using text
			retrieval systems. As an interface system, it is
			responsive to a wide variety of users, who have varying levels of ability in computer use and
			varying levels of ability in computer use and varying levels of knowledge about the topic
			being investigated. The I ³ R system can be
			being investigated. The I K system can be

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
			compared to the Hearsay-II system. Hearsay-II
			is a speech understanding system that
			synthesizes the partial interpretations of several
			diverse knowledge sources into a coherent
			understanding of a spoken sentence."
			Salton 1989, p. 290 "[S]tored records are
			identified by sets of single terms that are used
			collectively to represent the information
			content of each record Among the
			methods suggested to generate complex
			identifiers are linguistic-analysis procedures
			capable of recognizing linguistically related
			units in document texts."
			Salton 1989, p. 294-6 (see also fn. 28-30)(
			Linguistic methodologies including syntactic
			class indicators (adjective, noun, adverb, etc.)
			are assigned to the terms).
			Herz 13:24-27 teaches that, for the purposes of
			creating a profile, "one could break the text
			into overlapping word bigrams (sequences of 2
			adjacent words), or more generally, word n-
			grams."
			Herz See also Abstract; 1:18-43; 27:47-49;
			27:62-67; 4:49-8:8; 28:41-55:42; 55:44-
			56:14; 56:15-30; 58:57–60:9; Figures 1-16.
			Culliss 3:46-48 "Inferring Personal Data
			Users can explicitly specify their own personal
			data, or it can be inferred from a history of
			their search requests or article viewing habits.
			In this respect, certain key words or terms,
			such as those relating to sports (i.e. "football"
			and "soccer"), can be detected within search
		11	

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
			requests and used to classify the user as someone interested in sports."
			•
(d) separating, by the computer system, said text item into at	Braden 7:47-49 "each of the documents in the set is	Kurtzman, II 4:38-39 "Next, the individual words are passed	See Salton 1968, p. 96 (Fig. 3-19), above.
least one sentence;	subjected to natural language processing, specifically	through a stemming procedure to obtain words and word-	Salton 1968, p. 95 "The phrase finding process is completely straightforward and consists of
	morphological, syntactic and logical form, to produce logical	stems (block 708)."	matching the first component of a given phrase with each component of each word of a given
	forms for each sentence in that document."	Kurtzman, II, Figs. 6, 7, and 9.	sentence; the second phrase component is then matched, and so on."
		Kurtzman, II See, e.g., 5:31-41.	,
	Braden Abstract "Each such document is then subjected to		Dedrick See, e.g., 3:37–4:9, 5:34–6:3, 6:53–8:19, 14:65–15:10, Abstract, Figures 1–8.
	natural language processing to produce a set of logical forms.		Kupiec 4:27-29 "Continuing with Example 1,
	Each such logical form encodes, in a word-relation-		suppose that the retrieved documents contain the following additional noun phrases in
	word manner, semantic		proximity to the noun phrase "New York
	relationships, particularly argument and adjunct structure,		City.""
	between words in a phrase."		Kupiec 11:19-24 "In step 300 the input string
	Braden See, e.g., 11:62-14:61.		is analyzed to determine what part of speech each word of the string is. Each word of the
	Braden 800, 0181, 11102 1 11011		string is tagged to indicate whether it is a noun,
			verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that
			uses a hidden Markov model. The result
			produced by step 300 is a tagged input string."
			Kupiec 11:28-30 "In step 310, which
			comprises component steps 311 and 312, the tagged input string is analyzed to detect noun
			phrases. In step 315 the tagged input string is further analyzed to detect main verbs."
			-

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
The '067 Patent	Braden-Harder	Kurtzman, II	Kupiec 13:15-21 "The match sentences are analyzed in substantially the same manner as the input string is analyzed in step 220 above. The detected phrases typically comprise noun phrases and can further comprise title phrases or other kinds of phrases. The phrases detected in the match sentences are called preliminary hypotheses." Kupiec 14:45-54 "Hypotheses are verified in step 260 through lexico-syntactic analysis. Lexico-syntactic analysis comprises analysis of linguistic relations implied by lexico-syntactic patterns in the input string, constructing or generating match templates based on these relations, instantiating the templates using particular hypotheses, and then attempting to match the instantiated templates, that is, to find primary or secondary documents that contain text in which a hypothesis occurs in the context of a template." Herz 13:24-27 teaches that, for the purposes of creating a profile, "one could break the text
			creating a profile, "one could break the text into overlapping word bigrams (sequences of 2 adjacent words), or more generally, word ngrams."
			Herz See also Abstract; 1:18-43; 27:47-49; 27:62-67; 4:49–8:8; 28:41–55:42; 55:44–56:14; 56:15-30; 58:57–60:9; Figures 1-16.

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
(e) extracting, from each of	Braden 7:47-49 "each of the	Kurtzman, II 4:38-39 "Next,	See Salton 1968 p. 96 (Fig. 3-19), above.
said at least one sentence, by	documents in the set is	the individual words are passed	
the computer system, at least	subjected to natural language	through a stemming procedure	Salton 1968, p. 95 "The phrase finding process
one segment representative of a	processing, specifically	to obtain words and word-	is completely straightforward and consists of
linguistic pattern of each	morphological, syntactic and	stems (block 708)."	matching the first component of a given phrase
sentence of said at least one	logical form, to produce logical		with each component of each word of a given
sentence;	forms for each sentence in that	Kurtzman, II 5:31-41 "Each	sentence; the second phrase component is then
	document."	content file in the content stream is converted into	matched, and so on."
	Braden Abstract "Each such	individual words. Insignificant	Salton 1968, p. 95 "If a particular phrase is
	document is then subjected to	words such as HTML	found in a sentence, an appropriate entry is
	natural language processing to	formatting tags and stop words	made in a chained list of concept numbers,
	produce a set of logical forms.	are discarded. The individual	similar in format to the list of concepts derived
	Each such logical form	words are then passed through	by the thesaurus look-up. This concept list is
	encodes, in a word-relation-	a stemming procedure to obtain	kept sorted by concept number, and each
	word manner, semantic	words and word-stems. The	concept is stored together with its weight and
	relationships, particularly	word and word-stems are	with coded information identifying the given
	argument and adjunct structure,	counted to determine their	concept number as a phrase concept. A typical
	between words in a phrase."	frequencies. These frequencies are paired with the words and	entry in the chained list is shown in Fig. 3-18."
	Braden See, e.g., 11:62-14:61.	word-stems to create a	Salton 1968, p. 95 "If it is also desired to use
		multidimensional vector for	the syntactic option, those sentences containing
		each content file in the content	statistical phrases are separated from the
		stream."	remainder of the text in order to be used later
			as input for the syntactic analysis programs."
		Kurtzman, II, Figs. 6, 7, and 9.	G 1. 4050 470 // 1
			Salton 1968, p. 158 "Automatic phrase structure recognition. A number of operating
			automatic recognition procedures are based on
			context-free phase structure grammars [8].
			This is the case notably of all so-called
			"syntax-directed" compiling systems used in
			the computer field for the recognition and
			translation of computer programming
			languages. One of the best known systems for
			automatic analysis of the context-free
			languages is the predictive analyzer [9, 10].

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
			This system produces for a given sentence all possible syntactic interpretation compatible with the context-free grammar being used For example, the word <i>base</i> will have a homograph assignment corresponding to <i>noun</i> , <i>singular</i> , one corresponding to <i>transitive verb</i> , and one corresponding to <i>adjective</i> ."
			Salton 1968, p. 166 "It appears possible, therefore, as a first step toward a more complete linguistic analysis to attempt to combine a variety of grammatically related phrase components into larger entities, termed criterion phrases or criterion trees and to assign these phrases as document identifiers in the same way other concepts extracted from the thesaurus or from the statistical phrase dictionary [20]."
			Dedrick See, e.g., 3:37–4:9, 5:34–6:3, 6:53–8:19, 14:65–15:10, Abstract, Figures 1–8.
			Kupiec 4:27-29 "Continuing with Example 1, suppose that the retrieved documents contain the following additional noun phrases in proximity to the noun phrase "New York City.""
			Kupiec 11:19-24 "In step 300 the input string is analyzed to determine what part of speech each word of the string is. Each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
			Kupiec 11:28-30 "In step 310, which
			comprises component steps 311 and 312, the
			tagged input string is analyzed to detect noun
			phrases. In step 315 the tagged input string is
			further analyzed to detect main verbs."
			Kupiec 13:15-21 "The match sentences are
			analyzed in substantially the same manner as
			the input string is analyzed in step 220 above.
			The detected phrases typically comprise noun
			phrases and can further comprise title phrases
			or other kinds of phrases. The phrases detected
			in the match sentences are called preliminary
			hypotheses."
			Kupiec 14:45-54 "Hypotheses are verified in
			step 260 through lexico-syntactic analysis.
			Lexico-syntactic analysis comprises analysis of
			linguistic relations implied by lexico-syntactic
			patterns in the input string, constructing or
			generating match templates based on these
			relations, instantiating the templates using
			particular hypotheses, and then attempting to match the instantiated templates, that is, to find
			primary or secondary documents that contain
			text in which a hypothesis occurs in the context
			of a template."
			Belkin p. 402 "Knowledge sources
			communicate by reading and writing on a
			blackboard. The blackboard has several distinct levels which hold different representations of
			the problem space. Typical blackboard levels
			for speech understanding are sound segments,
			syllables, words, and phrases. The knowledge
			sources are pattern-action productions: if the
			information on the blackboard matches the

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
			pattern of a knowledge source then its action can be executed. At any time, many knowledge sources are likely to have patterns that match the contents of the blackboard." Herz 13:24-27 teaches that, for the purposes of creating a profile, "one could break the text into overlapping word bigrams (sequences of 2 adjacent words), or more generally, word n-grams." Herz <i>See also</i> Abstract; 1:18-43; 27:47-49; 27:62-67; 4:49–8:8; 28:41–55:42; 55:44–
			56:14; 56:15-30; 58:57–60:9; Figures 1-16.
(f) adding, by the computer system, at least one segment extracted at said step (e) to said user data profile;	Braden See, e.g., 11:62-14:61.	Kurtzman, II 6:31-33 "[C]reating a content data structure which indicates features of the content having particular characteristics converting the content data into individual words."	See Salton 1968, p. 96 (Fig. 3-19), above. Salton 1968, p. 95 "If a particular phrase is found in a sentence, an appropriate entry is made in a chained list of concept numbers, similar in format to the list of concepts derived by the thesaurus look-up. This concept list is kept sorted by concept number, and each concept is stored together with its weight and with coded information identifying the given concept number as a phrase concept. A typical entry in the chained list is shown in Fig. 3-18." Dedrick <i>See</i> , e.g., 3:37–4:9, 5:34–6:3, 6:53–
			8:19, 14:65–15:10, Abstract, Figures 1–8. Belkin p. 403 "For I ³ R to be adaptable, it must be able to assess the user's abilities so it can adjust the interface to match them.[22] This requires a user model builder. As each user may have his own view of the subject area

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
			being searched, it would be valuable to capture this information and remember it from session to session in a domain knowledge expert."
			Herz 13:24-27 teaches that, for the purposes of creating a profile, "one could break the text into overlapping word bigrams (sequences of 2 adjacent words), or more generally, word n-grams."
			Herz <i>See also</i> Abstract; 1:18-43; 27:47-49; 27:62-67; 4:49–8:8; 28:41–55:42; 55:44–56:14; 56:15-30; 58:57–60:9; Figures 1-16.
			Culliss 3:46-48 "Inferring Personal Data. Users can explicitly specify their own personal data, or it can be inferred from a history of their search requests or article viewing habits. In this respect, certain key words or terms, such as those relating to sports (i.e. "football" and "soccer"), can be detected within search requests and used to classify the user as someone interested in sports."
(g) repeating, by the computer system, said steps (c) to (f) for each text item of said at least one text item in said user linguistic data;	Braden 7:47-49 "each of the documents in the set is subjected to natural language processing, specifically morphological, syntactic and logical form, to produce logical forms for each sentence in that	Kurtzman, II 3:49-50 "The content stream to be analyzed includes the specific files selected and viewed by the user." Kurtzman, II 5:31-41 "FIG. 9	See Salton 1968, p. 96 (Fig. 3-19), above. Salton 1968, p. 95 "The phrase finding process is completely straightforward and consists of matching the first component of a given phrase with each component of each word of a given sentence; the second phrase component is then
	document." Braden See, e.g., 11:62-14:61.	shows the creation of content feature vectors from the content files in the content stream (block 620). Each content file in the content	matched, and so on." Salton 1968, p. 95 "After all phrases detected in a given document are entered into the chained list, this list is merged with the

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
		stream is converted into	concepts derived from other sources (for
		individual words (block 910).	example, from the regular thesaurus), as
		Insignificant words such as	previously seen in Fig. 3-16."
		HTML formatting tags (block	
		920) and stop words (block	Chislenko 4:40-50 "Ratings can be inferred by
		930) are discarded. The	the system from the user's usage pattern. For
		individual words are then	example, the system may monitor how long the
		passed through a stemming	user views a particular Web page and store in
		procedure to obtain words and	that user's profile an indication that the user
		word-stems (block 940). The	likes the page, assuming that the longer the
		word and word-stems are	user views the page, the more the user likes the
		counted to determine their	page. Alternatively, a system may monitor the
		frequencies (block 950). These	user's actions to determine a rating of a
		frequencies are paired with the	particular item for the user. For example, the
		words and word-stems to create	system may infer that a user likes an item
		a multidimensional vector for	which the user mails to many people and enter
		each content file in the content	in the user's profile and indication that the user
		stream (block 960)."	likes that item."
		Kurtzman, II, Figs. 6, 7, and 9.	Dedrick See, e.g., 3:37–4:9, 5:34–6:3, 6:53–
		, , , , , , , , , , , , , , , , , , ,	8:19, 14:65–15:10, Abstract, Figures 1–8.
			, , ,
			Salton 1989, p. 388-89 "This reduces the
			analysis to a pattern matching system in which
			the presence of particular patterns in the input
			leads to corresponding output responses
			As mentioned in Chapter 9, pattern-matching
			techniques have been widely used in automatic
			indexing for the assignment of complex
			content identifiers consisting of multiword
			phrases. [23-25] In that case, syntactic class
			markers, such as nominal, adjective, and
			pronoun, are first attached to the text words.
			Syntactic class patterns are then specified, such
			as "noun-noun," or "adjective-adjective-noun,"
			and groups of text words corresponding to
		10	permissible syntactic class patterns are
		19	

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
			assigned to the texts for content identification."
			Herz 12:55-64 teaches that textual documents
			may be profiled using word frequencies. "[A]
			textual attribute, such as the full text of a
			movie review, can be replaced by a collection
			of numeric attributes that represent scores to
			denote the presence and significance of the words in that text. The score of a word in a
			text may be defined in numerous ways. The
			simplest definition is that the score is the rate
			of the word in the text, which is computed by
			computing the number of times the word
			occurs in the text, and dividing this number by
			the total number of words in the text."
			Herz 13:24-27 teaches that, for the purposes of
			creating a profile, "one could break the text
			into overlapping word bigrams (sequences of 2
			adjacent words), or more generally, word n-grams."
			grams.
			Herz See also Abstract; 1:18-43; 27:47-49;
			27:62-67; 4:49–8:8; 28:41–55:42; 55:44–
			56:14; 56:15-30; 58:57–60:9; Figures 1-16.
			Culliss 3:46-48 "Inferring Personal Data
			Users can explicitly specify their own personal
			data, or it can be inferred from a history of
			their search requests or article viewing habits.
			In this respect, certain key words or terms, such as those relating to sports (i.e. "football"
			and "soccer"), can be detected within search
			requests and used to classify the user as
			someone interested in sports."
			1

[b] generating at least one user segment group, by the computer system, by grouping together identical segments of said at least one segment; Braden Abstract "Each such document." Braden Abstract "Each such document is then subjected to natural language processing of each sentence in that document." Braden Abstract "Each such document is then subjected to natural language processing or each set of logical forms. Each such logical forms for each set minimum number of occurrences of each of the phrase in a given sentence determines the return. Such content file in the content site with the words and word-stems to read the content file in the content site with the words and word-stems to read to frems fire and to file in the content file in the content file in	The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
computer system, by grouping together identical segments of said at least one segment; said at least one segment in the following additional not phrase and is initially defined as the minimum number of occurrences of a phrase in a given sentence determines the weight assigned to that phrase and is initially defined as the minimum number of occurrences of a phrase in a given sentence determines the weight assigned to that phrase and is initially defined as the minimum number of occurrences of sach of the phrase components in the sentence. If a phrase and is initially defined as the minimum number of occurrences of each of the phrase components are sentence in the chained list is detected, appropriately increased. Salton 1968, p. 9. The manufer of o	(h) generating at least one user	Braden 7:47-49 "each of the	Kurtzman, II 5:38-41 "These	See Salton 1968, p. 96 (Fig. 3-19), above.
together identical segments of said at least one segment; by processing, specifically morphological, syntactic and logical form, to produce logical forms for each sentence in that document." Braden Abstract "Each such document is then subjected to natural language processing to produce a set of logical forms. Each such logical form encodes, in a word-relation-word manner, semantic relationships, particularly argument and adjunct structure, between words in a phrase." Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61.	segment group, by the	documents in the set is	1 1	
said at least one segment; morphological, syntactic and logical forms for each sentence in that document." Braden Abstract "Each such document is then subjected to natural language processing to produce a set of logical formencodes, in a word-relation-word manner, semantic relationships, particularly argument and adjunct structure, between words in a phrase." Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61. Braden Ward and adjunct structure, between words in a phrase." Braden See, e.g., 11:62-14:61.		subjected to natural language	words and word-stems to create	<u>-</u>
logical form, to produce logical forms for each sentence in that document." Braden Abstract "Each such document is then subjected to natural language processing to produce a set of logical form encodes, in a word-relation-word manner, semantic relationships, particularly argument and adjunct structure, between words in a phrase." Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61.	S S	1		
forms for each sentence in that document." Kurtzman, II, Figs. 6, 7, and 9. Kurtzman, II, Figs. 6, 7, and 9. Faden Abstract "Each such document is then subjected to natural language processing to produce a set of logical forms. Each such logical forms encodes, in a word-relation-word manner, semantic relationships, particularly argument and adjunct structure, between words in a phrase." Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61.	said at least one segment;			
document." Kurtzman, II, Figs. 6, 7, and 9. Braden Abstract "Each such document is then subjected to natural language processing to produce a set of logical forms. Each such logical form encodes, in a word-relation-word manner, semantic relationships, particularly argument and adjunct structure, between words in a phrase." Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61. Kurtzman, II, Figs. 6, 7, and 9. in the sentence. If a phrase already entered in the chained list is detected, appropriately increased. Since a given text word may correspond to many concept numbers, it is theoretically possible that a single word may be responsible for the generation of a complete phrase; such a condition is not allowed, and care is taken to eliminate "phrases" where the several components are detected in the same word." Dedrick See, e.g., 3:37–4:9, 5:34–6:3, 6:53–8:19, 14:65–15:10, Abstract, Figures 1–8. Kupice 4:27-29 "Continuing with Example 1, suppose that the retrieved documents contain the following additional noun phrases in proximity to the noun phrase "New York City."" Kupice 11:19-24 "In step 300 the input string is analyzed to determine what part of speech each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."			stream."	
the chained list is detected, appropriately increased. Since a given text word may correspond to many concept numbers, it is theoretically possible that a single word may be responsible for the generation of a complete phrase; such a condition is not allowed, and care is taken to eliminate "phrases" where the several components are detected in the same word." Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61. the chained list is detected, appropriately increased. Since a given text word may to correspond to many concept numbers, it is theoretically possible that a single word may be responsible for the generation of a complete phrase; such a condition is not allowed, and care is taken to eliminate "phrases" where the several components are detected in the same word." Dedrick See, e.g., 3:37–4:9, 5:34–6:3, 6:53–8:19, 14:65–15:10, Abstract, Figures 1–8. Kupiec 4:27-29 "Continuing with Example 1, suppose that the retrieved documents contain the following additional noun phrases in proximity to the noun phrase "New York City."" Kupiec 11:19-24 "In step 300 the input string is analyzed to determine what part of speech each word of the string is. Each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."			и с 7 10	
Braden Abstract "Each such document is then subjected to natural language processing to produce a set of logical forms. Each such logical form encodes, in a word-relation-word manner, semantic relationships, particularly argument and adjunct structure, between words in a phrase." Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61.		document."	Kurtzman, II, Figs. 6, /, and 9.	
document is then subjected to natural language processing to produce a set of logical forms. Each such logical form encodes, in a word-relation-word manner, semantic relationships, particularly argument and adjunct structure, between words in a phrase." Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61.		Draden Abetreet "Feeb such		
natural language processing to produce a set of logical forms. Each such logical form encodes, in a word-relation-word manner, semantic relationships, particularly argument and adjunct structure, between words in a phrase." Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61.				,
produce a set of logical forms. Each such logical form encodes, in a word-relation- word manner, semantic relationships, particularly argument and adjunct structure, between words in a phrase." Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61. Wupice 4:27-29 "Continuing with Example 1, suppose that the retrieved documents contain the following additional noun phrases in proximity to the noun phrase "New York City."" Kupice 11:19-24 "In step 300 the input string is analyzed to determine what part of speech each word of the string is. Each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."				· · ·
Each such logical form encodes, in a word-relationword manner, semantic relationships, particularly argument and adjunct structure, between words in a phrase." Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61.				
encodes, in a word-relation-word manner, semantic relationships, particularly argument and adjunct structure, between words in a phrase." Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61. Expose that the retrieved documents contain the following additional noun phrases in proximity to the noun phrase "New York City." Kupiec 11:19-24 "In step 300 the input string is analyzed to determine what part of speech each word of the string is. Each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."		-		
word manner, semantic relationships, particularly argument and adjunct structure, between words in a phrase." Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61. Kupiec 4:27-29 "Continuing with Example 1, suppose that the retrieved documents contain the following additional noun phrases in proximity to the noun phrase "New York City."" Kupiec 11:19-24 "In step 300 the input string is analyzed to determine what part of speech each word of the string is. Each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."				
argument and adjunct structure, between words in a phrase." Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61. Kupiec 4:27-29 "Continuing with Example 1, suppose that the retrieved documents contain the following additional noun phrases in proximity to the noun phrase "New York City."" Kupiec 11:19-24 "In step 300 the input string is analyzed to determine what part of speech each word of the string is. Each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."		word manner, semantic		
between words in a phrase." Braden See, e.g., 11:62-14:61. Braden See, e.g., 11:62-14:61. Kupiec 4:27-29 "Continuing with Example 1, suppose that the retrieved documents contain the following additional noun phrases in proximity to the noun phrase "New York City." Kupiec 11:19-24 "In step 300 the input string is analyzed to determine what part of speech each word of the string is. Each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."		relationships, particularly		word."
Braden See, e.g., 11:62-14:61. Kupiec 4:27-29 "Continuing with Example 1, suppose that the retrieved documents contain the following additional noun phrases in proximity to the noun phrase "New York City."" Kupiec 11:19-24 "In step 300 the input string is analyzed to determine what part of speech each word of the string is. Each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."				
Braden See, e.g., 11:62-14:61. Kupiec 4:27-29 "Continuing with Example 1, suppose that the retrieved documents contain the following additional noun phrases in proximity to the noun phrase "New York City."" Kupiec 11:19-24 "In step 300 the input string is analyzed to determine what part of speech each word of the string is. Each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."		between words in a phrase."		· ·
Kupiec 4:27-29 "Continuing with Example 1, suppose that the retrieved documents contain the following additional noun phrases in proximity to the noun phrase "New York City."" Kupiec 11:19-24 "In step 300 the input string is analyzed to determine what part of speech each word of the string is. Each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."				8:19, 14:65–15:10, Abstract, Figures 1–8.
suppose that the retrieved documents contain the following additional noun phrases in proximity to the noun phrase "New York City."" Kupiec 11:19-24 "In step 300 the input string is analyzed to determine what part of speech each word of the string is. Each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."		Braden See, e.g., 11:62-14:61.		W : 427.20 %G : : : : : : : : 1.1
the following additional noun phrases in proximity to the noun phrase "New York City."" Kupiec 11:19-24 "In step 300 the input string is analyzed to determine what part of speech each word of the string is. Each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."				
proximity to the noun phrase "New York City."" Kupiec 11:19-24 "In step 300 the input string is analyzed to determine what part of speech each word of the string is. Each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."				
City."" Kupiec 11:19-24 "In step 300 the input string is analyzed to determine what part of speech each word of the string is. Each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."				1
Kupiec 11:19-24 "In step 300 the input string is analyzed to determine what part of speech each word of the string is. Each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."				-
is analyzed to determine what part of speech each word of the string is. Each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."				City.
is analyzed to determine what part of speech each word of the string is. Each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."				Kupiec 11:19-24 "In step 300 the input string
each word of the string is. Each word of the string is tagged to indicate whether it is a noun, verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."				
verb, adjective, etc. Tagging can be accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."				
accomplished, for example, by a tagger that uses a hidden Markov model. The result produced by step 300 is a tagged input string."				
uses a hidden Markov model. The result produced by step 300 is a tagged input string."				
produced by step 300 is a tagged input string."				1 1 0
Kupiec 11:28-30 "In step 310, which				produced by step 300 is a tagged input string."
				Kupiec 11:28-30 "In step 310, which

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
			comprises component steps 311 and 312, the tagged input string is analyzed to detect noun
			phrases. In step 315 the tagged input string is
			further analyzed to detect main verbs."
			Kupiec 13:15-21 "The match sentences are analyzed in substantially the same manner as the input string is analyzed in step 220 above. The detected phrases typically comprise noun phrases and can further comprise title phrases or other kinds of phrases. The phrases detected in the match sentences are called preliminary hypotheses."
			Kupiec 14:45-54 "Hypotheses are verified in
			step 260 through lexico-syntactic analysis. Lexico-syntactic analysis comprises analysis of
			linguistic relations implied by lexico-syntactic
			patterns in the input string, constructing or
			generating match templates based on these relations, instantiating the templates using
			particular hypotheses, and then attempting to
			match the instantiated templates, that is, to find
			primary or secondary documents that contain text in which a hypothesis occurs in the context
			of a template."
			Belkin p. 402 "The knowledge sources are
			pattern-action productions: if the information
			on the blackboard matches the pattern of a
			knowledge source then its action can be executed. At any time, many knowledge
			sources are likely to have patterns that match
			the contents of the blackboard."
			Herz 12:55-64 teaches that textual documents
			may be profiled using word frequencies. "[A]

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
THE OUT LABOR.	Diauch-Haruei		textual attribute, such as the full text of a movie review, can be replaced by a collection of numeric attributes that represent scores to denote the presence and significance of the words in that text. The score of a word in a text may be defined in numerous ways. The simplest definition is that the score is the rate of the word in the text, which is computed by computing the number of times the word occurs in the text, and dividing this number by the total number of words in the text." Herz 13:24-27 teaches that, for the purposes of creating a profile, "one could break the text into overlapping word bigrams (sequences of 2 adjacent words), or more generally, word n-grams." Herz See also Abstract; 1:18-43; 27:47-49; 27:62-67; 4:49–8:8; 28:41–55:42; 55:44–56:14; 56:15-30; 58:57–60:9; Figures 1-16.
(i) determining a user segment count, by the computer system, for each user segment group of said at least one user segment group, each said user segment count being representative of a number of identical segments in the corresponding user segment group of said at least one user segment group, and linking each said user segment count to the corresponding user segment group of said at least one user segment group of said at least one user segment group;	Braden Abstract "Each document that has at least one matching logical forms is heuristically scored, with each different relation for a matching logical forms being assigned a different corresponding predefined weight. The score of each such document is, e.g., a predefined function of the weights of its uniquely matching logical forms. Finally, the retained documents are ranked in order	Kurtzman, II 6:39-41 "[C]reating a multidimensional vector comprised of the words and word-stems mapped to their respective frequencies." Kurtzman, II, Figs. 6, 7, and 9.	See Salton 1968, p. 96 (Fig. 3-19), above. Salton 1968, p. 95 "If a particular phrase is found in a sentence, an appropriate entry is made in a chained list of concept numbers, similar in format to the list of concepts derived by the thesaurus look-up. This concept list is kept sorted by concept number, and each concept is stored together with its weight and with coded information identifying the given concept number as a phrase concept. A typical entry in the chained list is shown in Fig. 3-18." Salton 1968, p. 95 "The number of occurrences

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
	of descending score and then		of a phrase in a given sentence determines the
	presented to a user in that		weight assigned to that phrase and is initially
	order."		defined as the minimum number of
			occurrences of each of the phrase components
	Braden 11:22-26 "Thereafter,		in the sentence. If a phrase already entered in
	through comparing the logical		the chained list is detected, appropriately
	form triples for the query		increased. Since a given text word may
	against those for each		correspond to many concept numbers, it is
	document, process 600 scores		theoretically possible that a single word may
	each document that contains at		be responsible for the generation of a complete
	least one matching logical form		phrase; such a condition is not allowed, and
	triple, then ranks these		care is taken to eliminate "phrases" where the
	particular documents based on		several components are detected in the same
	their scores."		word."
	Braden 17:44-53 "Of these		Dedrick See, e.g., 3:37–4:9, 5:34–6:3, 6:53–
	triples, two are identical, i.e.,		8:19, 14:65–15:10, Abstract, Figures 1–8.
	"HAVE-Dsub-OCTOPUS". A		0.125, 1.1.00 12.120, 1.20312.000, 1.1.50.12.00
	score for a document is		Belkin teaches determining a user segment
	illustratively a numeric sum of		count through its scheduler. p. 402-404
	the weights of all uniquely		-
	matching triples in that		Herz 78:47-50 "The method generates sets of
	document. All duplicate		search profiles for the users based on such
	matching triples for any		attributes as the relative frequency of
	document are ignored. An		occurrence of words in the articles readby the
	illustrative ranking of the		users, and uses these search profiles to
	relative weightings of the		efficiently identify future articles of interest."
	different types of relations that		
	can occur in a triple, in		Herz 12:55-64 teaches that textual documents
	descending order from their		may be profiled using word frequencies. "[A]
	largest to smallest weightings		textual attribute, such as the full text of a
	are: first, verb-object		movie review, can be replaced by a collection
	combinations (Dobj); verb-		of numeric attributes that represent scores to
	subject combinations (Dsub);		denote the presence and significance of the
	prepositions and operators (e.g.		words in that text. The score of a word in a
	Ops), and finally modifiers		text may be defined in numerous ways. The
	(e.g. Nadj)."		simplest definition is that the score is the rate

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
	Braden 25:41-48 "Rather than using fixed weights for each different attribute in a logical form triple, these weights can dynamically vary and, in fact, can be made adaptive. To accomplish this, a learning mechanism, such as, e.g., a Bayesian or neural network, could be appropriately incorporated into our inventive process to vary the numeric weight for each different logical form triple to an optimal value based upon learned experiences."		of the word in the text, which is computed by computing the number of times the word occurs in the text, and dividing this number by the total number of words in the text." Herz 13:24-27 teaches that, for the purposes of creating a profile, "one could break the text into overlapping word bigrams (sequences of 2 adjacent words), or more generally, word n-grams." Herz <i>See also</i> Abstract; 1:18-43; 27:47-49; 27:62-67; 4:49–8:8; 28:41–55:42; 55:44–56:14; 56:15-30; 58:57–60:9; Figures 1-16.
(j) sorting the user segment groups of said at least one user segment group, by the computer system, in an descending order of user segment counts starting from a user segment group having a highest user segment count, and recording said user segment groups and corresponding user segment counts in said user data profile; and	Braden Abstract "Each document that has at least one matching logical forms is heuristically scored, with each different relation for a matching logical forms being assigned a different corresponding predefined weight. The score of each such document is, e.g., a predefined function of the weights of its uniquely matching logical forms. Finally, the retained documents are ranked in order of descending score and then presented to a user in that order."	Kurtzman, II 6:12-13 "[C]reating a content data structure which indicates features of the content having particular characteristics." Kurtzman, II, Figs. 6, 7, and 9.	See Salton 1968, p. 96 (Fig. 3-19), above. Salton teaches sorting segment counts. See Salton 1968, p. 91 (Fig. 3-16)(Concept Nos.) Dedrick See, e.g., 3:37–4:9, 5:34–6:3, 6:53–8:19, 14:65–15:10, Abstract, Figures 1–8. Belkin teaches determining a user segment count through its scheduler. p. 402-404. Herz 78:47-50 "The method generates sets of search profiles for the users based on such attributes as the relative frequency of occurrence of words in the articles readby the users, and uses these search profiles to efficiently identify future articles of interest." Herz 12:55-64 teaches that textual documents

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
	Braden 11:22-26 "Thereafter,		may be profiled using word frequencies. "[A]
	through comparing the logical		textual attribute, such as the full text of a
	form triples for the query		movie review, can be replaced by a collection
	against those for each		of numeric attributes that represent scores to
	document, process 600 scores		denote the presence and significance of the
	each document that contains at		words in that text. The score of a word in a
	least one matching logical form		text may be defined in numerous ways. The
	triple, then ranks these		simplest definition is that the score is the rate
	particular documents based on		of the word in the text, which is computed by
	their scores."		computing the number of times the word
	5 1 15 14 50 ((0.0.1)		occurs in the text, and dividing this number by
	Braden 17:44-53 "Of these		the total number of words in the text."
	triples, two are identical, i.e.,		H 12.24.27
	"HAVE-Dsub-OCTOPUS". A		Herz 13:24-27 teaches that, for the purposes of
	score for a document is		creating a profile, "one could break the text
	illustratively a numeric sum of		into overlapping word bigrams (sequences of 2
	the weights of all uniquely		adjacent words), or more generally, word n-grams."
	matching triples in that document. All duplicate		grams.
	matching triples for any		Herz See also Abstract; 1:18-43; 27:47-49;
	document are ignored. An		27:62-67; 4:49–8:8; 28:41–55:42; 55:44–
	illustrative ranking of the		56:14; 56:15-30; 58:57–60:9; Figures 1-16.
	relative weightings of the		30.14, 30.13-30, 30.37-00.7, Figures 1-10.
	different types of relations that		
	can occur in a triple, in		
	descending order from their		
	largest to smallest weightings		
	are: first, verb-object		
	combinations (Dobj); verb-		
	subject combinations (Dsub);		
	prepositions and operators (e.g.		
	Ops), and finally modifiers		
	(e.g. Nadj)."		
	D 1 25 41 40 "D 1 1		
	Braden 25:41-48 "Rather than		
	using fixed weights for each		
	different attribute in a logical		

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
	form triple, these weights can dynamically vary and, in fact, can be made adaptive. To accomplish this, a learning mechanism, such as, e.g., a Bayesian or neural network, could be appropriately incorporated into our inventive process to vary the numeric weight for each different logical form triple to an optimal value based upon learned experiences."		
(k) storing, by the computer system, said user data profile, representative of an overall linguistic pattern of the user, in the data storage system, said overall linguistic pattern substantially corresponding to the user's social, cultural, educational, economic background and to the user's psychological profile.	Braden 7:19-23 "Generally speaking and in accordance with our present invention, we have recognized that precision of a retrieval engine can be significantly enhanced by employing natural language processing to process, i.e., specifically filter and rank, the records, i.e., ultimately the documents, provided by a search engine used therein." Braden See, e.g., 11:62-14:61.	Kurtzman, II 3:47-49 "These specific files selected and viewed by the user are recorded by the affinity server." Kurtzman, II 3:21- 23 "User profiling uses content stream analysis, as well as demographic, geographic, psychographic, digital identification, and HTTP information."	See Salton 1968, p. 96 (Fig. 3-19), above. Salton 1968, p. 91 "After dictionary look-up, weight assignment, and the merging of concepts derived from various sources, the document is reduced to a merged concept vector, as shown for a typical document in Fig. 3-16." Chislenko 4:15-18 "For example, the system may assume that Web sites for which the user has created "bookmarks" are liked by that user and may use those sites as initial entries in the user's profile." Chislenko 4:40-50 "Ratings can be inferred by the system from the user's usage pattern. For example, the system may monitor how long the user views a particular Web page and store in that user's profile an indication that the user likes the page, assuming that the longer the user views the page, the more the user likes the

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
			page. Alternatively, a system may monitor the
			user's actions to determine a rating of a
			particular item for the user. For example, the
			system may infer that a user likes an item
			which the user mails to many people and enter
			in the user's profile and indication that the user
			likes that item."
			Dedrick 3:54–4:4 "The GUI may also have
			hidden fields relating to "consumer variables."
			Consumer variables refer to demographic,
			psychographic and other profile information.
			Demographic information refers to the vital
			statistics of individuals, such as age, sex,
			income and marital status. Psychographic
			information refers to the lifestyle and
			behavioral characteristics of individuals, such
			as likes and dislikes, color preferences and
			personality traits that show consumer
			behavioral characteristics. Thus, the consumer
			variables refer to information such as marital
			status, color preferences, favorite sizes and shapes, preferred learning modes, employer,
			job title, mailing address, phone number,
			personal and business areas of interest, the
			willingness to participate in a survey, along
			with various lifestyle information. This
			information will be referred to as user profile
			data, and is stored on a consumer owned
			portable profile device such as a Flash
			memory-based PCMClA pluggable card."
			Dadwids C 2,27, 4,0, 5,24, 6,2, 6,52
			Dedrick <i>See</i> , <i>e.g.</i> , 3:37–4:9, 5:34–6:3, 6:53–8:19, 14:65–15:10, Abstract, Figures 1–8.
			6.17, 14.03–13.10, Austract, Figures 1–6.
			Belkin p. 399 "In the general information
			seeking interaction, the IR system needs to

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
			have (see Table 1 for a brief listing of the ten functions and their acronyms): a model of the user himself, including goals, intentions and
			experience (UM)." Belkin p. 403 "For I ³ R to be adaptable, it must
			be able to assess the user's abilities so it can adjust the interface to match them.[22] This requires a user model builder. As each user
			may have his own view of the subject area being searched, it would be valuable to capture this information and remember it from session to session in a domain knowledge expert."
			Herz 27:62-66 teaches generating user profiles based on a wide variety of attributes. "In a variation, each user's user profile is subdivided
			into a set of long-term attributes, such as demographic characteristics, and a set of shortterm attributes that help to identify the user's temporary desires and emotional state."
			Herz 20:35-37 "User profiles may make use of any attributes that are useful in characterizing humans."
			Herz 11:31-38 "written response[s] to Rorschach inkblot test," "multiple-choice responses by [the person] to 20 self-image questions," as well as "their literary tastes and
			psychological peculiarities." Herz 32:39-49 "A second function of the proxy
			server is to record user-specific information associated with user U All of this user-specific information is stored in a database that
			is keyed by user U's pseudonym (whether

The '067 Patent	Braden-Harder	Kurtzman, II	Additional Prior Art References
			secure or non-secure) on the proxy server."
			Herz 66:65-67; 67:1-3 "The system uses the method of section 'Searching for Target Objects' above to automatically locate a small set of one or more clusters with profiles similar to the query profile, for example, the articles they contain are written at roughly an 8th-grade level and tend to mention Galileo and the Medicis."
			Herz <i>See also</i> Abstract; 1:18-43; 27:47-49; 27:62-67; 4:49–8:8; 28:41–55:42; 55:44–56:14; 56:15-30; 58:57–60:9; Figures 1-16.
			Culliss 3:46-48 "Inferring Personal Data. Users can explicitly specify their own personal data, or it can be inferred from a history of their search requests or article viewing habits. In this respect, certain key words or terms, such as those relating to sports (i.e. "football" and "soccer"), can be detected within search requests and used to classify the user as someone interested in sports."