Exhibit A

US006199067B1

(12) **United States Patent**
Geller

(10) **Patent No.: US 6,199,067 B1**
(45) **Date of Patent: Mar. 6, 2001**

(54) **SYSTEM AND METHOD FOR GENERATING PERSONALIZED USER PROFILES AND FOR UTILIZING THE GENERATED USER PROFILES TO PERFORM ADAPTIVE INTERNET SEARCHES**

(75) Inventor: **Ilya Geller**, Brooklyn, NY (US)

(73) Assignee: **Mightiest Logicon Unisearch, Inc.**, Brooklyn, NY (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days

(21) Appl. No.: **09/422,286**

(22) Filed: **Oct. 21, 1999**

Related U.S. Application Data

(60) Provisional application No. 60/116,582, filed on Jan. 20, 1999

(51) **Int. Cl.**[7] **G06F 17/30**

(52) **U.S. Cl.** **707/10**; 707/3; 707/5; 707/100; 704/4; 704/9; 704/247; 700/17

(58) **Field of Search** 707/1–5, 100, 707/10, 200; 704/1, 4, 9, 270.1, 247, 250; 382/181, 190, 209; 700/11, 17, 56, 83, 86; 369/13

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,914,590	*	4/1990	Loatman et al.	704/8
5,483,650	*	1/1996	Pedersen et al.	707/2
5,696,962	*	12/1997	Kupiec	707/4

(List continued on next page.)

OTHER PUBLICATIONS

Helm, Richard et al., "Integrating Information Retrieval and Domain Specific Approaches for Browsing and Retrieval in Object–Oriented Class Libraries," Conference proceedings on Object–Oriented programming systems, languages, and applications May 1991 p *

Maass, Henning. "Location–aware mobile applications based on dictionary services," Mobile Networks and Applications, vol. 3, Issue 2, Aug. 1998, pp. 157–173 *

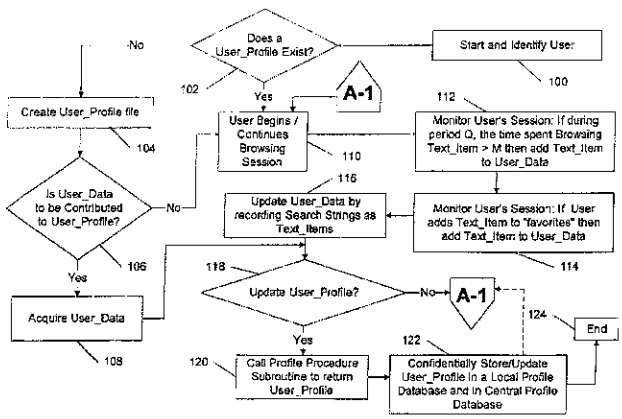
Xu, Yaowu et al., "Hierarchical Content Description and Object Formation by Learning," Proceedings of the IEEE Workshop on Content–Based Access of Image and Video Libraries, (CBAIVL '99), Jun. 22, 1999, pp. 84–88, May 1991 *

Primary Examiner—Hosain T. Alam
Assistant Examiner—Shahid Alam
(74) *Attorney, Agent, or Firm*—Edward Etkin, Esq.

(57) **ABSTRACT**

A system and method for automatically generating personalized user profiles and for utilizing the generated profiles to perform adaptive Internet or computer data searches is provided. In accordance with the present invention, particular linguistic patterns and their frequency of recurrence are extracted from personal texts provided by the users of the system of the present invention and stored in a user profile data file such that the user profile data file is representative of the user's overall linguistic patterns and the frequencies of recurrence thereof. All documents in a remote computer system, such as the Internet, are likewise analyzed and their linguistic patterns and pattern frequencies are also extracted and stored in corresponding document profiles. When a search for particular data is initiated by the user, linguistic patterns are also extracted from a search string provided by the user into a search profile. The user profile is then cross matched with the search profile and the document profiles to determine whether any linguistic patterns match in all three profiles and to determine the magnitude of the match based on summation of respective frequencies of recurrence of the matching patterns. The documents with document profiles having the highest matching magnitudes are presented to the user as not only matching the subject of the search string, but also as corresponding to the user's cultural, educational, and social backgrounds as well as the user's psychological profile.

62 Claims, 8 Drawing Sheets



U S PATENT DOCUMENTS

5,822,748	*	10/1998	Cohen et al	7074/2
5,848,408	*	12/1998	Jakobsson et al	707/3
5,943,648	*	8/1999	Tel	704/260
5,974,408	*	10/1999	Cohen et al	707/2
6,016,487	*	1/2000	Rioux et al	707/2
6,018,734	*	1/2000	Zhang et al	707/3
6,081,750	*	6/2000	Hoffberg et al	700/17

* cited by examiner

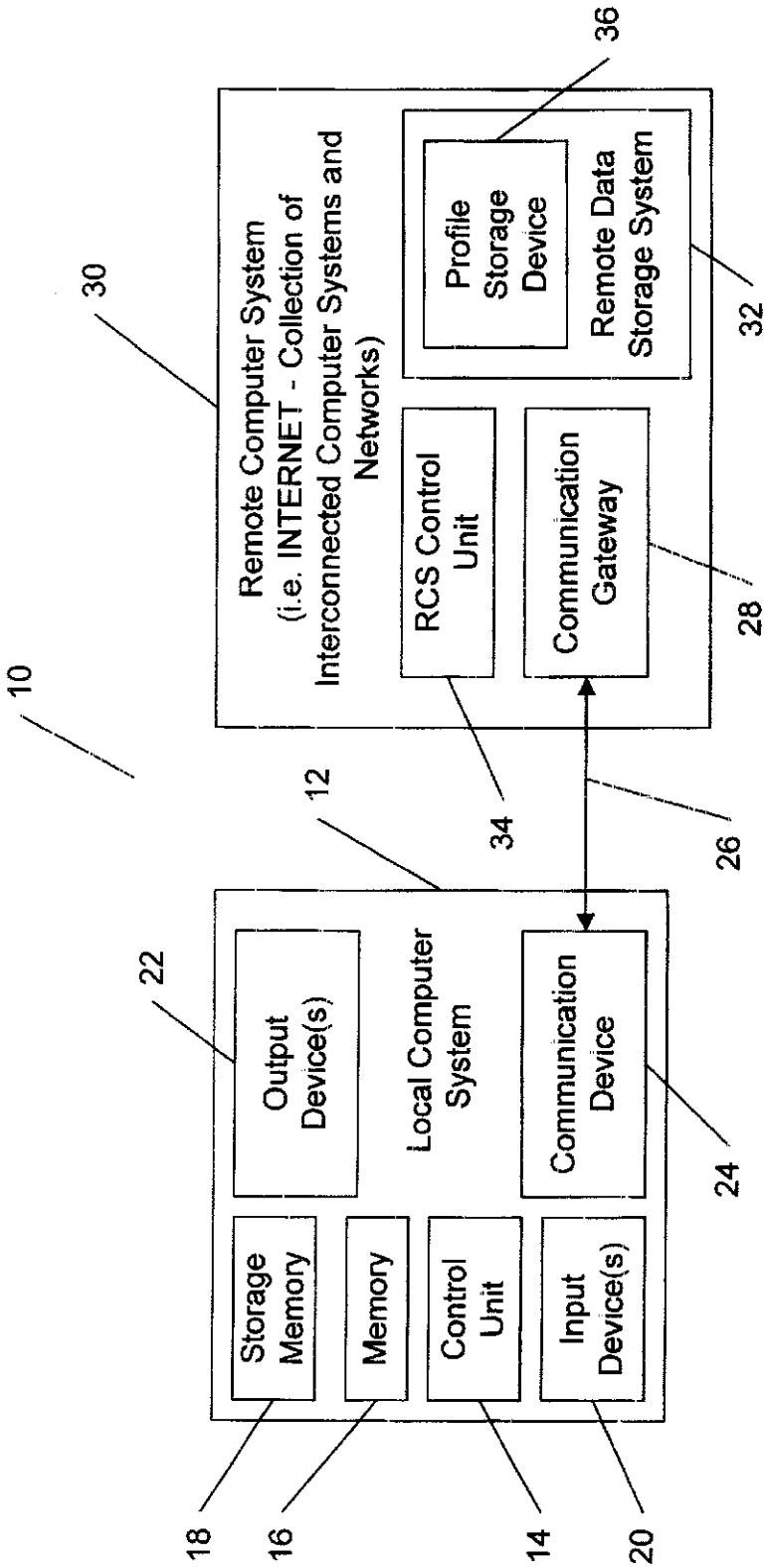

10
12
14
16
18
20
22
24
26
28
30
32
34
36
Remote Computer System (i.e. INTERNET - Collection of Interconnected Computer Systems and Networks)
Profile Storage Device
Remote Data Storage System
RCS Control Unit
Communication Gateway
Local Computer System
Output Device(s)
Communication Device
Storage Memory
Memory
Control Unit
Input Device(s)

FIG. 1

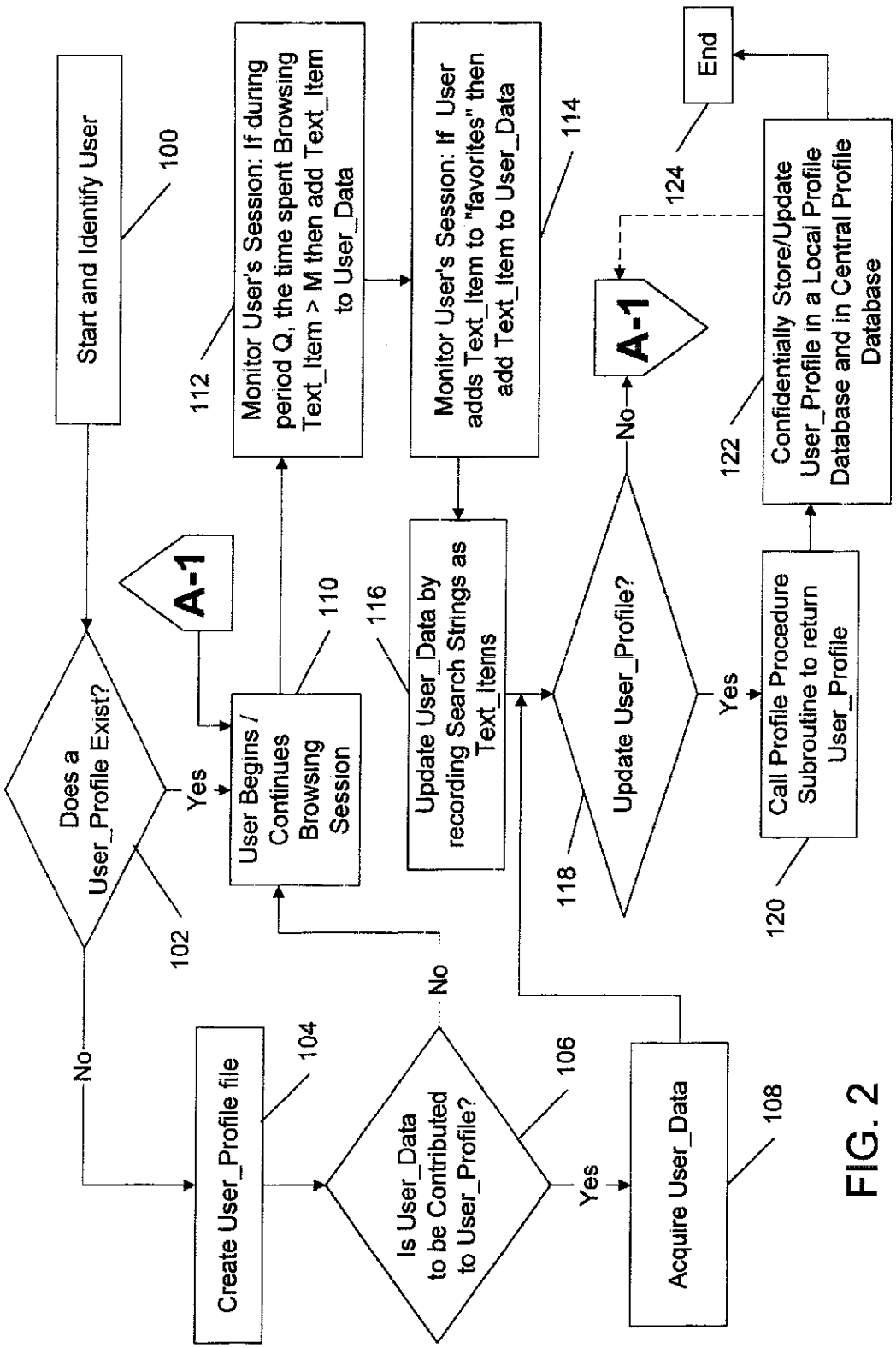

Start and Identify User
100
Does a User_Profile Exist?
102
Yes
No
A-1
User Begins / Continues Browsing Session
110
Monitor User's Session: If during period Q, the time spent Browsing Text_Item > M then add Text_Item to User_Data
112
Monitor User's Session: If User adds Text_Item to "favorites" then add Text_Item to User_Data
114
Update User_Data by recording Search Strings as Text_Items
116
Update User_Profile?
118
Yes
No
A-1
Call Profile Procedure Subroutine to return User_Profile
120
Confidentially Store/Update User_Profile in a Local Profile Database and in Central Profile Database
122
End
124
Create User_Profile file
104
Is User_Data to be Contributed to User_Profile?
106
No
Yes
Acquire User_Data
108

FIG. 2

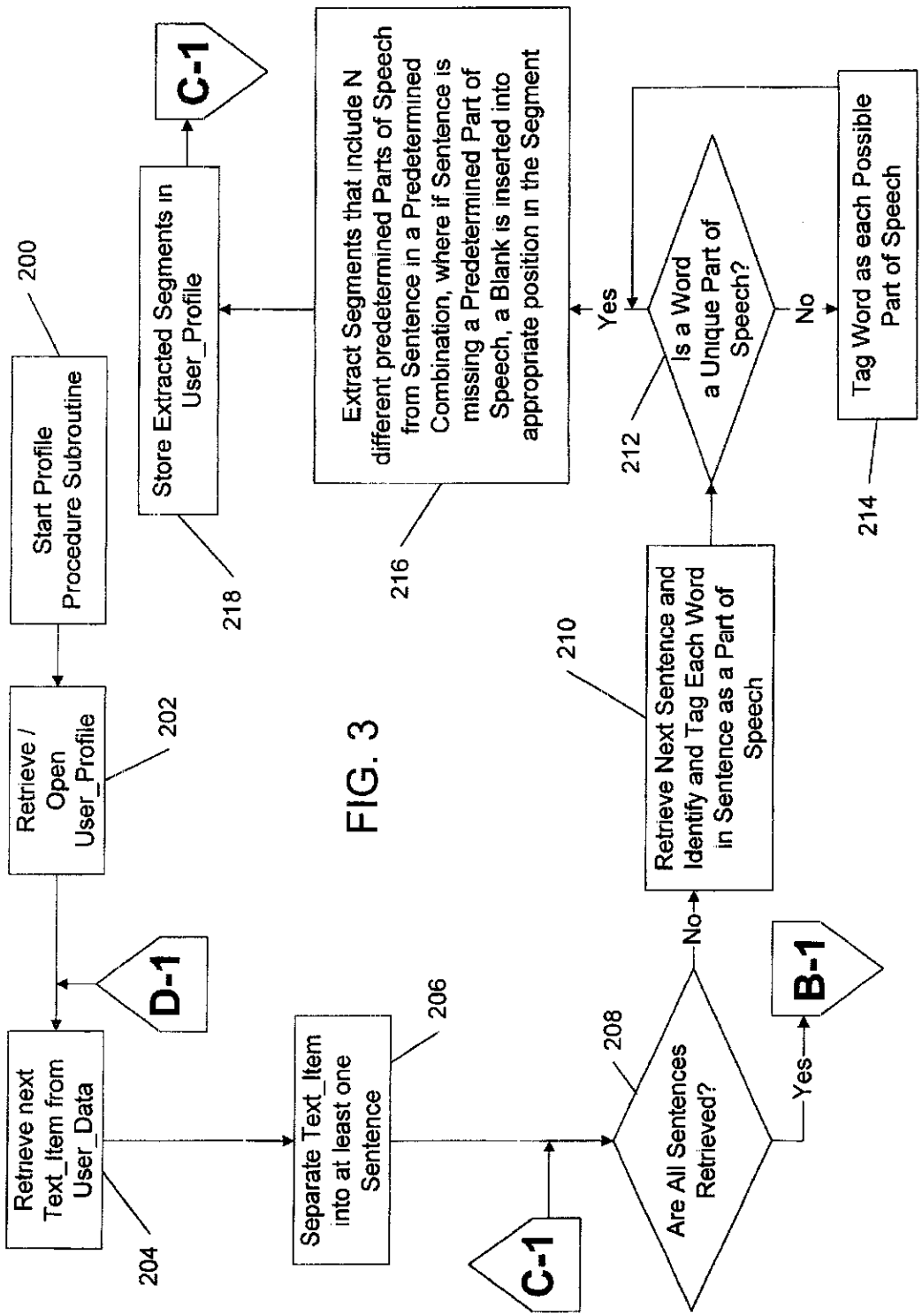
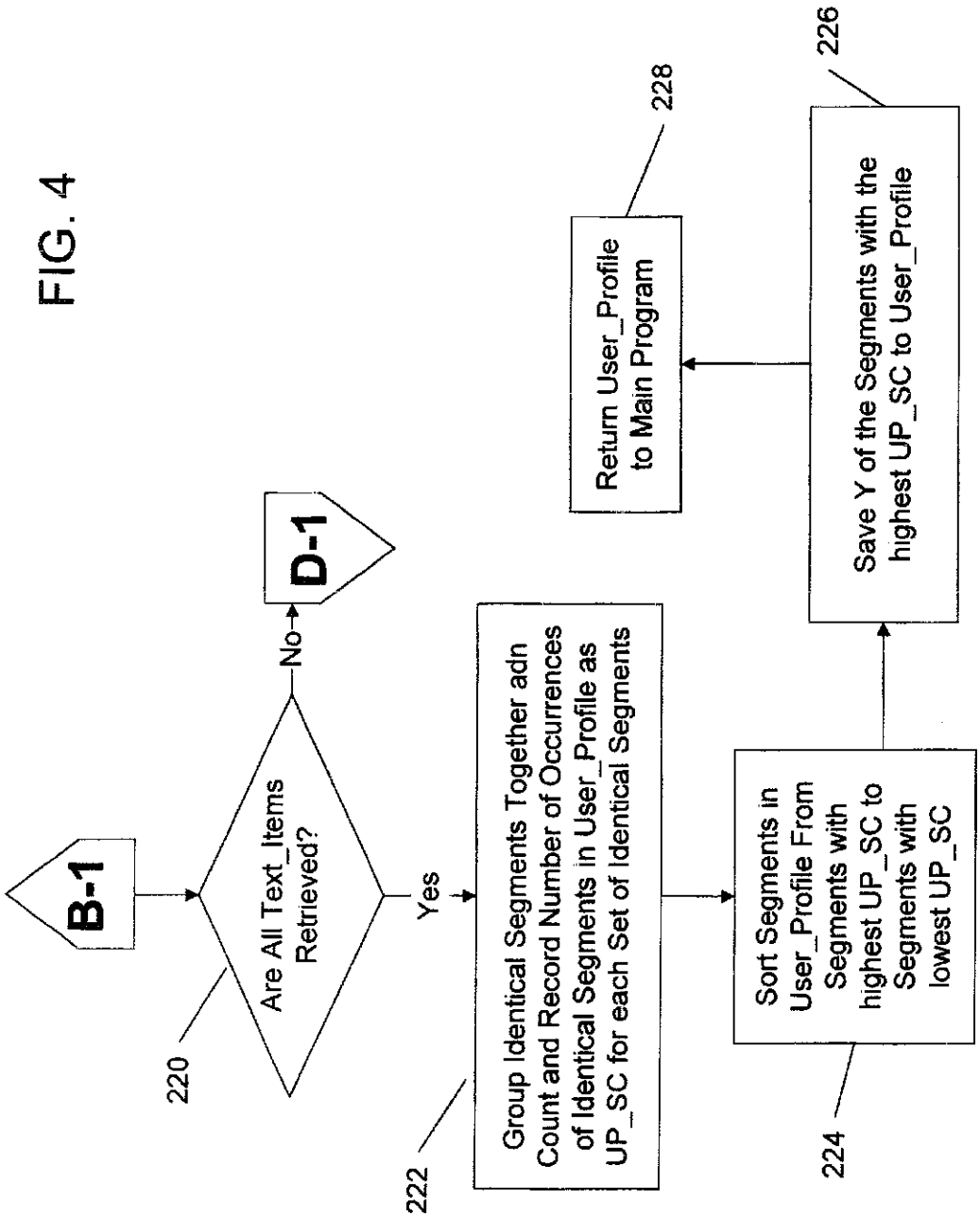

Start Profile Procedure Subroutine
200
Retrieve / Open User_Profile
202
D-1
Retrieve next Text_Item from User_Data
204
Separate Text_Item into at least one Sentence
206
C-1
Are All Sentences Retrieved?
208
No
Yes
B-1
Retrieve Next Sentence and Identify and Tag Each Word in Sentence as a Part of Speech
210
Is a Word a Unique Part of Speech?
212
Yes
No
Tag Word as each Possible Part of Speech
214
Extract Segments that include N different predetermined Parts of Speech from Sentence in a Predetermined Combination, where if Sentence is missing a Predetermined Part of Speech, a Blank is inserted into appropriate position in the Segment
216
Store Extracted Segments in User_Profile
218
C-1

FIG. 3

FIG. 4


B-1
Are All Text_Items Retrieved?
220
No
D-1
Yes
Group Identical Segments Together adn Count and Record Number of Occurrences of Identical Segments in User_Profile as UP_SC for each Set of Identical Segments
222
Sort Segments in User_Profile From Segments with highest UP_SC to Segments with lowest UP_SC
224
Save Y of the Segments with the highest UP_SC to User_Profile
226
Return User_Profile to Main Program
228

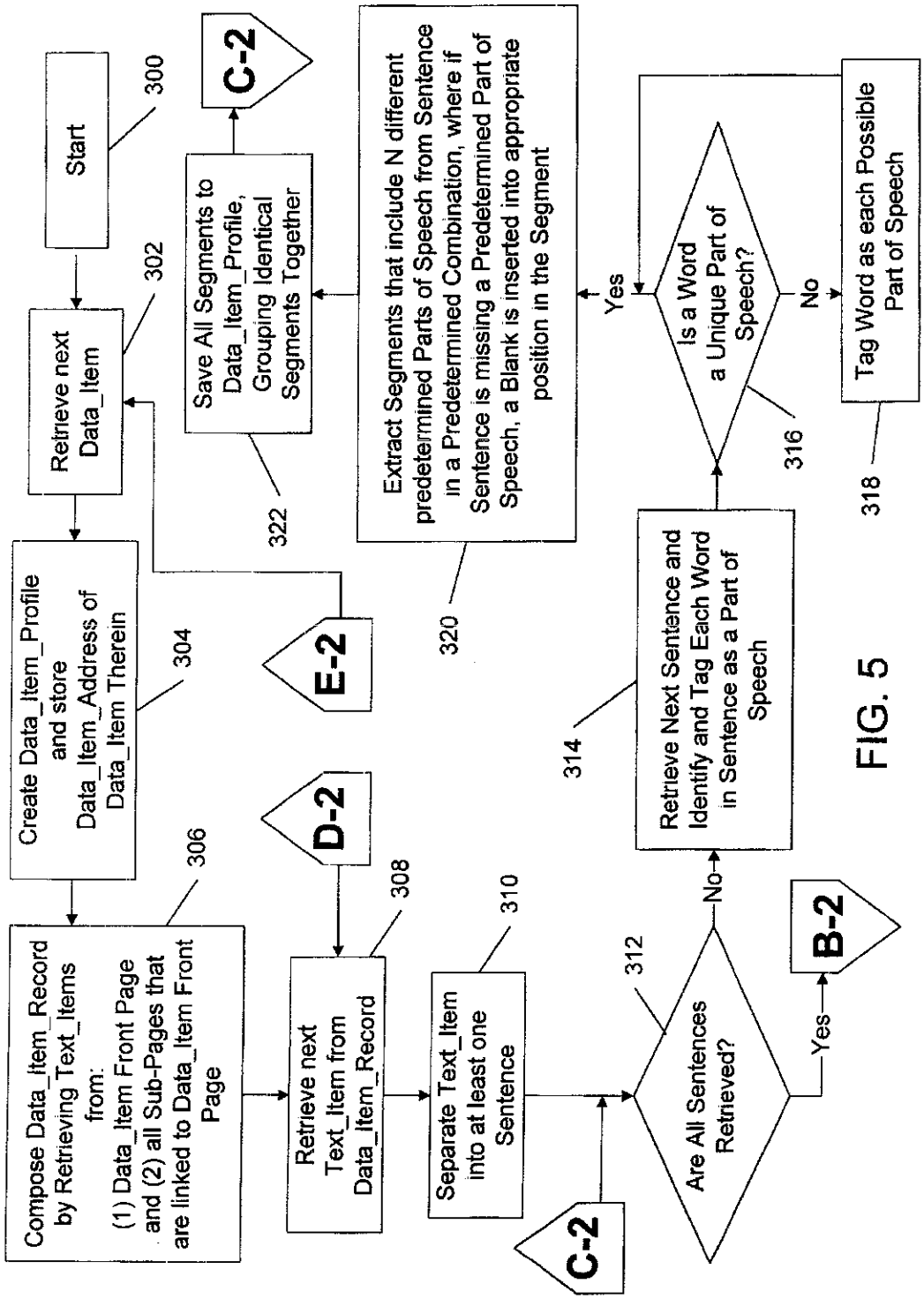
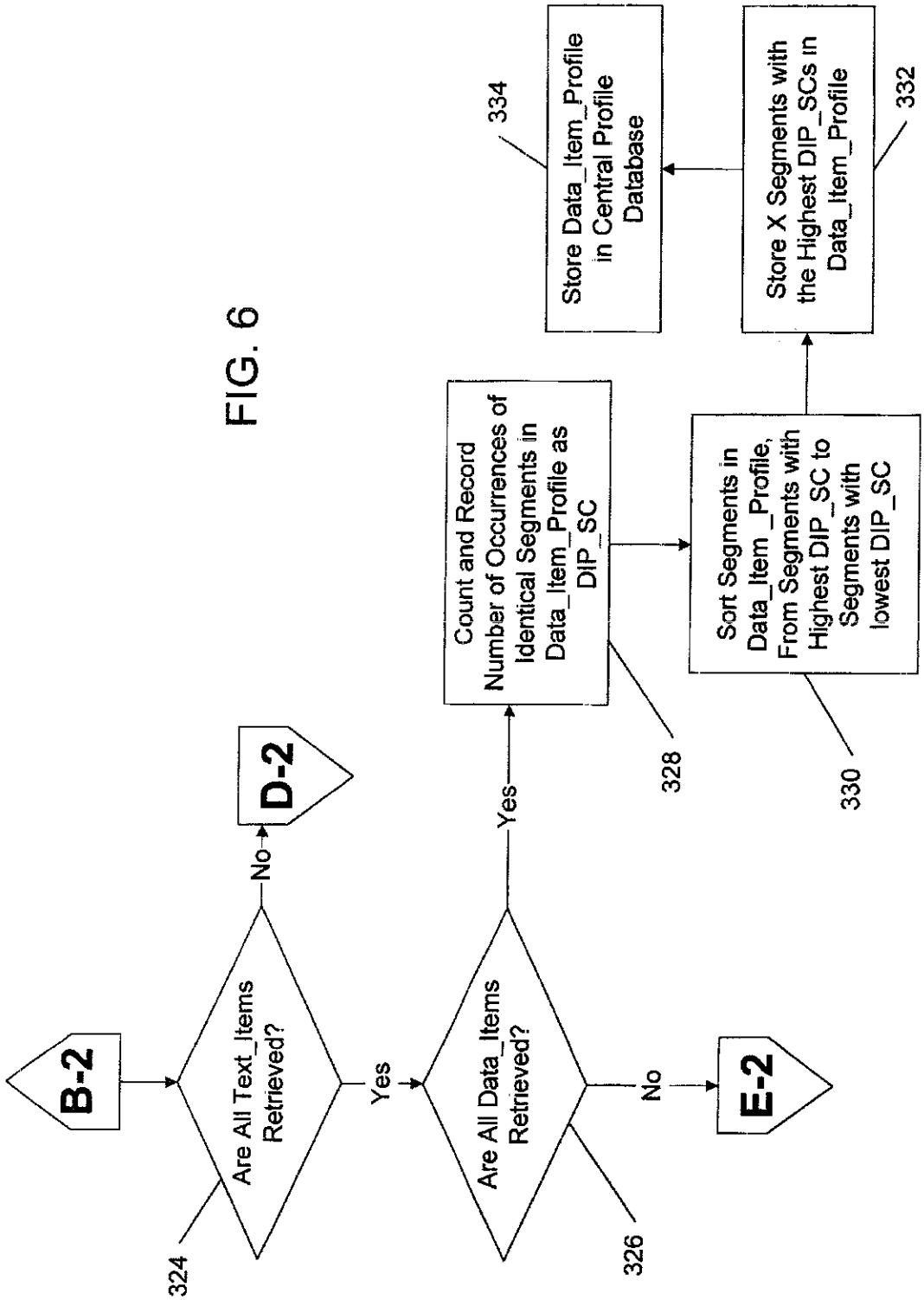

Start
300
Retrieve next Data_Item
302
Create Data_Item_Profile and store Data_Item_Address of Data_Item Therein
304
Compose Data_Item_Record by Retrieving Text_Items from: (1) Data_Item Front Page and (2) all Sub-Pages that are linked to Data_Item Front Page
306
Retrieve next Text_Item from Data_Item_Record
308
Separate Text_Item into at least one Sentence
310
Are All Sentences Retrieved?
312
Retrieve Next Sentence and Identify and Tag Each Word in Sentence as a Part of Speech
314
Is a Word a Unique Part of Speech?
316
Tag Word as each Possible Part of Speech
318
Extract Segments that include N different predetermined Parts of Speech from Sentence in a Predetermined Combination, where if Sentence is missing a Predetermined Part of Speech, a Blank is inserted into appropriate position in the Segment
320
Save All Segments to Data_Item_Profile, Grouping Identical Segments Together
322
Yes
No
No
Yes
C-2
C-2
D-2
E-2
B-2

FIG. 5

FIG. 6


B-2
Are All Text_Items Retrieved?
No
D-2
Yes
Are All Data_Items Retrieved?
No
E-2
Yes
Count and Record Number of Occurrences of Identical Segments in Data_Item_Profile as DIP_SC
Sort Segments in Data_Item_Profile, From Segments with Highest DIP_SC to Segments with lowest DIP_SC
Store X Segments with the Highest DIP_SCs in Data_Item_Profile
Store Data_Item_Profile in Central Profile Database
324
326
328
330
332
334

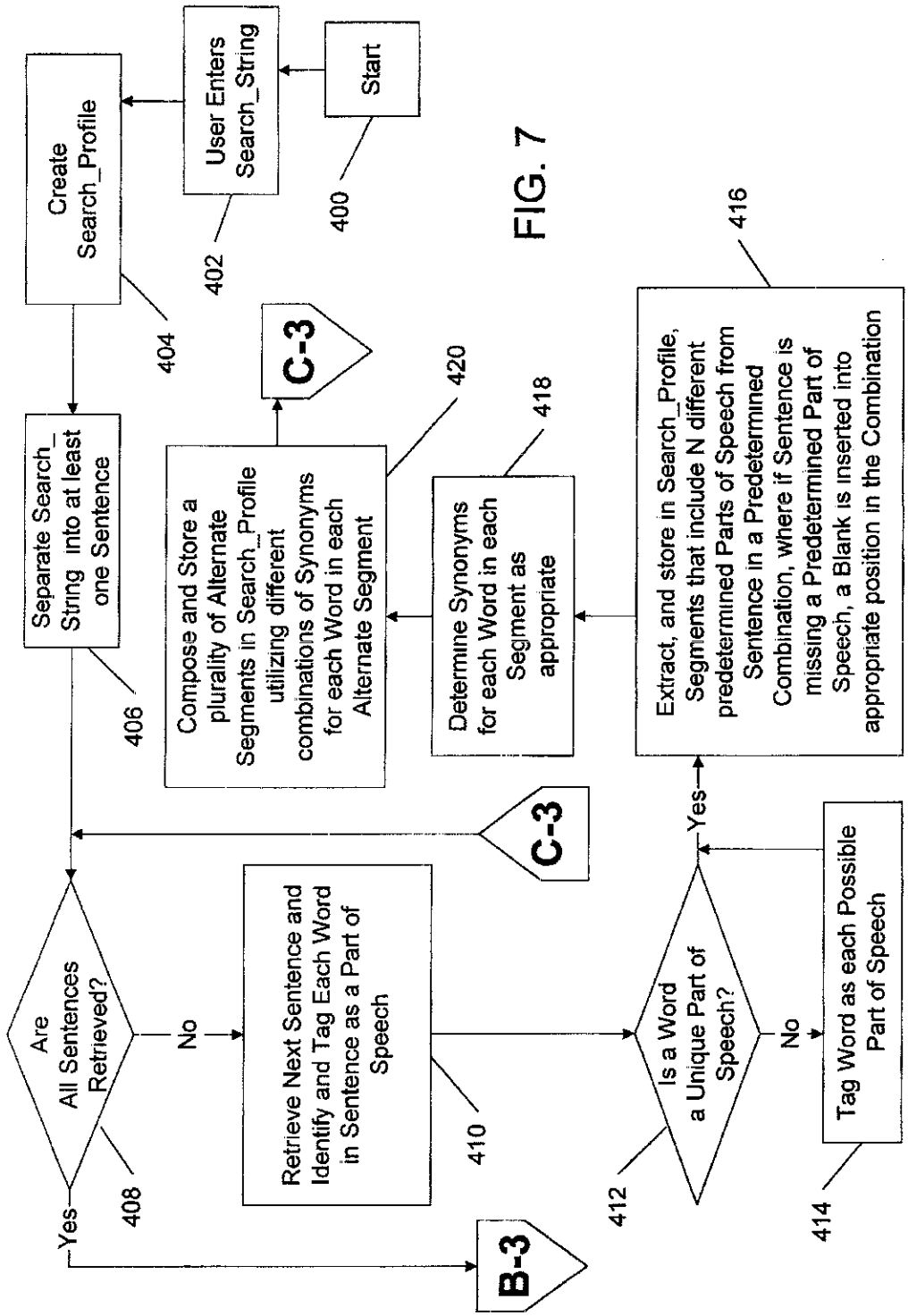

Start
400
User Enters Search_String
402
Create Search_Profile
404
Separate Search_ String into at least one Sentence
406
Are All Sentences Retrieved?
408
Yes
No
Retrieve Next Sentence and Identify and Tag Each Word in Sentence as a Part of Speech
410
Is a Word a Unique Part of Speech?
412
Yes
No
Tag Word as each Possible Part of Speech
414
Extract, and store in Search_Profile, Segments that include N different predetermined Parts of Speech from Sentence in a Predetermined Combination, where if Sentence is missing a Predetermined Part of Speech, a Blank is inserted into appropriate position in the Combination
416
Determine Synonyms for each Word in each Segment as appropriate
418
Compose and Store a plurality of Alternate Segments in Search_Profile utilizing different combinations of Synonyms for each Word in each Alternate Segment
420
C-3
C-3
B-3

FIG. 7

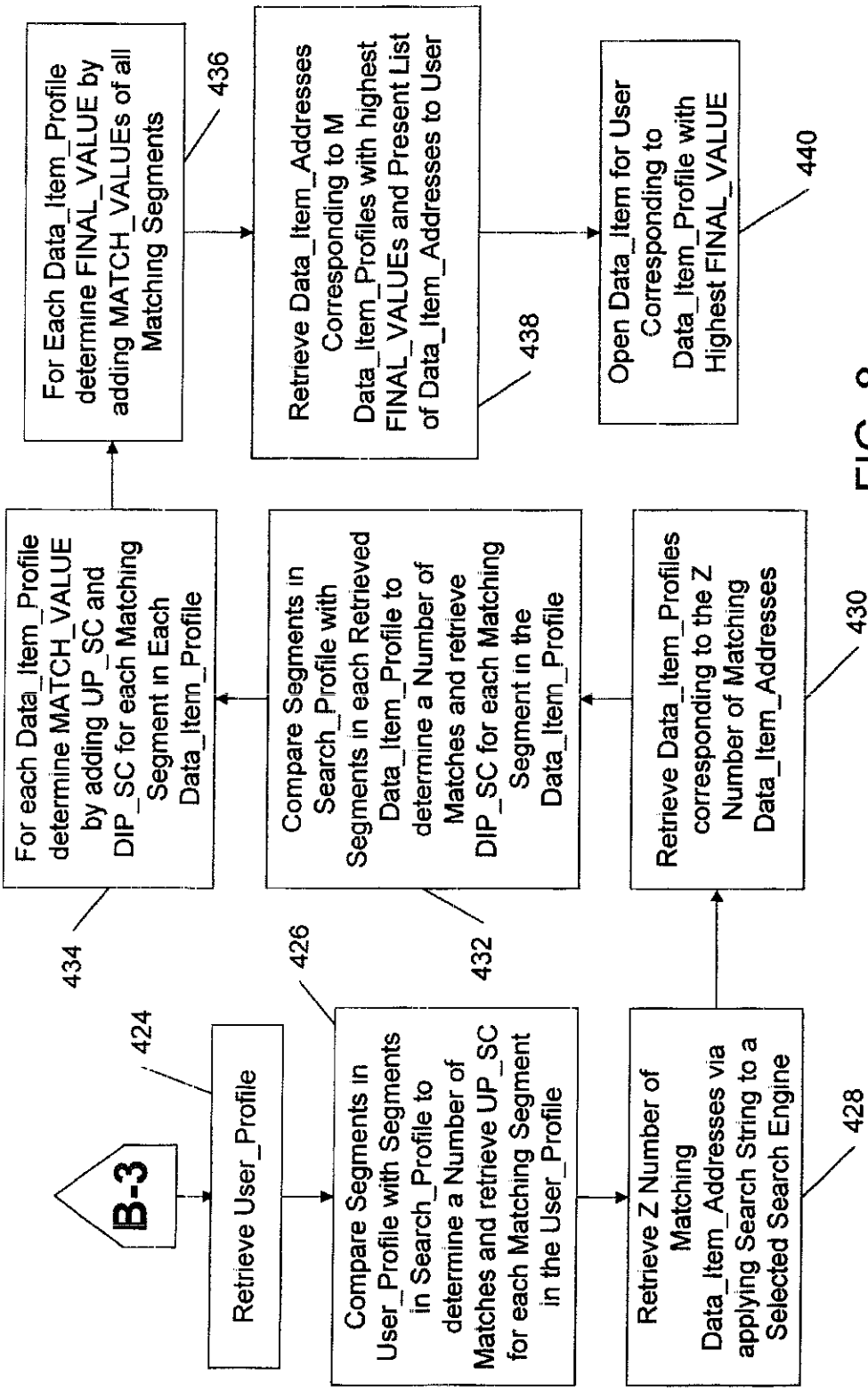

B-3
Retrieve User_Profile
424
Compare Segments in User_Profile with Segments in Search_Profile to determine a Number of Matches and retrieve UP_SC for each Matching Segment in the User_Profile
426
Retrieve Z Number of Matching Data_Item_Addresses via applying Search String to a Selected Search Engine
428
For each Data_Item_Profile determine MATCH_VALUE by adding UP_SC and DIP_SC for each Matching Segment in Each Data_Item_Profile
434
Compare Segments in Search_Profile with Segments in each Retrieved Data_Item_Profile to determine a Number of Matches and retrieve DIP_SC for each Matching Segment in the Data_Item_Profile
432
Retrieve Data_Item_Profiles corresponding to the Z Number of Matching Data_Item_Addresses
430
For Each Data_Item_Profile determine FINAL_VALUE by adding MATCH_VALUEs of all Matching Segments
436
Retrieve Data_Item_Addresses Corresponding to M Data_Item_Profiles with highest FINAL_VALUEs and Present List of Data_Item_Addresses to User
438
Open Data_Item for User Corresponding to Data_Item_Profile with Highest FINAL_VALUE
440

FIG. 8