US006185614B1

(12) **United States Patent** (10) Patent No.: **US 6,185,614 B1**
Cuomo et al. (45) **Date of Patent:** **Feb. 6, 2001**

(54) **METHOD AND SYSTEM FOR COLLECTING USER PROFILE INFORMATION OVER THE WORLD-WIDE WEB IN THE PRESENCE OF DYNAMIC CONTENT USING DOCUMENT COMPARATORS**

(75) Inventors: **Gennaro A. Cuomo**, Apex; **Binh Q. Nguyen**, Cary; **Sandeep K. Singhal**, Raleigh, all of NC (US)

(73) Assignee: **International Business Machines Corp.**, Armonk, NY (US)

(*) Notice: Under 35 U.S.C. 154(b), the term of this patent shall be extended for 0 days.

(21) Appl. No.: **09/084,452**

(22) Filed: **May 26, 1998**

(51) **Int. Cl.^{7}** **G06F 15/173**; G06F 15/16; G06F 7/00
(52) **U.S. Cl.** **709/224**; 709/203; 707/104
(58) **Field of Search** 709/203, 224; 707/6, 10, 104, 501, 513, 3, 5

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,649,186	*	7/1997	Ferguson	707/10
5,732,218		3/1998	Bland	709/204
5,740,430	*	4/1998	Rosenberg et al.	707/200
5,745,900	*	4/1998	Burrows	707/102
5,813,007	*	9/1998	Nielsen	707/10
5,890,164	*	3/1999	Nielsen	707/201
5,892,917		4/1999	Myerson	709/204
5,893,908	*	4/1999	Cullen et al.	707/5
5,895,470	*	4/1999	Pirolli et al.	707/102
5,898,836	*	4/1999	Freivald et al.	709/218
5,909,677	*	6/1999	Broder et al.	707/3
5,913,208	*	6/1999	Brown et al.	707/3
5,941,944	*	8/1999	Messerly	709/203
5,978,842	*	11/1999	Noble et al.	709/218
5,983,268	*	11/1999	Freivald et al.	709/218
5,987,480	*	11/1999	Donohue et al.	707/501
5,999,929	*	12/1999	Goodman	707/7
6,012,087	*	1/2000	Freivald et al.	709/218

FOREIGN PATENT DOCUMENTS

9831155   7/1998   (WO) .

OTHER PUBLICATIONS

Brin, S., et al., "Copy Detection Mechansims for Digital Documents," Proc. Of the 1995 ACM SIGMOD Int'l. Conf. on Management of Data, ACM, pp. 398–409, May, 1995.*
Garcia–Molina, H., et al, "dSCAM: Finding Document Copies Across Multiple Databases," Proc. of the 4th Int'l. Conf.on Parallel and Distributed Information Systems, IEEE, pp. 68–79, May 1995.*

* cited by examiner

Primary Examiner—Ahmad F. Matar
Assistant Examiner—Andrew Caldwell
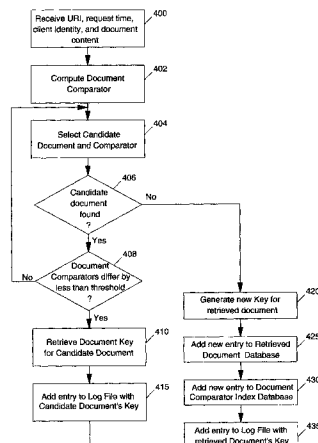(74) Attorney, Agent, or Firm—A. Bruce Clay

(57) **ABSTRACT**

Disclosed is a method and system for collecting profile information about users accessing dynamically generated content from one or more servers. In a specific embodiment, a server dynamically generates a web page in response to a user request. The server customizes the web page content based on the requested universal resource identifier (URI) and one or more of: the user's identity, access permissions, demographic information, and previous behavior at the site. The web server then passes the URI, user identity, and dynamically generated web page to an access information collector. The access information collector generates document comparators from the current web page content and compares them to document comparators associated with previously retrieved web pages. If the current web page is sufficiently similar to some previously retrieved web page, the access information collector logs the URI, user identity, and a document key associated with the matching previously retrieved page. Otherwise, the access information collector generates a new key; stores the new key and the document comparators in a database; and logs the URI, user identity, and the newly generated document key.

27 Claims, 4 Drawing Sheets

FIG. 1

Local Area Network

Local Area Network

Gateway Server

Local Area Network

FIG. 2

FIG. 3

```
        ┌─────────────────────────┐  400
        │ Receive URI, request time,│
        │ client identity, and document│
        │         content          │
        └─────────────────────────┘
                    │
                    ▼
        ┌─────────────────────────┐  402
        │   Compute Document       │
        │     Comparator           │
        └─────────────────────────┘
                    │
                    ▼
        ┌─────────────────────────┐  404
        │   Select Candidate       │
        │ Document and Comparator  │
        └─────────────────────────┘
                    │
                    ▼
              ◇ 406
           Candidate            No
           document  ─────────────────┐
           found                      │
             ?                        │
             │ Yes                    │
             ▼                        │
           ◇ 408                      │
     No   Document                    │
  ◄────── Comparators differ by       │
          less than threshold         │
             ?                        ▼
             │ Yes         ┌──────────────────────┐ 420
             ▼             │  Generate new Key for │
   ┌──────────────────┐ 410│   retrieved document  │
   │ Retrieve Document │    └──────────────────────┘
   │ Key for Candidate │               │
   │     Document      │               ▼
   └──────────────────┘    ┌──────────────────────┐ 425
             │             │ Add new entry to Retrieved│
             ▼             │  Document  Database   │
   ┌──────────────────┐ 415└──────────────────────┘
   │ Add entry to Log File │           │
   │ with Candidate        │           ▼
   │ Document's Key        │ ┌──────────────────────┐ 430
   └──────────────────┘    │ Add new entry to Document│
             │             │ Comparator Index Database│
             │             └──────────────────────┘
             │                         │
             │                         ▼
             │             ┌──────────────────────┐ 435
             │             │ Add entry to Log File │
             │             │ with retrieved        │
             │             │ Document's Key        │
             │             └──────────────────────┘
             │                         │
             ▼◄────────────────────────┘
          ╭─────────╮ 490
          │   End   │
          ╰─────────╯
```
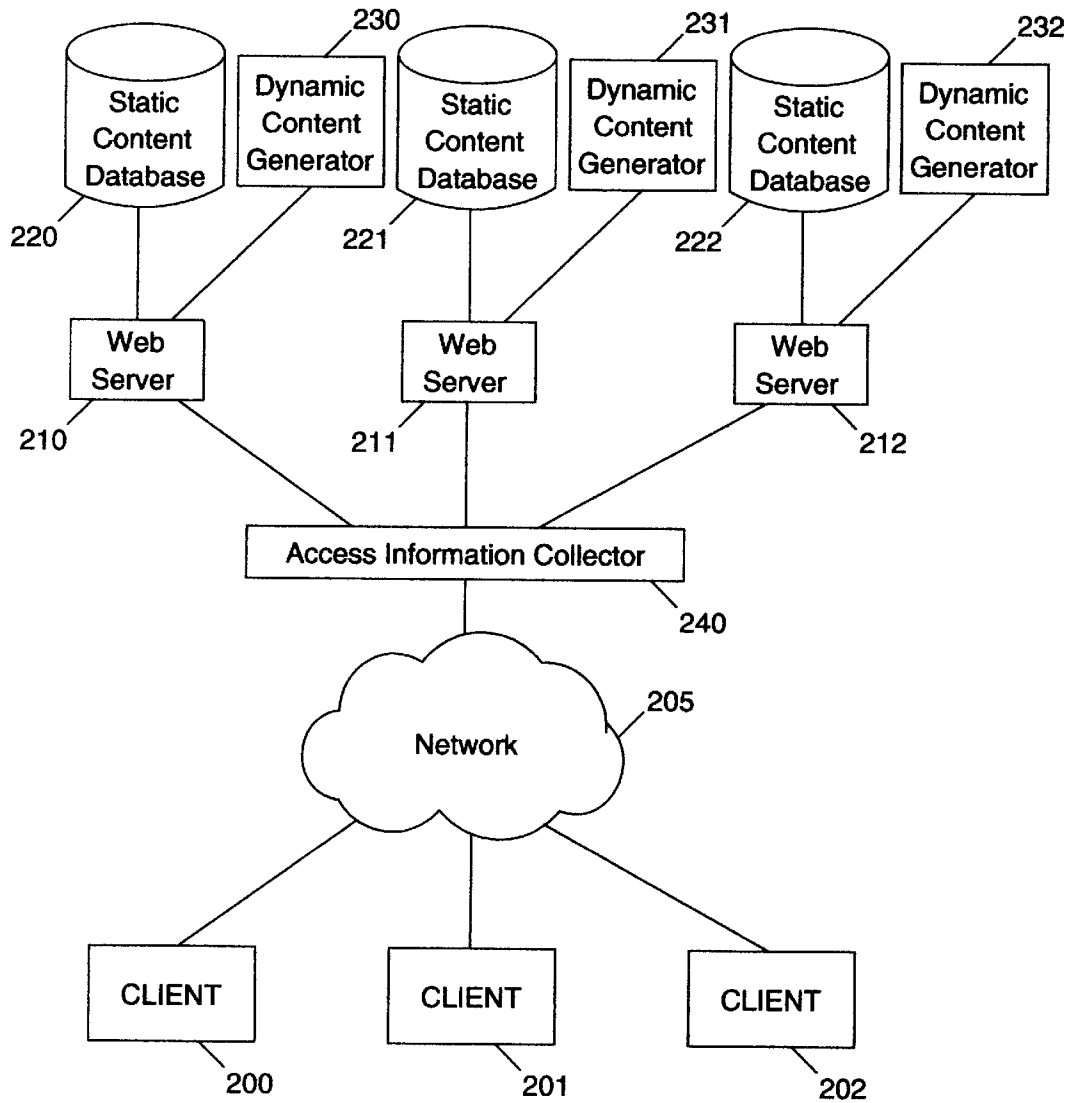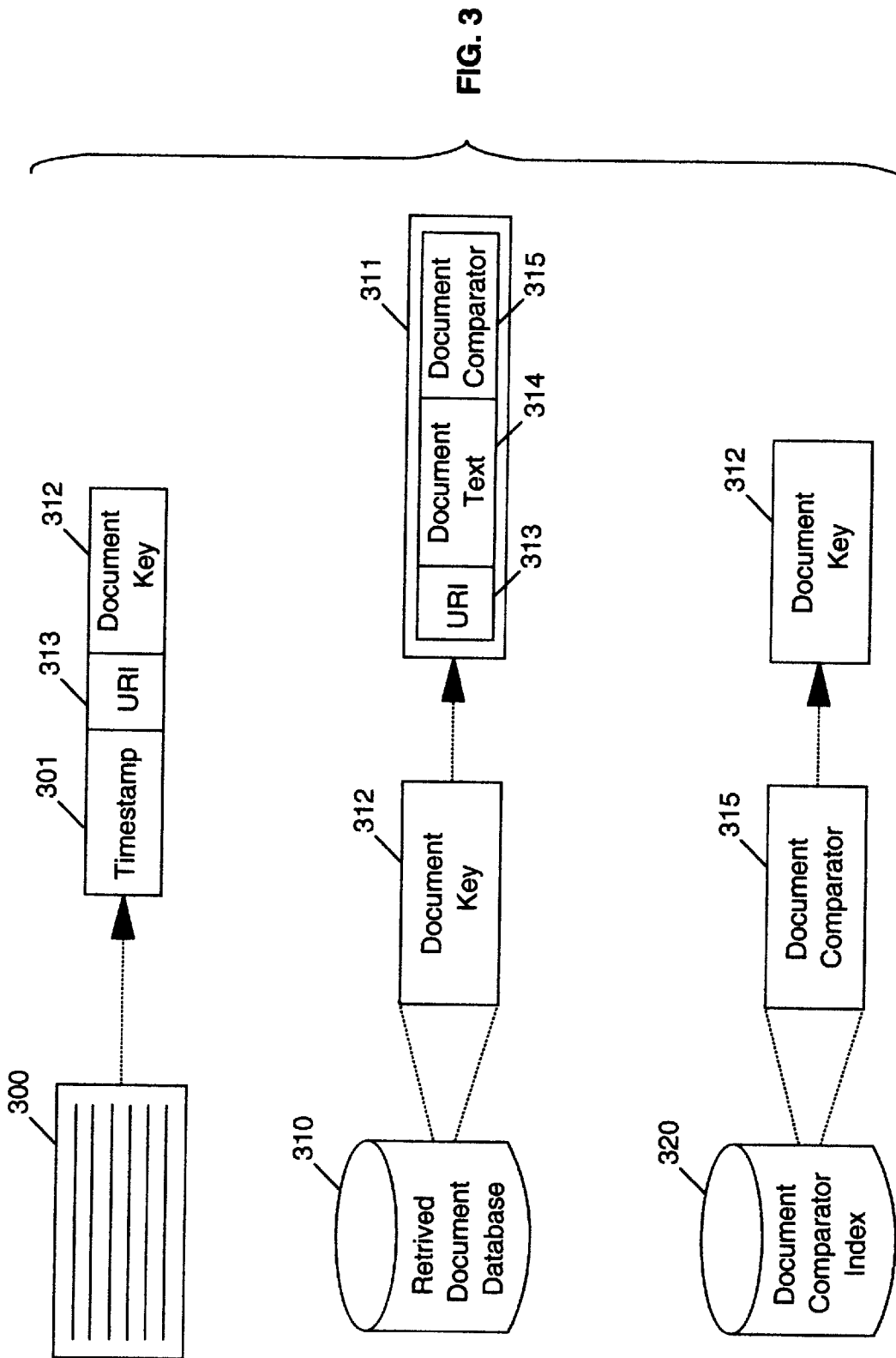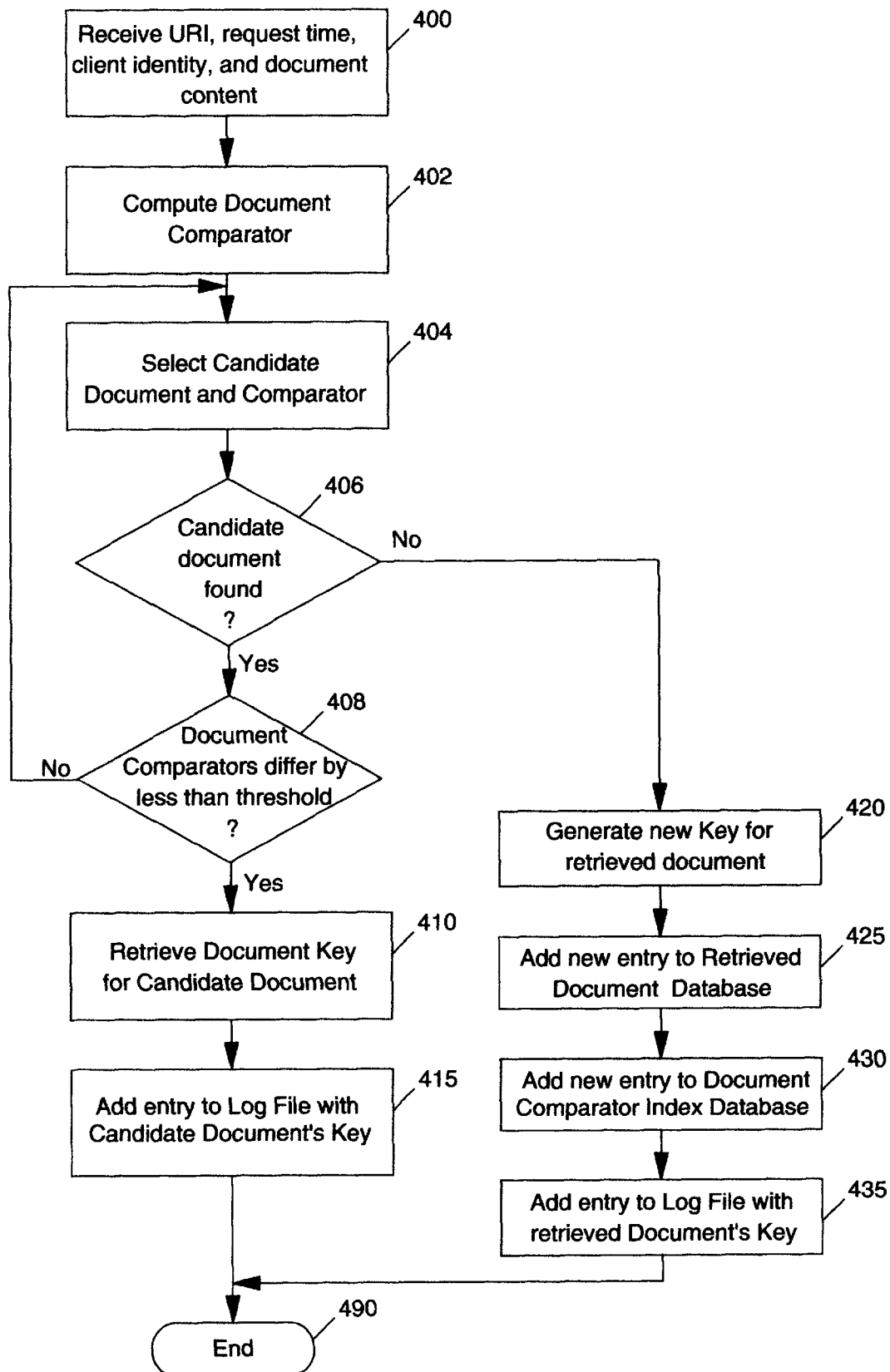
FIG. 4

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

fastcase®
Smarter legal research.