(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 December 2012 (20.12.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/173672 A1

(51) International Patent Classification:
G06F 17/30 (2006.01)

(21) International Application Number:
PCT/US2012/026544

(22) International Filing Date:
24 February 2012 (24.02.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/496,463 13 June 2011 (13.06.2011) US

(71) Applicant *(for all designated States except US)*: **UNITED VIDEO PROPERTIES, INC.** [US/US]; 2830 de La Cruz Boulevard, Santa Clara, California 95050 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **GREEN, Benjamin** [US/US]; 5 Linnaean Street, Apt. 35, Cambridge, Massachusetts 02138 (US). **HELSINGER, Alex** [US/US]; 60 Ossipee Road, Unit 2, Somerville, Massachusetts 02144 (US). **PAPISH, Michael** [US/US]; 1271 North Randolph Road, Randolph Center, Vermont 05061 (US).

(74) Agents: **FEUSTEL, Richard M.** et al.; Ropes & Gray LLP, Prudential Tower, 800 Boylston Street, Boston, Massachusetts 02199-3600 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

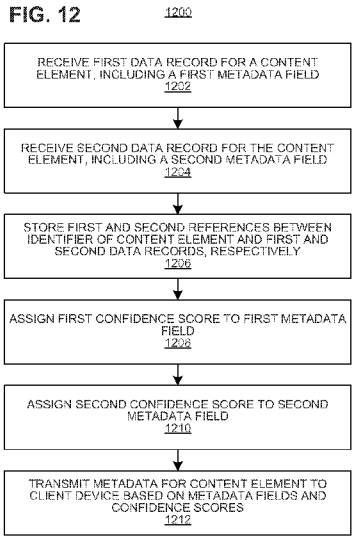
(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

(54) Title: SYSTEMS AND METHODS FOR TRANSMITTING CONTENT METADATA FROM MULTIPLE DATA RECORDS

FIG. 12



(57) Abstract: Systems and methods are provided for cataloging content metadata from a variety of sources and providing metadata to client devices. A processing device receives inconsistent data records representative of a common content element, with different values for a metadata field descriptive of a common attribute of the content element. The processor assign confidence scores metadata fields from each data record, and use these confidence scores to select the metadata that is transmitted to the client device.

WO 2012/173672 A1

SYSTEMS AND METHODS FOR TRANSMITTING CONTENT METADATA FROM MULTIPLE DATA RECORDS

Cross-Reference to Related Application

[0001] This application claims priority to U.S. Provisional Application No. 61/496,463, filed June 13, 2011 and incorporated by reference herein in its entirety.

Background

[0002] Today, there are many separate catalogs and data sources for information about media content such as videos, books, music, and games. The quality of these catalogs varies dramatically, with different catalogs differing in their data structures, naming conventions, content domains covered, level of metadata detail, and accuracy. Previous efforts to aggregate multiple catalogs have relied on re-cataloging, standardizing and reconciling these disparate formats and data, often via automated pattern matching or manual effort by human editors. The challenge of aggregating these catalogs is compounded by the fact that many of these catalogs undergo constant updates with new media content and new metadata.

Summary

[0003] Described herein are systems and methods for cataloging content metadata from a variety of sources so that useful metadata can be quickly and accurately transmitted to individual users or user applications. These systems and methods may be used to import, store, manage and export content identifiers and metadata without loss of editorial control or decline in quality.

[0004] In some aspects, systems and methods are disclosed for resolving the differences between inconsistent data records (*e.g.*, from multiple different data catalogs) in order to provide metadata descriptive of media content to one or more client devices. A processing device of a metadata cataloging system stores, in a memory, first and second data records that are representative of a common content element (*e.g.*, a particular actor, a particular book, a particular genre of music, *etc.*). In some implementations, the processing device receives the first and second data records from two different cataloging systems. However, the data records are inconsistent; though both purport to represent the same content element and have metadata fields descriptive of the same attribute of the content element, the records include different values for those fields (*e.g.*, two different years published, two different running times, two different plot summaries plot summary, *etc.*). In some implementations, the name of the metadata field in the first data record may be the same as the name of the metadata field in the second data record, or may be substantially the same, or similar. The processing device also stores a first confidence score for the metadata field of the first data record, and a second confidence score for the metadata field of the second data record. The processing device identifies which of the metadata fields has the greater confidence score, and when the processing

device receives a request from a client device for metadata descriptive of the common content element, the processing device transmits the value of the identified metadata field to the client device for display. The value may be transmitted to the client device over a computer network, such as the Internet.

[0005] In some implementations, the processing device retains the value of the metadata field with the lesser confidence score in the memory device after selecting the value of the metadata field with the greater confidence score, and may change the confidence scores of metadata fields in response to an instruction from an administrator input device. In some implementations, the processing device receives a lock instruction directed to the first data record from an administrator input device, and in response, stores a locked status for the first data record, indicating that data included in the first data record will not be changed by subsequent change instructions.

[0006] In some implementations, the request from the client device is transmitted in response to receiving, at the client device, a request for a recommendation from a user device. The client device may be configured to transmit the value of the metadata field to a user device for display as a recommendation for the content element, for example.

Brief Description of the Drawings

[0007] The above and other objects and advantages of the systems and methods of the present disclosure will be apparent upon consideration of the following detailed description, taken in conjunction with the accompanying drawings, in which like reference characters refer to like parts throughout, and in which:

[0008] FIGS. 1A and 1B depict illustrative displays that may be used to provide interactive application items;

[0009] FIG. 2 depicts an illustrative recommendation display;

[0010] FIG. 3A is a block diagram of an illustrative interactive media system which may be used with various embodiments;

[0011] FIG. 3B depicts an illustrative client device;

[0012] FIG. 4 is a block diagram of a catalog and recommendation system;

[0013] FIG. 5 is a block diagram of a service processor;

[0014] FIG. 6 depicts an example content data structure;

[0015] FIG. 7 is a flow diagram of a process for including new records and creating new global identifiers in an media data catalog;

[0016] FIG. 8 is a flow diagram of a process for merging records associated with two global identifiers;

[0017] FIGS. 9A-9D illustrate an example of the inclusion and merging processes of FIGS. 7 and 8, respectively;

[0018] FIG. 10 is a flow diagram of a process for splitting a record from a global identifier;

[0019] FIGS. 11A-11B illustrate an example of the splitting process of FIG. 10;

[0020] FIG. 12 is a flow diagram of a process for transmitting content metadata;

[0021] FIG. 13 is a flow diagram of a process for transmitting attribute information using confidence scores;

[0022] FIG. 14A depicts an example content data structure with multiple global identifiers;

[0023] FIG. 14B depicts an example aggregate set of metadata associated with one of the global identifiers of the example media data structure FIG. 14A;

[0024] FIG. 15 depicts an example of a content data structure including editorial metadata, locked global identifiers and locked references; and

[0025] FIG. 16 is an illustrative administrator-facing display for receiving locking instructions.

Detailed Description of Embodiments

[0026] As discussed above, aggregating catalogs of media data presents a number of challenges. As an illustrative example, consider two different content catalogs (*e.g.*, one associated with an online streaming video service and the other curated by an online user community), each of which have a record for the movie "Wizard of Oz." Each catalog will likely have a different identification number or code for this record, along with different metadata fields that describe the movie (*e.g.*, "Title," "Year Produced," "Cast," "Director," "Run Time," *etc.*). Some of these metadata fields may be used in both catalogs and referred to by the same field name (*e.g.*, "Title"), some metadata fields may be used in both catalogs and referred to by different field names (*e.g.*, "Year" and "Year Produced"), and some metadata fields may be used in only a single catalog. Moreover, the quality of the data may differ between the catalogs, and even between metadata fields within one catalog (*e.g.*, the video service catalog has more reliable run time information, while the user-curated catalog has more complete cast information).

[0027] Described herein are systems and methods for making sense of this data in a way that improves the usability of metadata and maintains flexibility in the structure of the catalog. Different implementations of the content cataloging systems and methods provide one or more of the following advantages: allowing multiple records to be associated with a single content identifier that represents a content element (*e.g.*, a movie, an author, a genre), allowing data from each record to be selectively provided to client applications interested in the content element, and maintaining separate identities for the different records to facilitate transfer of records between content identifiers when errors or better matches are discovered. The term "content element" is used herein to refer to any asset, category, feature, property or other characteristic of content that is catalogued by catalog and recommendation system 400 according to the methods described herein. Examples of content elements include particular assets (*e.g.*, the Beatles' "White Album"), or descriptors such as categories (*e.g.*, detective novels, role-playing games), attributes (*e.g.*, actors, directors, language), or any other piece of information that may be catalogued or used to classify content.

[0028] The content cataloging systems and methods disclosed herein may be readily applied to any interactive application (*e.g.*, interactive software, interactive websites, interactive television programs, and interactive presentations) or static application that includes aggregating data for transmitting recommendations to one or more users (*e.g.*, a magazine feature providing product recommendations to different types of readers). As used herein, the term "recommendation" should be understood to mean information chosen to appeal to a user or group of users. Recommendations may be explicit (*e.g.*, by presenting a particular book in a "Recommended For You" display on a website) or implicit (*e.g.*, by presenting an advertisement for a particular product expected to appeal to a particular user or group of users). For illustrative purposes, this disclosure will often discuss exemplary embodiments of these systems and methods as applied in media guidance applications, but it will be understood that these illustrative examples do not limit the range of applications which may be improved by the use of the systems and methods disclosed herein.

[0029] The amount of information available to users in any given search, recommendation or content delivery system can be substantial. Consequently, many users desire a form of media guidance through an interface that allows users to efficiently navigate content selections and easily identify content that they may desire. An application that provides such guidance is referred to herein as an interactive media guidance application or, sometimes, a media guidance application or a guidance application. In particular, the cataloging techniques disclosed herein may be advantageously utilized by guidance applications (*e.g.*, as part of the guidance data source from which the guidance application draws information).

[0030] Interactive media guidance applications may take various forms depending on the content for which they provide guidance. One typical type of media guidance application is an interactive television program guide. Interactive television program guides (sometimes referred to as electronic program guides) are well-known guidance applications that, among other things, allow users to navigate among and locate many types of content. As referred to herein, the term "content" should be understood to mean an electronically consumable user asset, such as television programming, as well as pay-per-view programs, on-demand programs (as in video-on-demand (VOD) systems), Internet content (*e.g.*, streaming content, downloadable content, webcasts, *etc.*), video clips, audio, content information, pictures, rotating images, documents, playlists, websites, articles, books, electronic books, blogs, advertisements, chat sessions, social media, applications, games, descriptions of media assets (*e.g.*, year made, genre, ratings, reviews, *etc.*) and/or any other media or multimedia and/or combination of the same. Guidance applications also allow users to navigate among and locate content. As referred to herein, the term "multimedia" should be understood to mean content that utilizes at least two different content forms described above, for example, text, audio, images, video, or interactivity content forms. Content may be recorded, played, displayed or accessed by client devices, but can also be part of a live performance.

[0031] One of the functions of the media guidance application is to provide media guidance data to users. As referred to herein, the phrase, "media guidance data" or "guidance data" should be understood to mean any data related to content, such as metadata, recommendations, media listings, media-related information (*e.g.*, broadcast times, broadcast channels, titles, descriptions, ratings information (*e.g.*, parental control ratings, critic's ratings, *etc.*), genre or category information, actor information, logo data for broadcasters' or providers' logos, *etc.*), media format (*e.g.*, standard definition, high definition, 3D, *etc.*), advertisement information (*e.g.*, text, images, media clips, *etc.*), on-demand information, blogs, websites, and any other type of guidance data that is helpful for a user to navigate among and locate desired content selections.

[0032] FIGS. 1 and 2 show illustrative display screens that may be used to provide media guidance data organized according to the cataloging systems and techniques disclosed herein. The display screens shown in FIGS. 1-2 may be implemented on any suitable client device or platform. While the displays of FIGS. 1-2 are illustrated as full screen displays, they may also be fully or partially overlaid over content being displayed. A user may indicate a desire to access content information by selecting a selectable option provided in a display screen (*e.g.*, a menu option, a listings option, an icon, a hyperlink, *etc.*) or pressing a dedicated button (*e.g.*, a GUIDE button) on a remote control or other user input interface or device. In response to the user's indication, the media guidance application may provide a display screen with media guidance data organized in one of several ways, such as by time and channel in a grid, by time, by channel, by source, by content type, by category (*e.g.*, movies, sports, news, children, or other categories of programming), or other predefined, user-defined, or other organization criteria. The organization of the media guidance data is determined by guidance application data. As referred to herein, the phrase, "guidance application data" should be understood to mean data used in operating the guidance application, such as program information, guidance application settings, user preferences, or user profile information. In some implementations, the guidance application data is based on data from a catalog of data assembled and maintained in accordance with the techniques described herein. For example, information about the particular channels displayed in FIG. 1 may be those channels for which sufficient metadata is available in a catalog to which the guidance application has access.

[0033] FIG. 1A shows illustrative grid program listings display 100 arranged by time and channel that also enables access to different types of content in a single display. Display 100 may include grid 102 with: (1) a column of channel/content type identifiers 104, where each channel/content type identifier (which is a cell in the column) identifies a different channel or content type available; and (2) a row of time identifiers 106, where each time identifier (which is a cell in the row) identifies a time block of programming. Grid 102 also includes cells of program listings, such as program listing 108, where each listing provides the title of the program provided on the listing's associated channel and time. With a user input device, a user can select program listings by moving highlight region 110. Information relating to the program listing selected by highlight region 110 may be

provided in program information region 112. Region 112 may include, for example, the program title, the program description, the time the program is provided (if applicable), the channel the program is on (if applicable), the program's rating, and other desired information. Program and channel information used in grid 102 may come from a content catalog assembled and maintained according to the techniques described herein.

[0034] Display 100 may also include advertisement 124, video region 122, and options region 126. The item advertised in advertisement 124 and/or the format of advertisement 124 (*e.g.*, interactive or passive, animated or static) may be selected using the recommendation techniques described herein. Video region 122 may allow the user to view and/or preview programs that are currently available, will be available, or were available to the user. The content of video region 122 may correspond to, or be independent from, one of the listings displayed in grid 102. Grid displays including a video region are sometimes referred to as picture-in-guide (PIG) displays. PIG displays and their functionalities are described in greater detail in Satterfield et al. U.S. Patent No. 6,564,378, issued May 13, 2003 and Yuen et al. U.S. Patent No. 6,239,794, issued May 29, 2001, which are hereby incorporated by reference herein in their entireties. PIG displays may be included in other media guidance application display screens of the embodiments described herein.

[0035] Options region 126 may allow the user to access different types of content, media guidance application displays, and/or media guidance application features. Options region 126 may be part of display 100 (and other display screens described herein), or may be invoked by a user by selecting an on-screen option or pressing a dedicated or assignable button on a user input device. The selectable options within options region 126 may concern features related to program listings in grid 102 or may include options available from a main menu display. Features related to program listings may include searching for other air times or ways of receiving a program, requesting programs similar to or recommended based on a program, recording a program, enabling series recording of a program, setting a program and/or a channel as a favorite, providing negative or positive feedback regarding the accuracy of program or channel information for use in adjusting confidence scores associated with metadata, as described in detail below), purchasing a program, or other features. Options available from a main menu display may include search options, VOD options, parental control options, Internet options, cloud-based options, device synchronization options, second screen device options, options to access various types of media guidance data displays, options to subscribe to a premium service, options to edit a user's profile, options to access a browse overlay, or other options.

[0036] Another display arrangement for providing media guidance is shown in FIG. 1B. Video mosaic display 200 includes selectable options 202 for content information organized based on content type, genre, and/or other organization criteria. In display 200, television listings option 204 is selected, thus providing listings 206, 208, 210, and 212 as broadcast program listings. The information in one or more of listings 206, 208, 210 and 212 may include information from an aggregated content catalog, such as those disclosed herein. In display 200 the listings may provide

graphical images including cover art, still images from the content, video clip previews, live video from the content, or other types of content that indicate to a user the content being described by the media guidance data in the listing. Each of the graphical listings may also be accompanied by text to provide further information about the content associated with the listing. For example, listing 208 may include more than one portion, including media portion 214 and text portion 216. Media portion 214 and/or text portion 216 may be selectable to view content in full-screen or to view information related to the content displayed in media portion 214 (*e.g.*, to view listings for the channel on which the video is displayed). A user may also select recommendations option 218 to be provided with recommendations, as discussed below.

[0037] FIG. 2 is an illustrative display for providing recommendations which may be generated, for example, in response to a user selection of recommendation option 218, or in response to any other suitable user action (*e.g.*, logging in to a search service, or launching a media guidance application), and may be based on the data stored in an aggregated catalog assembled and maintained using the techniques described herein. Display 250 includes a set of navigation elements 260, each of which may be selected by a user to change the information displayed (*e.g.*, personal recommendations). In the current display, navigation element 262 is highlighted, indicating that "For You" information is displayed. In display 250, the "For You" information includes an array of content element indicators 252, each of which indicates a particular content element that is recommended for the user in an associated content domain 254 (*e.g.*, "movies," "music," "TV," *etc.*). Each content element indicator may indicate an asset (*e.g.*, the movie "Top Gun"), a genre (*e.g.*, the musical genre of "Death Metal"), an artist (*e.g.*, the author J.K. Rowling) or any other content element that is expected to appeal to the user. A user may select "more" icon 256 to view more recommendations on a display in a particular content domain. In some embodiments, recommendations are not displayed by content domain, but are displayed according to chronology, in order of user preference, clustered by common elements (*e.g.*, common actions, common themes or common rating, or are arranged randomly. Each of the indicators 252 may be user-selectable (*e.g.*, via mouse click, double-touch, or hover-over), the recommendation systems described herein may provide additional information about the content element and/or allow the user to access assets associated with the content element. Display 250 includes advertisement 258, which may advertise a product, service or other purchasable item. As described above with reference to advertisement 124 of FIG. 1A, the item advertised by advertisement 258 may be selected based on a user's preference or by a determination that the advertised item is related to content elements that the media guidance application has determined that a client may like. Further discussion of various configurations for the display screens of FIG. 1-2, as well as several other exemplary displays, are presented elsewhere herein.

[0038] FIG. 3A is a block diagram of an illustrative interactive media system 350. System 350 includes media content source 366 and media guidance data source 368 coupled to communications network 364 via communication paths 370 and 372, respectively. Paths 370 and 372 may separately

or together include one or more communications paths, such as, a satellite path, a fiber-optic path, a cable path, a path that supports Internet communications (*e.g.*, IPTV), free-space connections (*e.g.*, for broadcast or other wireless signals), or any other suitable wired or wireless communications path or combination of such paths. Communications with media content source 366 and media guidance data source 368 may be exchanged over one or more communications paths, but are shown as a single path in FIG. 3A to avoid overcomplicating the drawing. In addition, there may be more than one of each of media content source 366 and media guidance data source 368, but only one of each is shown in FIG. 3A to avoid overcomplicating the drawing. Different possible types of each of these sources are discussed below. If desired, media content source 366 and media guidance data source 368 may be integrated as one source device. Media content source 366 and media guidance data source 358 include inputs 384 and 386, respectively, for receiving data from external sources. The cataloging systems and techniques disclosed herein may be implemented by media guidance data source 358, for example, which may be configured to aggregate content metadata received from multiple metadata sources via input 386.

[0039] In some implementations, a media guidance application is implemented on a client server, which receives data from a media guidance data source (such as media guidance data source 368) and uses that data to provide a media guidance application to one or more client devices. In some implementations, the media guidance application executes directly on the client device; in this case, the client device is itself a client of the media guidance data source. As used herein, the term "client" or "client device" should be understood to mean any device that receives media guidance data (such as recommendations) from a media guidance data source. A user device, then, is a particular example of a client device. Client devices 374 may be coupled to communications network 364. Namely, user television equipment 352, user computer equipment 354, and wireless user communications device 356 are coupled to communications network 364 via communications paths 358, 360, and 362, respectively. Client devices 374 may include client data server 376, which has additional client devices: user television equipment 378, user computer equipment 380, and wireless user communications device 382. Communications network 364 may be one or more networks including the Internet, a mobile phone network, mobile voice or data network (*e.g.*, a 4G or LTE network), cable network, public switched telephone network, or other types of communications network or combinations of communications networks. Paths 358, 360, 362, 388, 390, 392 and 394 may include any of the communication paths described above in connection with paths 370 and 372. Paths 362 and 394 are drawn with dotted lines to indicate that, in the exemplary embodiment shown in FIG. 3A, they are wireless paths, and paths 358, 360, 388, 390 and 392 are drawn as solid lines to indicate they are wired paths (although these paths may be wireless paths, if desired). Various network configurations of devices may be implemented and are discussed in more detail below. Although communications between sources 366 and 368 and client devices 374 are shown as through communications network 364, in an embodiment, sources 366 and 368 may communicate directly

with client devices 374 via communication paths (not shown) such as those described above in connection with paths 370 and 372. Additional discussion of suitable configurations of system 350 is presented elsewhere herein.

[0040] Client devices 374 of FIG. 3A can be implemented in system 350 as any type of equipment suitable for accessing content and/or media guidance data, such as a non-portable gaming machine. Client devices, on which a media guidance application may be implemented, may function as standalone devices or may be part of a network of devices. FIG. 3B shows a generalized embodiment of illustrative client device 300. More specific implementations of client devices are discussed below in connection with FIG. 3A. Client device 300 may receive content and data (such as metadata from a catalog) via input/output (hereinafter "I/O") path 302. I/O path 302 may provide content (*e.g.*, broadcast programming, on-demand programming, Internet content, content available over a local area network (LAN) or wide area network (WAN), and/or other content) and data (such as media guidance data) to control circuitry 304, which includes processing circuitry 306 and storage 308. Control circuitry 304 may be used to send and receive commands, requests, and other suitable data using I/O path 302. I/O path 302 may connect control circuitry 304 (and specifically processing circuitry 306) to one or more communications paths (described below). I/O functions may be provided by one or more of these communications paths, but are shown as a single path in FIG. 3B to avoid overcomplicating the drawing. In some implementations, client device 300 is a user device through which a user may access content and the media guidance application (and its display screens described above and below). In some implementations, client device 300 is a server or other processing system that acts as an intermediary between media guidance data (such as content metadata) and one or more user devices.

[0041] An operator may send instructions to control circuitry 304 using input interface 310. Input interface 310 may be any suitable interface, such as a remote control, mouse, trackball, keypad, keyboard, touch screen, touchpad, stylus input, joystick, voice recognition interface, or other input interfaces. Display 312 may be provided as a stand-alone device or integrated with other elements of client device 300. Display 312 may be one or more of a monitor, a television, a liquid crystal display (LCD) for a mobile device, or any other suitable equipment for displaying visual images. In some embodiments, display 312 may be HDTV-capable. In some embodiments, display 312 may be a 3D display, and the interactive media guidance application and any suitable content may be displayed in 3D. A video card or graphics card may generate the output to the display 312. The video card may offer various functions such as accelerated rendering of 3D scenes and 2D graphics, MPEG-2/MPEG-4 decoding, TV output, or the ability to connect multiple monitors. The video card may be any processing circuitry described above in relation to control circuitry 304. The video card may be integrated with the control circuitry 304. Speakers 314 may be provided as integrated with other elements of client device 300 or may be stand-alone units. The audio component of videos and other content displayed on display 312 may be played through speakers 314. In some embodiments, the

audio may be distributed to a receiver (not shown), which processes and outputs the audio via speakers 314. In some implementations, client device 300 may not include one or more of display 312 and speakers 314.

[0042] FIG. 4 is a block diagram of catalog and recommendation system 400, one embodiment of media guidance data source 368 of FIG. 3A. In some implementations, the components of catalog and catalog and recommendation system 400 are distributed between multiple processing and storage devices; for example, the components of catalog and recommendation system 400 may be divided between media guidance data source 368, media content source 366 and client data server 376 (FIG. 3A). Catalog and recommendation system 400 is illustrated as divided into three functional components, each of which include one or more processing devices and storage devices (such as these described above with reference to client device 300 of FIG. 3B): orchestration component 406, offline component 402 and real-time component 404. Offline component 402 may be configured to perform many of the back-end cataloging processes described herein. In particular, offline component 402 includes content information database 414, which may receive media data records from one or more data sources via input 438 (which may correspond to input 386 of media guidance data source 368 of FIG. 3A). Content information database 414 includes memory hardware configured to operate in any of a number of database architectures, such as a relational database management system or a document-based database architecture like NoSQL. Content information database 414 also includes a processing engine executed on one or more database servers to receive, store and serve data stored in memory. Any of the database hardware and architecture configurations described herein, including those described above with reference to content information database 414, may be used for any of the databases or data storage systems described herein. In some embodiments, the media data records received at input 438 are electronic signals representative of media content or information about media content (referred to herein as "content metadata" or "metadata"). Signals received at input 438 may be provided by third-party data providers (such as cable television head-ends, web-based data sources, catalog management organizations, or real-time or other data feeds) or from users supplying content or metadata to catalog and recommendation system 400. Signals received at input 438 may take the form of a file of multiple data records, or through a message bus that provides new data records and updates to previous data records as changes are made, for example. In some implementations, content information database 414 is coupled with one or more processing devices configured to extract metadata from data records arranged in a tabular format and to store that metadata in content information database 414. In some implementations, content information database 414 may "catalog" the information received at input 438 in a memory (*e.g.*, local, remote or distributed) according to a data structure, as described in additional detail below.

[0043] Information from database 414 may be transmitted (by one or more servers associated with database 414) to data mining processor 412. Data mining processor 412 is configured to extract information from database 414 and process the extracted information to reconcile information from

multiple sources (*e.g.*, data records from multiple catalog management systems). In some implementations, data mining processor 412 includes a memory device configured as a database for storing one or more tokens used in performing the domain-based tokenization techniques described in co-pending application _____, entitled "Systems and methods for domain-specific tokenization" (Attorney Docket No. 003597-0618-101), which is incorporated by reference herein in its entirety. Data mining processor 412 may also transmit the reconciled information to core content relations management ("CCRM") module 408. As used herein, the term "module" should be understood to mean a processing device executing programming logic, such as source code, or higher-level code (*e.g.*, Java code executed via a Java compiler), stored in a memory device (*e.g.*, RAM, ROM, removable memory media, Flash memory, optical dishes, *etc.*). In some implementations, CCRM module 408 includes a MySQL database of reconciled data. Systems and methods for reconciling data in an aggregate catalog, which may be implemented by data mining processor 412 in conjunction with CCRM module 408 and the rest of offline component 402, are described in detail below.

[0044] CCRM module 408 may also receive information from editorial influence module 410. In some embodiments, editorial influence module 410 receives metadata from human or computer editors, and augments the information that is automatically catalogued with this "editorial" metadata. Editorial influence module 410 includes a server configured to provide a web-based interface between human editors and the database of CCRM module 408. In some implementations, editorial influence module 410 includes a Java application running on an Apache Tomcat web server, but may be executed on any processing device or devices with a user interface. Human editors may interact with the web-based interface using a personal computer connected to the Internet, a hand-held device, or any of the client devices (such as client device 300 of FIG. 3B) described herein. Editorial metadata is described in additional detail below with reference to FIG. 14.

[0045] Information from database 414 may also be transmitted (*e.g.*, by one or more servers associated with database 414) to export/index processor 416. Export/index processor 416 queries CCRM module 408 to extract catalog information from CCRM module 408 and formats this information for use in different modules of real-time component 404 (as described in detail below). Export/index processor 416 may be configured to extract information in batches on a regular interval (*e.g.*, every twenty-four hours) and format and transmit this batched information to a dependent module, or may be configured to extract information as it is updated in CCRM module 408. As shown in FIG. 4, export/index processor 416 transmits information to domain relations module 420, search indices module 424, and metadata module 422. These modules serve as "quick" sources of certain common types of information for real-time service processor 418 (described in detail below); instead of requiring real-time service processor 418 to query CCRM module 408 whenever a particular kind of metadata is desired, real-time service processor 418 may instead query one of these modules to obtain the information. Domain relations module 420 includes a data storage device configured as a database for storing metadata about relationships between media content and

descriptors of media content (such as genre, actors, media domain, rating, *etc.*). Content metadata module 422 includes a data storage device configured as a database for storing frequently requested metadata. Additionally, metadata module 422 may include only those metadata fields that are commonly used for the recommendation techniques executed by real-time service processor 418. In some implementations, metadata module 422 stores a subset of the data stored in CCRM module 408 in a format that can be easily filtered according to the parameters of a search or recommendation request (*e.g.*, a tabular format that can be quickly filtered to exclude movies rated "R" and above). Search indices module 424 includes a data storage device configured as a database for storing search heuristics that may be used by real-time service processor 418 to improve search performance. Many search techniques utilize heuristics such as removing spaces from search queries, transforming queries into lower-case characters, comparing a search query against a list of common variations and misspellings, and identifying one or more n-grams within a search query, among others.

[0046] Real-time service processor 418 receives information from domain relations module 420, metadata module 422 and search indices module 424, as described above, and provides recommendation information to client devices (such as client device 300). The components of catalog and recommendation system 400 may be distributed between multiple processing and storage devices; for example, the components of catalog and recommendation system 400 may be divided between media guidance data source 368, media content source 366 and client data service 376 (FIG. 3B) via device gateway 434 in orchestration component 406. Device gateway 434 may include any transmission path suitable for communicating recommendation information, such as the path 372 between media guidance data source 368 and communication network 364 and paths 358, 360, 362 and 388 between communication network 364 and client devices 374 (FIG. 3A). In particular, real-time service processor 418 is configured to provide metadata in response to various types of client queries (*e.g.*, for metadata matching a search term, for metadata on content related to a particular content element, *etc.*). Real-time service processor 418 may provide, for example, identifiers of particular content as well as metadata for that content (*e.g.*, album art in response to a music search request). Real-time service processor 418 may execute any of a number of recommendation techniques, such as those described in co-pending application _____, titled "Systems and methods for providing media recommendations" (Attorney Docket No. 003597-0603-101), which is incorporated by reference in its entirety herein. Real-time service processor 418 may also query CCRM module 408 directly, or provide feedback to CCRM module 408 as an application on a client device interacts with catalog and recommendation system 400 through device gateway 434. In some implementations, real-time service processor 418 is implemented as a web service executing on an Apache Tomcat or other server.

[0047] Real-time service processor 418 also communicates with profiles database 426, which may include a data storage device configured as a database for storing information about client preferences (including preference values calculated by real-time service processor 418 as described herein with

reference to FIG. 6), client equipment, client event history, or other information relevant for transmitting recommendations and data to a client. In some implementations, profiles database 426 stores profile information individual users (who may be users of an intermediate client service). Media guidance applications (such as recommendation applications) may be personalized based on a client's preferences as stored in profiles database 426. A personalized media guidance application allows a client to customize displays and features to create a personalized "experience" with the media guidance application. The customizations may include preferred sources of content metadata (which may be assigned higher confidence scores in the cataloging techniques described in detail below), varying presentation schemes (*e.g.*, color scheme of displays, font size of text, *etc.*), aspects of content listings displayed (*e.g.*, only HDTV or only 3D programming, client-specified broadcast channels based on favorite channel selections, re-ordering the display of channels, recommended content, *etc.*), desired recording features (*e.g.*, recording or series recordings for particular users, recording quality, *etc.*), parental control settings, customized presentation of Internet content (*e.g.*, presentation of social media content, e-mail, electronically delivered articles, *etc.*) and other desired customizations. This personalized experience may be created by allowing a client (such as a user) to input these customizations and/or by the media guidance application monitoring user activity to determine various user preferences.

[0048] Clients may access their personalized guidance application by logging in, communicating with catalog and recommendation system 400 using a designated protocol over path 440, or otherwise identifying themselves to the guidance application. The media guidance application may allow a client to provide profile information for profiles database 426 or may automatically compile profile information. The media guidance application may, for example, monitor the content the client accesses and/or other interactions the user may have with the guidance application, including responses to and feedback based on recommended content. Profiles database 426 may communicate with event database 436, which may store event records that contain information about client interactions with catalog and recommendation system 400. Profiles database 426 may access event database 436 to reconstruct a client's history of use of catalog and recommendation system 400 and to determine content preferences. Additionally, the media guidance application may obtain all or part of other profiles that are related to a particular client (*e.g.*, from other web sites on the Internet the client accesses, such as www.allrovi.com, from other media guidance applications the client accesses, from other interactive applications the client accesses, from another device of the client, *etc.*), and/or obtain information about the client from other sources that the media guidance application may access. As a result, a client can be provided with a unified guidance application experience across the client's different devices. This type of experience is described in greater detail below. Additional personalized media guidance application features are described in greater detail in Ellis et al., U.S. Patent Application Publication No. 2005/0251827, filed July 11, 2005; Boyer et al., U.S. Patent No. 7,165,098, issued January 16, 2007; and Ellis et al., U.S. Patent Application Publication

No. 2002/0174430, filed February 21, 2002, which are hereby incorporated by reference herein in their entireties.

[0049] Real-time service 418 transmits information to and receives information from client devices by way of path 440 and device gateway 434. As described above with reference to paths 370 and 372 of FIG. 3A, path 440 may include one or more communication paths such as a satellite path, a fiber-optic path, a cable path, a path that supports Internet communications, free-space connections, or any other suitable wired or wireless communications path or combination of such paths. Device gateway 434 may be, for example, a web service implemented on one or more server devices, configured to receive requests from client devices via path 440. The client devices that communicate with device gateway 434 may be client devices 374 of FIG. 3A. These client devices may take the form of client device 300 (FIG. 3B), for example, and may communicate with device gateway 434 via I/O path 302 (FIG. 3B). The data provided to client devices via path 440 may be supplemented by data from supplemental database 428, which may store metadata and media content that is provided along with the information transmitted from real-time service processor 418 to device gateway 434. Supplemental database 428 may include, for example, media content source 366 (FIG. 3A), and may include or be in communication with media guidance data source 358 or another content metadata catalog. For example, in response to a call to device gateway 434 from a client device, device gateway 434 may send a search or recommendation request to real-time service processor 418. Real-time service processor 418 may respond by sending a list of content identifiers that satisfy the request back to device gateway 434, at which point device gateway 434 will request appropriate supplementary information from supplemental database 428 (*e.g.*, clips of videos whose identifiers are included in the data provided to device gateway 434 by real-time service processor 418). In some implementations, supplemental database 428 is populated with information from content information database 414, domain relations module 420, metadata module 422 or search indices module 422, or may be the same as one or more of these databases or modules.

[0050] FIG. 5 is a block diagram of service processor 500, one possible implementation of real-time service processor 418 (FIG. 4A) and media guidance data source 368 (FIG. 3A). Service processor 500 may be functionally organized into web service tier 532 and orchestration tier 534. Orchestration tier 534 includes dispatcher processor 506, one or more service context modules 508, sources 510 and cache 516. Dispatcher processor 506 manages the flow of data between web service tier 532 and other components in orchestration tier 534, and in particular, responds to requests from web service tier 532 by checking to see whether data stored in cache 516 satisfies the request or determining which of service context modules 508 to call to satisfy the request. Requests from web service tier 532 may represent requests from client devices (such as client devices 374 of FIG. 3A) received via path 440 and device gateway 434 (FIG. 4), for example. In some implementations, dispatcher processor 506 includes processing hardware configured to execute a Java application to perform the operations described herein. Cache 516 includes a memory device that stores data

recently received from or transmitted to the web service tier 532, thus providing a "quick" source for data that may be requested or used multiple times. Each of sources 510 includes computer-executable code (*e.g.*, Java code) for performing a particular search or recommendation operation. For example, a source may include code that may be executed to perform a search of a particular database, or may include code that may be executed to identify similar items to a specified item within a catalog. Sources 510 include primary sources 512, which include basic or common search or recommendation operations, and secondary sources 514, which include custom implementations of particular search or recommendation operations (*e.g.*, for particular clients) or implementations of search or recommendation operations that build on or use primary source operations stored as primary sources 512. Collections of one or more sources 510 are stored as service context modules 508, each of which specifies a particular set of one or more of sources 510 to use when satisfying requests (*e.g.*, from particular regions like North America or Europe, or from particular customers). In some implementations, service context modules 510 are represented as XML files.

[0051] Dispatcher processor 506 is also in communication with a number of service modules in web service tier 532, including REST v1 service module 520, REST v2 service module 522, and SOAP service module 524. These different service modules provide interfaces and transport mechanisms for accessing the "back-end" processing and data of orchestration tier 532. REST and SOAP are two different ways of packaging input and output data, and any other such protocols may be used. In some embodiments, service processor 500 includes processing and networking hardware configured with a software platform for serving dynamically generated recommendations applications in XML and JSON.

[0052] In some embodiments, index/export module 416 (FIG. 4) or service processor 500 (FIG. 5) includes one or more data contracts. Data contracts are electronic data files, encoded in a data definition language, that define a type and structure of data available to an application that accesses a service that operates according to the contract. When an application accesses a contracted service, the application can parse the contract to determine what data (*e.g.*, assets, metadata, recommendations, *etc.*) the service can provide. A single service may be instantiated multiple times with different contracts, with each contract governing a different type of data. A service may advertise the data types and structures that, according to the contract, the service can provide to applications. The application may receive this information and determine which services provide data of a type and structure that is compatible with the application's own purpose and architecture. Multiple services, each with its own contract or contracts, may communicate with each other, passing data through the services and transforming the data, repackaging the data, or adding content along the way. In some embodiments, a service determines the contracts that it advertises based on the contracts that it reads in from other services (indicating the data types and structure to which the service has access), plus additional fields and operations that represent additional functionality provided by the service itself.

[0053] The types and structure of data specified in a contract may take any of a number of forms. For example, a recommendation system may receive a search or other query and may return pointers to media assets and fields containing metadata about those media assets. Thus, in some configurations of service processor 500 (FIG. 5), each of services 520, 522 and 524 advertises the operations the service supports, the types of data that the service can return per operation, and the fields it can return per data type per operation. A field may be single- or multi-valued, optional or required, and stored as strings or more complex data objects (such via a map to an internal object through JSON). The same contract can also be used in different service contexts, which utilize different sets of underlying source data (*e.g.*, different third-party metadata catalogs).

[0054] The above systems may be configured to assemble and maintain catalogs of content metadata, and to transmit that metadata to users and client applications, according to the techniques described below. These systems are particularly advantageous when multiple sources of data records are aggregated into a single, global identification space. For example, these systems may assign a single global identifier to a content element (*e.g.*, a movie) and reference multiple data records for that content element from multiple data sources to the single global identifier, while maintaining the ability to merge identifiers, split identifiers, remove records and extract the most useful data from the records to present to client devices or downstream applications. Although the techniques described below may be described as executed by data mining processor 412 (FIG. 4) for clarity of illustration, it will be understood that the cataloging and metadata provision processes of the present disclosure may be performed by any device or group of devices configured to do so; for example, any special- or general-purpose processing circuitry located within media guidance data source 358 (FIG. 3A), client device 300 (FIG. 3B), or any appropriately-configured component of catalog and recommendation system 400 (FIG. 4) such as processing circuitry associated with CCRM module 408. In some implementations, these processes are performed by multiple processing devices operating in series, in parallel, or a combination.

[0055] In some embodiments of the data cataloging systems and methods described herein, data are stored in content data structures that associate each global content identifier (which identifies a particular content element, such as a movie or author) with one or more data records representative of the content element using three different types of associations: original references, references, and original relations. These reference types are illustrated in FIG. 6, which depicts an example content data structure 600 which may be maintained, for example, in CCRM module 408 of catalog and recommendation system 400 (FIG. 4). A detailed discussion of how such content data structure may be created and modified is given below with references to FIGS. 7-11. Content data structure 600 includes global identifier 602 (representing the movie "Titanic") which is associated with record 604 via original reference 606. Original reference 606 indicates that global identifier 602 was the identifier originally associated with record 604 when record 604 was first stored in CCRM module 408. Global identifier 602 is also associated with record 610 via reference 608. Record 610

is associated with global identifier 612 via original relation 614 and with global identifier 602 via reference 608. Original relation 614 indicates that global identifier 612 was the global identifier associated with record 610 when record 610 was first stored in CCRM module 408 but that record 610 was subsequently associated with a different global identifier (global identifier 602) as indicated by reference 608. Original relation 614 is maintained in a memory of CCRM module 408.

[0056] FIG. 7 is a flow diagram 700 of a process for including new records in a content data structure. At step 702, data mining processor 412 creates a new global identifier. The global identifier may be, for example, an entry in CCRM module 408 associated with a unique serial number or a pointer to a designated database entry or memory location. At step 704, data mining processor 412 associates the new global identifier created at step 702 with a new record to be included in CCRM module 408. This association may occur, for example, by a common entry in a database, a link in memory, or by any other data structure. At step 706, data mining processor 412 sets the new record as the original record associated with the new global identifier created at step 702. Unless and until the new record is associated with a different global identifier, CCRM module 408 will continue to store the original reference between the new record and the new global identifier as discussed above with reference to FIG. 6. If the new record becomes associated with a different global identifier (*i.e.*, through a merge operation as described below), the global identifier will maintain an original relation with the record.

[0057] FIG. 8 is a flow diagram of a process for merging records associated with two global identifiers in CCRM module 408. At step 802, data mining processor 412 selects one of two candidate global identifiers that are to be merged. In some embodiments, data mining processor 412 selects the global identifier with the greater number of references to data records. If the global identifiers are associated with the same number of data records, data mining processor 412 may select the older of the two global identifiers. Though selecting the "larger" global identifier at step 802 allows more rapid consolidation of data records associated with the same content element than selecting the "smaller" global identifier, any selection technique may be used.

[0058] At step 804, data mining processor 412 associates all of the records of the non-selected global identifier with the selected global identifier. This association is accomplished by including a reference from the selected global identifier to each of the records of the non-selected global identifier. At step 806, data mining processor 412 de-associates all records from the non-selected global identifier. If the non-selected global identifier was the original identifier for any of the records, CCRM module 408 continues to store, in memory, the original relation between the non-selected global identifier and those records.

[0059] FIGS. 9A-9D illustrate an example of the inclusion and merging processes of FIGS. 7 and 8, respectively. To begin, FIG. 9A represents the state of a portion of a data catalog in CCRM module 408 after five new records 922-930 are first imported. To catalog the new records, data mining processor 412 creates five new global identifiers 902-910 and associates each of the new

records 922-930 with one of the new global identifiers 902-910 through an original reference (original references 912-920). FIG. 9B depicts the result of transferring a record from its original global identifier to a different global identifier. In FIG. 9B, record 930 is associated with global identifier 908 via reference 932. In lieu of original reference 920, CCRM module 408 maintains original relation 934 between record 930 and global identifier 910 (*e.g.*, by storing in a memory).

[0060] FIG. 9C depicts the result of two additional transfers of records between global identifiers. In FIG. 9C, record 924 is associated with global identifier 902 via reference 936 (while original relation 938 between record 924 and global identifier 904 is maintained), and record 926 has become associated with global identifier 908 via reference 942 (while original relation 940 between record 926 global identifier 906 is maintained). In the embodiment illustrated in FIGS. 9A-9D, each of the transfers of records depicted in FIGS. 9B and 9C can also be thought of as mergers between the associated global identifiers because merger and transfer operations have the same result when both global identifiers are each associated with a single record. FIG. 9D depicts the result of merging two global identifiers, one of which has more than one associated record. Here the records associated with global identifier 902 and the records associated with global identifier 908 have been merged under global identifier 908. In this example, the global identifier with the larger number of records prior to the merger (here, global identifier 908) is the global identifier to which the new records are transferred in the merger, though as discussed above with reference to FIG. 8, any other transfer rule may be used. As a result of this merger, records 922 and 924 are now associated with global identifier 908 via references 946 and 948, respectively, and the original relation between global identifier 902 and record 922 is maintained in memory.

[0061] In some embodiments, once a record is associated with a global identifier via an active reference (*e.g.*, a reference or an original reference), the record may be split from the global identifier and returned to its original global identifier. FIG. 10 is a flow diagram of a process for splitting a record from a global identifier. At step 1002, data mining processor 412 identifies the record to be split from a global identifier with which the record is currently associated. This record may be identified by a human editor as incorrectly associated with the global identifier, or may be identified as such by an automated aggregation process (*e.g.*, that uses pattern matching or other computational techniques). At step 1004, data mining processor 412 determines which global identifier serves as the original global identifier for the record identified at step 1002. In other words, data mining processor 412 determines which global identifier has an original relation with the record to be split. At step 1006, data mining processor 412 re-associates the record identified at step 1002 with its original global identifier (as determined at step 1004). Thus, the original reference between the identified record and its original global identifier is restored as an active reference. At step 1008, data mining processor 412 de-associates the identified record from its former global identifier.

[0062] FIGS. 11A-11B illustrate an example of the splitting process of FIG. 10. In FIG. 11A, global identifier 1108 is associated with records 1122, 1126 and 1130 via references 1132, 1134

and 1136, respectively. Global identifier 1108 is also associated with record 1128 via original reference 1118. Global identifier 1104 is associated with record 1124 via original reference 1114, and global identifiers 1102, 1106 and 1110 have original relations 1112, 1116 and 1120 with records 1122, 1126 and 1130, respectively. FIG. 11B depicts the result of splitting record 1122 from global identifier 1108. In FIG. 11B, the reference between record 1122 and global identifier 1108 is removed, and the original reference between record 1122 and global identifier 1102 is reinstated.

[0063] In some embodiments, a content catalog (such as CCRM module 408 and other databases that implement the content data structures and management techniques described above with reference to FIGS. 6-11) may transmit content data to media guidance applications (such as those described herein with reference to FIGS. 1-3). In some such embodiments, a client application requests information from the catalog by global identifier, content element name, attribute of a particular content element, metadata field name, content element type, or any other characteristic, and the requested data is retrieved and passed to the requesting application (*e.g.*, in accordance with a contract as described above with reference to FIGS. 4 and 5). In some embodiments, the content catalog stores additional confidence information that describes the reliability and/or quality of different data records, metadata fields and/or references. As discussed in detail below, confidence scores may be assigned to many different elements in a content data catalog and may be used in combination to transmit aggregate metadata from multiple (potentially inconsistent) records. The aggregate metadata and confidence scores may be stored in CCRM module 408, and subsets of this data may be transmitted to domain relations module 420 and asset metadata module 422 as discussed above with reference to FIG. 4.

[0064] FIG. 12 is a flow diagram of a process for transmitting content metadata, which may be executed by one or more of the processing devices described herein, such as data mining processor 412 of FIG. 4. At step 1202, data mining processor 412 receives a first data record representative of a content element from content information database 414 (FIG. 4). The first data record may identify a particular content element (*e.g.*, by name; by ISBN or other serial number; by URL; by time, date and channel information, *etc.*) and includes a first metadata field with a value descriptive of an attribute of the content element. For example, when the content element is the movie "Titanic," the name of the first metadata field may be "Title" and it may have a value of "Titanic." Other examples of attributes include title, running time, synopsis, price, format, maturity level, game type, or any characteristics or descriptions of content elements that may be stored as metadata. The first data record may also include additional metadata fields with values descriptive of other attributes of the content element.

[0065] At step 1204, catalog and recommendation system 400 receives, from content information database 414 (FIG. 4), a second data record representative of the same content element as the first data record. The second data record may come into catalog and recommendation system 400 via input 438 from a different data source than the first data record, such as a different website, a different

third party catalog, or a different television guide system. The second data record may identify the content element in the same way that the first data record did, or in a different way, and includes a second metadata field with a value descriptive of the same attribute of the content element as described by the first metadata field of the first data record. The name of the second metadata field may be the same as the name of the first metadata field, or it may be different (*e.g.*, the name of the second metadata field may be "Working Title" or "Content Name"). In this illustration of the process of FIG. 12, the value of the second metadata field is different than the value of the first metadata field (*e.g.*, when the second data record is representative of the movie "Titanic," the value of "Working Title" may be a variant such as "Titanic: The Movie" or a completely incorrect title like "Monsters, Inc.").

[0066] At step 1206, data mining processor 412 stores, with an identifier for the content element (*e.g.*, a global identifier as described above with reference to FIG. 7), a first reference between an identifier for the content element (*e.g.*, a global identifier as described above with reference to FIG. 7) and the first data record, and a second reference between the identifier for the content element and the second data record. These references are stored in CCRM module 408 (FIG. 4). Data mining processor 412 may store the references at step 1206 using any of the processes and embodiments described herein (*e.g.*, those discussed above with reference to FIGS. 7 and 8). At step 1208, data mining processor 412 assigns a first confidence score to the first metadata field. A confidence score may take any of a number of forms, such as a numerical value (*e.g.*, between 0 and 1) or a qualitative value (*e.g.*, "good" or "bad"). In some embodiments, the first confidence score represents the quality, accuracy, usefulness or other characteristic of the first metadata field. The confidence score may be manually entered into catalog and recommendation system 400 via editorial influence module 410 (which may not require the use of data mining processor 412), imported into catalog and recommendation system 400 along with the first data record via input 438 and content information database 414, assigned to the first data record by a third-party service, determined by data mining processor 412 by comparing the metadata to other metadata for the content element and using similarity as a measure of confidence, or any other method or combination of methods. Data mining processor 412 may assign the confidence score at step 1208 in any of a number of ways, such as by storing the confidence score in a designated portion of a memory of CCRM MODULE 408 with a pointer to the first data record or including the confidence score as additional metadata in the first data record, for example.

[0067] At step 1210, data mining processor 412 assigns a second confidence score to the second metadata field of the second data record (*e.g.*, using any of the techniques described above with reference to step 1208). At step 1212, data mining processor 412 identifies and responds to inconsistencies in the different data records and metadata fields that describe a particular content element in order to transmit information about that content element to a client device. To do so, data mining processor 412 may use the first and second confidence scores assigned at steps 1208 and 1210

as described in detail below with reference to FIG. 13. Metadata for the content element may be transmitted to the client device via CCRM module 408, real-time service processor 418 and device gateway 434, and may be supplemented by information from domains relations module 420 and/or asset metadata module 422 (FIG. 4). In some embodiments, data mining processor 412 transmits the metadata at step 1212 via CCRM MODULE 408 by selecting different pieces of metadata from each of the portions of the first and second metadata according to the confidence scores. For example, data mining processor 412 may be configured to include a particular attribute in metadata transmitted at step 1212 (*e.g.*, "Title") and may select the corresponding metadata value from either the first data record or the second data record according to which has the higher confidence score assigned to the metadata fields associated with the attribute "Title." In some embodiments, the metadata transmitted at step 1212 includes each unique metadata field from the first data record and each unique metadata field from the second data record, and the confidence scores are used to determine which value is included if both the first and second data records have values for a particular attribute.

[0068] A technique for reconciling inconsistent data records using confidence scores is illustrated in flow diagram 1300 of FIG. 13. At step 1302, data mining processor 412 receives a request for information about a particular attribute of a particular content element (referred to as the "desired attribute" and the "desired content element," respectively). This request may be part of a request for information about multiple attributes of the desired content element, and data mining processor 412 may be configured to respond to the request by transmitting information attribute by attribute, or as a batch. In some implementations, the request received at step 1302 is a request from a client device received via path 440 (FIG. 4). When the client device is part of an intermediate service that provides content information to downstream user devices, the request from the client device may be sent in response to a request from a user device further downstream. In some implementations, the request from the user device may be a request for a recommendation; the client device may transmit that request to catalog and recommendation system 400, which may identify content elements (including the desired content element) for the recommendation and provide attribute information (including the desired attribute) for those content elements. At step 1304, data mining processor 412 identifies the one or more data records that represent the desired content element within CCRM MODULE 408. Data mining processor may identify these data records by looking for which data records are associated with a global identifier for the content element, as described above with reference to FIGS. 6-11. At step 1306, data mining processor 412 indexes the metadata fields within the data records identified at step 1304 by assigning each metadata field an index value to form an ordered list of index values. Data mining processor 412 uses this index to sequentially analyze each metadata field in order to determine the highest confidence information about the desired attribute. Within the index, the ordering of the metadata fields from a single data record or across multiple data records is arbitrary, and may be assigned in any manner.

[0069] After indexing the metadata fields at step 1306, data mining processor 412 may initialize three variables. At step 1308, data mining processor 412 sets an INDEX variable equal to 1 (or any other suitable initial value). Data mining processor 412 may use the INDEX variable to keep track of which metadata field is currently being analyzed within the list of metadata fields indexed at step 1306. At step 1310, data mining processor 412 establishes an empty TEMP_WINNER variable (or sets the TEMP_WINNER variable to any predetermined default value). As the analysis of the different metadata fields progresses, data mining processor may use the TEMP_WINNER variable to store the metadata field (or a pointer to the metadata field) that provides the highest confidence information about the desired attribute. At step 1312, data mining processor 412 sets a TEMP_CONF variable equal to 0 (or any other suitable initial value). Data mining processor 412 may use the TEMP_CONF variable to store the confidence score of the metadata field stored in or pointed to by the TEMP_WINNER variable.

[0070] At step 1314, data mining processor 412 determines whether the metadata field associated with the index value INDEX is descriptive of the desired attribute. Data mining processor may make this determination using any of a number of techniques, including comparing the name of the desired attribute and the name of the metadata field (*e.g.*, for an exact match, a partial match, or using any other matching heuristic). For example, if the desired attribute is "Title," data mining processor may determine that a metadata field with name "Working Title" is descriptive of the desired attribute. If the metadata field associated with the index value INDEX is determined to be descriptive of the desired attribute at step 1314, data mining processor 412 determines whether the confidence score of the metadata field is greater than the value of the TEMP_CONF variable at step 1316. If yes, data mining processor 412 sets the TEMP_WINNER variable equal to (or pointing to) the metadata field associated with the index value INDEX at step 1318, and sets the TEMP_CONF variable equal to the confidence score of the metadata field associated with the index value INDEX at step 1320. In some implementations, the initial value of the TEMP_CONF value (set at step 1312) may be larger than the minimum possible confidence score; in such implementations, a metadata field must have confidence exceeding this initial threshold in order to be considered for transmission to the client device.

[0071] If data mining processor 412 determines at step 1314 that the metadata field associated with the index value INDEX is not descriptive of the desired attribute, or if data mining processor 412 determines at step 1316 that the confidence score of that metadata field is less than or equal to the value of the TEMP_CONF variable, data mining processor determines whether all metadata fields for the desired content element have been processed (step 1322). If no, data mining processor 412 increments the value of the INDEX variable at step 1324 and returns to step 1314 to evaluate the next metadata field in the list. If yes, data mining processor 412 transmits the value of the metadata field stored in or pointed to by the TEMP_WINNER variable to the client device at step 1326. Data mining processor 412 may transmit this value in conjunction with CCRM module 408, real-time service processor 418, and device gateway 434 (FIG. 4). In this manner, data mining processor 412

may select, for eventual display on a client or user device, metadata field values that have the highest confidence values among other metadata field values descriptive of the same attribute.

[0072] In some implementations of the process of FIG. 13, data mining processor 412 may index the metadata fields at step 1306 by confidence score in descending order. In such implementations, when data mining processor 412 first determines that a metadata field is descriptive of the desired attribute (and has at least a threshold confidence score, if one is specified), data mining processor 412 may proceed directly to step 1326 and transmit the value of that metadata field to the client device. Because the remaining metadata fields have confidence scores that are less than or equal to the score of that metadata field, data mining processor 412 need not consider additional metadata fields.

[0073] FIGS. 14A and 14B provide an illustrative example of the processes of FIGS. 12 and 13. FIG. 14A depicts an example content data structure with multiple global identifiers 1402 and 1404 and multiple records 1406, 1408, 1410 and 1412. Content data structure 1400 also includes several sets of metadata from different metadata sources associated with the records. In FIG. 14A, record 1406 includes Source1 Metadata 1414, Netflix Ratings 1416 and Wikipedia Entry 1418; record 1408 includes Source2 Metadata 1420; record 1410 includes Source1 Metadata 1422, Amazon Reviews 1424 and Source1 Metadata 1426; and record 1412 includes Source1 Metadata 1426 and Source1 Metadata 1428. As used herein, a record may include a set of metadata by storing the metadata in a memory structure associated with that record, or by pointing to the metadata from a data structure associated with the record for example. As illustrated by Source1 Metadata 1426 in FIG. 14A, it is possible for a single set of metadata to be included in two or more data records. Similarly, it is possible for a single data record to be referenced by two or more global identifiers.

[0074] FIG. 14B depicts an example aggregate set of metadata 1430 associated with global identifier 1402 of FIG. 14A. Aggregate set of metadata 1430 includes a list of metadata fields 1432, a list of the records 1434 from which the metadata fields are drawn, a list of metadata sources 1436 from which the metadata fields are drawn, a list of values 1438 for the metadata fields, and a list of confidence scores 1440 for each of the metadata fields. As shown in FIG. 14B, different metadata fields from the same data record, or from the same metadata source, can have different confidence scores. Moreover, metadata fields with the same name (*e.g.*, "Name") may be included in different metadata sources associated with the same or different data records. FIG. 14B also illustrates that several of the metadata fields have been designated as "top" fields. These fields have the highest confidence scores among all fields with the same name (as determined, for example, using the process illustrated in FIG. 13). In some embodiments, when catalog and recommendation system 400 transmits metadata to a client device at step 1212 of the flow diagram 1200 (FIG. 12), these "top" fields are included in the transmitted metadata.

[0075] The method of FIG. 12 may be used with any number of data records, any number of metadata fields, and may use confidence scores to select and transmit an aggregate set of metadata in any of a number of ways. In some embodiments, metadata is transmitted to a processor of a content

recommendation system (*e.g.*, real-time service processor 418 of FIG. 4, or media guidance data source 368 of FIG. 3B) and is used as the basis for transmitting a recommendation for the content element. In some such embodiments, the metadata is compared to a user input (*e.g.*, a search query, a "More Like This" request," a log-in, a user profile entry, or any other input to which content data is transmitted in response). The metadata may be displayed or otherwise communicated to the user (*e.g.*, via a media guidance application or other application running on user equipment as described herein with reference to FIGS. 1-3).

[0076] In some embodiments, data mining processor 412 assigns a confidence score to a reference between a global identifier and a data record instead of or in addition to assigning a confidence score to metadata. This confidence score may indicate, for example, the likelihood that the data record is truly descriptive of the content element represented by the global identifier. In some embodiments of the method of FIG. 12, data mining processor 412 assigns a third confidence score to the first reference and a fourth confidence score to the second reference. Then, at step 1212, data mining processor 412 draws the transmitted metadata from the first set of metadata based on the first and third confidence scores, and from the second set of metadata based on the second and fourth confidence scores. For example, data mining processor 412 may multiply the first and third confidence scores to obtain an aggregate confidence score for the portion of the first metadata, or may apply thresholds to both the first and third confidence scores and only use the portion of the first metadata if both thresholds are exceeded (and analogously for the second and fourth confidence scores). These techniques are simply illustrative, and any techniques for using two or more confidence scores may be used.

[0077] In some embodiments of catalog and recommendation system 400, an administrator or automated process can transmit an instruction to "lock" one or more portions of a content data catalog. Examples of lockable portions include references, data records, and metadata fields. When a lock instruction is issued for a portion, further changes to that portion (including removal of the portion) are prohibited until the lock instruction is revoked or overridden. In some embodiments, transmitting a lock instruction is equivalent to assigning a confidence score of 100% (or other score value indicating that the element is correct and should not be changed). Locked portions, then, may have the highest possible confidence, and will therefore "best" any other elements during an aggregation process like the metadata provision process of FIG. 12.

[0078] One particular type of locked data is editorial metadata. Editorial metadata is metadata included in data records that are designated as locked (or as having the highest possible confidence as discussed above). Editorial metadata may be changed by catalog administrators, and may represent metadata that has been entered or validated by trusted human editors. Editorial metadata may be input to CCRM module 408 via editorial influence module 410 via a personal computer or other device such as client device 300 of FIG. 3A.

[0079] FIG. 15 depicts an example of a content data structure including editorial metadata, locked global identifiers and locked references. Content data structure 1500 includes global identifiers 1502 and 1504. Global identifier 1502 is depicted as locked. A locked global identifier may indicate that all data records (including editorial metadata) with which the global identifier is associated should not be changed, nor shall references be removed. Content data structure 1500 also includes a locked reference 1512 from global identifier 1504 to record 1510. A locked reference may not be removed or changed, for example. Content data structure 1500 also includes editorial metadata 1514 and editorial metadata 1516 (as described above). In some embodiments, instructions to perform a prohibited operation to a locked portion may trigger an error message or other warning. An administrator may be given the opportunity to unlock the locked portion. In some embodiments, when the content of locked metadata (*e.g.*, editorial metadata) is in conflict (*e.g.*, two sets of editorial metadata associated with a movie provide different view times for the movie), an administrator may be prompted to reconcile the conflict.

[0080] FIG. 16 is an illustrative administrator-facing display 1600 for receiving locking instructions. This display may be presented on display 312 of client device 300 of FIG. 3A, and locking instructions may be received through user input interface 310 and passed to CCRM module 408 or data mining processor 412 via I/O path 302. An administrator may specify one or more portions of the content data catalog to lock (such as global identifiers, references, and metadata sources or fields). The administrator may specify these portions from one or more lists, or editorial influence module 410 may automatically populate display 1600 using information from portions selected prior to display 1600. In some embodiments (as discussed above,) once a portion is designated as "locked," it can only be unlocked or changed by an administrator or other individual or process with unlocking/overriding authority.

[0081] The following discussion addresses further embodiments of display screens, client devices and systems suitable for use with the asset cataloging, search, and recommendation techniques described herein. As noted above, the following discussion will often be presented in the context of media guidance applications, but it will be understood that these illustrative examples do not limit the range of interactive applications which may be improved by the use of the asset cataloging, search, and recommendation techniques of the present disclosure.

[0082] With the advent of the Internet, mobile computing, and high-speed wireless networks, users are accessing media on client devices on which they traditionally did not. As referred to herein, the phrase "client device," "user equipment," "user device," "electronic device," "electronic equipment," "media equipment device," or "media device" should be understood to mean any device for accessing the content described above, such as a television, a Smart TV, a set-top box, an integrated receiver decoder (IRD) for handling satellite television, a digital storage device, a digital media receiver (DMR), a digital media adapter (DMA), a streaming media device, a DVD player, a DVD recorder, a connected DVD, a local media server, a BLU-RAY player, a BLU-RAY recorder, a personal

computer (PC), a laptop computer, a tablet computer, a WebTV box, a personal computer television (PC/TV), a PC media server, a PC media center, a hand-held computer, a stationary telephone, a personal digital assistant (PDA), a mobile telephone, a portable video player, a portable music player, a portable gaming machine, a smart phone, or any other television equipment, computing equipment, or wireless device, and/or combination of the same. In some embodiments, the client device may have a front facing screen and a rear facing screen, multiple front screens, or multiple angled screens. In some embodiments, the client device may have a front facing camera and/or a rear facing camera. On these client devices, users may be able to navigate among and locate the same content available through a television. Consequently, media guidance may be available on these devices, as well. The guidance provided may be for content available only through a television, for content available only through one or more of other types of client devices, or for content available both through a television and one or more of the other types of client devices. The media guidance applications may be provided as on-line applications (*i.e.*, provided on a web-site), or as stand-alone applications or clients on client devices. The various devices and platforms that may implement media guidance applications are described in more detail below.

[0083] In addition to providing access to linear programming (*e.g.*, content that is scheduled to be transmitted to a plurality of client devices at a predetermined time and is provided according to a schedule), the media guidance application also provides access to non-linear programming (*e.g.*, content accessible to a client device at any time and is not provided according to a schedule). Non-linear programming may include content from different content sources including on-demand content (*e.g.*, VOD), Internet content (*e.g.*, streaming media, downloadable media, *etc.*), locally stored content (*e.g.*, content stored on any client device described above or other storage device), or other time-independent content. On-demand content may include movies or any other content provided by a particular content provider (*e.g.*, HBO On Demand providing "The Sopranos" and "Curb Your Enthusiasm"). HBO ON DEMAND is a service mark owned by Time Warner Company L.P. et al. and THE SOPRANOS and CURB YOUR ENTHUSIASM are trademarks owned by the Home Box Office, Inc. Internet content may include web events, such as a chat session or Webcast, or content available on-demand as streaming content or downloadable content through an Internet web site or other Internet access (*e.g.*, FTP).

[0084] Grid 102 may provide media guidance data for non-linear programming including on-demand listing 114, recorded content listing 116, and Internet content listing 118. A display combining media guidance data for content from different types of content sources is sometimes referred to as a "mixed-media" display. The various permutations of the types of media guidance data that may be displayed that are different than display 100 may be based on user selection or guidance application definition (*e.g.*, a display of only recorded and broadcast listings, only on-demand and broadcast listings, *etc.*). As illustrated, listings 114, 116, and 118 are shown as spanning the entire time block displayed in grid 102 to indicate that selection of these listings may provide access to a

display dedicated to on-demand listings, recorded listings, or Internet listings, respectively. In some embodiments, listings for these content types may be included directly in grid 102. Additional media guidance data may be displayed in response to the user selecting one of the navigational icons 120. Pressing an arrow key on a user input device may affect the display in a similar manner as selecting navigational icons 120.

[0085] Advertisement 124 may provide an advertisement for content that, depending on a viewer's access rights (*e.g.*, for subscription programming), is currently available for viewing, will be available for viewing in the future, or may never become available for viewing, and may correspond to or be unrelated to one or more of the content listings in grid 102. Advertisement 124 may also be for products or services related or unrelated to the content displayed in grid 102. Advertisement 124 may be selectable and provide further information about content, provide information about a product or a service, enable purchasing of content, a product, or a service, provide content relating to the advertisement, *etc.* Advertisement 124 may be targeted based on a user's profile/preferences, monitored user activity, the type of display provided, or on other suitable targeted advertisement bases.

[0086] While advertisement 124 is shown as rectangular or banner shaped, advertisements may be provided in any suitable size, shape, and location in a guidance application display. For example, advertisement 124 may be provided as a rectangular shape that is horizontally adjacent to grid 102. This is sometimes referred to as a panel advertisement. In addition, advertisements may be overlaid over content or a guidance application display or embedded within a display. Advertisements may also include text, images, rotating images, video clips, or other types of content described above. Advertisements may be stored in a client device having a guidance application, in a database connected to the client, in a remote location (including streaming media servers), or on other storage means, or a combination of these locations. Providing advertisements in a media guidance application is discussed in greater detail in, for example, Knudson et al., U.S. Patent Application Publication No. 2003/0110499, filed January 17, 2003; Ward, III et al. U.S. Patent No. 6,756,997, issued June 29, 2004; and Schein et al. U.S. Patent No. 6,388,714, issued May 14, 2002, which are hereby incorporated by reference herein in their entireties. It will be appreciated that advertisements may be included in other media guidance application display screens of the embodiments described herein.

[0087] In an embodiment, display 200 of FIG. 2 may be augmented by any of the items and features described above for display 100 of FIG. 1. For example, advertisement 205 may take the form of any of the embodiments described above for advertisement 124. The listings in display 200 are of different sizes (*i.e.*, listing 206 is larger than listings 208, 210, and 212), but if desired, all the listings may be the same size. Listings may be of different sizes or graphically accentuated to indicate degrees of interest to the user or to emphasize certain content, as desired by the content provider or based on user preferences. Various systems and methods for graphically accentuating content listings

are discussed in, for example, Yates, U.S. Patent Application Publication No. 2010/0153885, filed December 29, 2005, which is hereby incorporated by reference herein in its entirety.

[0088] As discussed above, the systems and methods of the present disclosure may be implemented in whole or in part by client 300 of FIG. 3B, which includes control circuitry 304. Control circuitry 304 may be based on any suitable processing circuitry such as processing circuitry 306. As referred to herein, processing circuitry should be understood to mean circuitry based on one or more microprocessors, microcontrollers, digital signal processors, programmable logic devices, field-programmable gate arrays (FPGAs), application-specific integrated circuits (ASICs), *etc.*, and may include a multi-core processor (*e.g.*, dual-core, quad-core, hexa-core, or any suitable number of cores) or supercomputer. In some embodiments, processing circuitry may be distributed across multiple separate processors or processing units, for example, multiple of the same type of processing units (*e.g.*, two Intel Core i7 processors) or multiple different processors (*e.g.*, an Intel Core i5 processor and an Intel Core i7 processor). In some embodiments, control circuitry 304 executes instructions for a media guidance application stored in memory (*i.e.*, storage 308).

[0089] In client-server based embodiments, control circuitry 304 may include communications circuitry suitable for communicating with a guidance application server or other networks or servers. The instructions for carrying out the above mentioned functionality may be stored on the guidance application server. Communications circuitry may include a cable modem, an integrated services digital network (ISDN) modem, a digital subscriber line (DSL) modem, a telephone modem, Ethernet card, or a wireless modem for communications with other equipment, or any other suitable communications circuitry. Such communications may involve the Internet or any other suitable communications networks or paths (which is described in more detail in connection with FIG. 3A). In addition, communications circuitry may include circuitry that enables peer-to-peer communication of client devices, or communication of client devices in locations remote from each other (described in more detail below). Server-centric and/or peer-to-peer communication may enable the pooling of preferences and behaviors between users, for use with the systems and techniques disclosed herein.

[0090] Memory may be an electronic storage device provided as storage 308 that is part of control circuitry 304. As referred to herein, the phrase "electronic storage device" or "storage device" should be understood to mean any device for storing electronic data, computer software, or firmware, such as random-access memory, read-only memory, hard drives, optical drives, digital video disc (DVD) recorders, compact disc (CD) recorders, BLU-RAY disc (BD) recorders, BLU-RAY 3D disc recorders, digital video recorders (DVR, sometimes called a personal video recorder, or PVR_), solid state devices, quantum storage devices, gaming consoles, gaming media, or any other suitable fixed or removable storage devices, and/or any combination of the same. Storage 308 may be used to store various types of content described herein as well as media guidance information, described above, and guidance application data, described above. Nonvolatile memory may also be used (*e.g.*, to launch a

boot-up routine and other instructions). Cloud-based storage, described in relation to FIG. 3A, may be used to supplement storage 308 or instead of storage 308.

[0091] Control circuitry 304 may include video generating circuitry and tuning circuitry, such as one or more analog tuners, one or more MPEG-2 decoders or other digital decoding circuitry, high-definition tuners, or any other suitable tuning or video circuits or combinations of such circuits. Encoding circuitry (*e.g.*, for converting over-the-air, analog, or digital signals to MPEG signals for storage) may also be provided. Control circuitry 304 may also include scaler circuitry for upconverting and downconverting content into the preferred output format of the client device 300. Circuitry 304 may also include digital-to-analog converter circuitry and analog-to-digital converter circuitry for converting between digital and analog signals. The tuning and encoding circuitry may be used by the client device to receive and to display, to play, or to record content. The tuning and encoding circuitry may also be used to receive guidance data. The circuitry described herein, including for example, the tuning, video generating, encoding, decoding, encrypting, decrypting, scaler, and analog/digital circuitry, may be implemented using software running on one or more general purpose or specialized processors. Multiple tuners may be provided to handle simultaneous tuning functions (*e.g.*, watch and record functions, picture-in-picture (PIP) functions, multiple-tuner recording, *etc.*). If storage 308 is provided as a separate device from client device 300, the tuning and encoding circuitry (including multiple tuners) may be associated with storage 308.

[0092] The guidance application may be implemented using any suitable architecture. For example, it may be a stand-alone application wholly implemented on client device 300. In such an approach, instructions of the application are stored locally, and data for use by the application is downloaded on a periodic basis (*e.g.*, from an out-of-band feed, from an Internet resource, or using another suitable approach). In some embodiments, the media guidance application is a client-server based application. Data for use by a thick or thin client implemented on client device 300 is retrieved on-demand by issuing requests to a server remote to the client device 300. In one example of a client-server based guidance application, control circuitry 304 runs a web browser that interprets web pages provided by a remote server.

[0093] In some embodiments, the media guidance application is downloaded and interpreted or otherwise run by an interpreter or virtual machine (run by control circuitry 304). In some embodiments, the guidance application may be encoded in the ETV Binary Interchange Format (EBIF), received by control circuitry 304 as part of a suitable feed, and interpreted by a user agent running on control circuitry 304. For example, the guidance application may be an EBIF application. In some embodiments, the guidance application may be defined by a series of JAVA-based files that are received and run by a local virtual machine or other suitable middleware executed by control circuitry 304. In some of such embodiments (*e.g.*, those employing MPEG-2 or other digital media encoding schemes), the guidance application may be, for example, encoded and transmitted in an MPEG-2 object carousel with the MPEG audio and video packets of a program.

[0094] User television equipment 352 may include a set-top box, an integrated receiver decoder (IRD) for handling satellite television, a television set, a digital storage device, a DVD recorder, a video-cassette recorder (VCR), a local media server, or other user television equipment. One or more of these devices may be integrated into a single device, if desired. User computer equipment 354 may include a PC, a laptop, a tablet, a WebTV box, a personal computer television (PC/TV), a PC media server, a PC media center, or other user computer equipment. WEBTV is a trademark owned by Microsoft Corp. Wireless user communications device 356 may include PDAs, a mobile telephone, a portable video player, a portable music player, a portable gaming machine, or other wireless devices.

[0095] A client device utilizing at least some of the system features described above in connection with FIG. 3A may not be classified solely as user television equipment 352, user computer equipment 354, or a wireless user communications device 356. For example, user television equipment 352 may, like some user computer equipment 354, be Internet-enabled allowing for access to Internet content, while user computer equipment 354 may, like some television equipment 352, include a tuner allowing for access to television programming. The media guidance application may have the same layout on the various different types of client device or may be tailored to the display capabilities of the client device. For example, on user computer equipment 354, the guidance application may be provided as a web site accessed by a web browser. In another example, the guidance application may be scaled down for wireless user communications devices 356.

[0096] In system 350, there is typically more than one of each type of client device but only one of each is shown in FIG. 3A to avoid overcomplicating the drawing. In addition, each user may utilize more than one type of client device and also more than one of each type of client device.

[0097] In some embodiments, a client device (*e.g.*, user television equipment 352, user computer equipment 354, wireless user communications device 356) may be referred to as a "second screen device." For example, a second screen device may supplement content presented on a first client device. The content presented on the second screen device may be any suitable content that supplements the content presented on the first device. In some embodiments, the second screen device provides an interface for adjusting settings and display preferences of the first device. In some embodiments, the second screen device is configured for interacting with other second screen devices or for interacting with a social network. The second screen device can be located in the same room as the first device, a different room from the first device but in the same house or building, or in a different building from the first device.

[0098] The user may also set various settings to maintain consistent media guidance application settings across in-home devices and remote devices. Settings include those described herein, as well as channel and program favorites, programming preferences that the guidance application utilizes to make programming recommendations, display preferences, and other desirable guidance settings. For example, if a user sets a channel as a favorite on, for example, the web site www.allrovi.com on their personal computer at their office, the same channel would appear as a favorite on the user's in-home

devices (*e.g.*, user television equipment and user computer equipment) as well as the user's mobile devices, if desired. Therefore, changes made on one client device can change the guidance experience on another client device, regardless of whether they are the same or a different type of client device. In addition, the changes made may be based on settings input by a user, as well as user activity monitored by the guidance application.

[0099] Although communications paths are not drawn between client devices, these devices may communicate directly with each other via communication paths, such as those described above in connection with paths 358, 360, and 362, as well other short-range point-to-point communication paths, such as USB cables, IEEE 1394 cables, wireless paths (*e.g.*, Bluetooth, infrared, IEEE 802-11x, *etc.*), or other short-range communication via wired or wireless paths. BLUETOOTH is a certification mark owned by Bluetooth SIG, INC. The client devices may also communicate with each other directly through an indirect path via communications network 364.

[0100] System 350 includes content source 366 and media guidance data source 358 coupled to communications network 364 via communication paths 370 and 372, respectively. Paths 370 and 372 may include any of the communication paths described above in connection with paths 358, 360, and 362. Communications with the content source 366 and media guidance data source 358 may be exchanged over one or more communications paths, but are shown as a single path in FIG. 3A to avoid overcomplicating the drawing. In addition, there may be more than one of each of content source 366 and media guidance data source 358, but only one of each is shown in FIG. 3A to avoid overcomplicating the drawing. (The different types of each of these sources are discussed below.) If desired, content source 366 and media guidance data source 358 may be integrated as one source device. Although communications between sources 366 and 358 with client devices 352, 354, and 356 are shown as through communications network 364, in some embodiments, sources 366 and 358 may communicate directly with client devices 352, 354, and 356 via communication paths (not shown) such as those described above in connection with paths 358, 360, and 362.

[0101] Content source 366 may include one or more types of content distribution equipment including a television distribution facility, cable system headend, satellite distribution facility, programming sources (*e.g.*, television broadcasters, such as NBC, ABC, HBO, *etc.*), intermediate distribution facilities and/or servers, Internet providers, on-demand media servers, and other content providers. NBC is a trademark owned by the National Broadcasting Company, Inc., ABC is a trademark owned by ABC, Inc., and HBO is a trademark owned by Home Box Office, Inc. Content source 366 may be the originator of content (*e.g.*, a television broadcaster, a Webcast provider, *etc.*) or may not be the originator of content (*e.g.*, an on-demand content provider, an Internet provider of content of broadcast programs for downloading, *etc.*). Content source 366 may include cable sources, satellite providers, on-demand providers, Internet providers, over-the-top content providers, or other providers of content. Content source 366 may also include a remote media server used to store different types of content (including video content selected by a user), in a location remote from any

of the client devices. Systems and methods for remote storage of content, and providing remotely stored content to client devices are discussed in greater detail in connection with Ellis et al., U.S. Patent No. 7,761,892, issued July 20, 2010, which is hereby incorporated by reference herein in its entirety.

[0102] Media guidance data source 358 may provide media guidance data, such as the media guidance data described above. Media guidance application data may be provided to the client devices using any suitable approach. In some embodiments, the guidance application may be a stand-alone interactive television program guide that receives program guide data via a data feed (*e.g.*, a continuous feed or trickle feed). Program schedule data and other guidance data may be provided to the client device on a television channel sideband, using an in-band digital signal, using an out-of-band digital signal, or by any other suitable data transmission technique. Program schedule data and other media guidance data may be provided to client devices on multiple analog or digital television channels.

[0103] In some embodiments, guidance data from media guidance data source 358 may be provided to users' equipment using a client-server approach. For example, a client device may pull media guidance data from a server, or a server may push media guidance data to a client device. In some embodiments, a guidance application client residing on the user's equipment may initiate sessions with source 358 to obtain guidance data when needed, *e.g.*, when the guidance data is out of date or when the client device receives a request from the user to receive data. Media guidance may be provided to the client device with any suitable frequency (*e.g.*, continuously, daily, a user-specified period of time, a system-specified period of time, in response to a request from a client device, *etc.*). Media guidance data source 358 may provide, to user equipment devices 352, 354, and 356, the media guidance application itself or software updates for the media guidance application.

[0104] Media guidance applications may be, for example, stand-alone applications implemented on client devices. In some embodiments, media guidance applications may be client-server applications where only the client resides on the client device. For example, media guidance applications may be implemented partially as a client application on control circuitry 304 of client device 300 and partially on a remote server as a server application (*e.g.*, media guidance data source 358). The guidance application displays may be generated by the media guidance data source 358 and transmitted to the client devices. The media guidance data source 358 may also transmit data for storage on the client, which then generates the guidance application displays based on instructions processed by control circuitry.

[0105] Content and/or media guidance data delivered to client devices 374, such as user equipment devices 352, 354, and 356 may be over-the-top (OTT) content. OTT content delivery allows Internet-enabled user devices, including any client device described above, to receive content that is transferred over the Internet, including any content described above. OTT content is delivered via an Internet connection provided by an Internet service provider (ISP), but a third party distributes the

content. The ISP may not be responsible for the viewing abilities, copyrights, or redistribution of the content, and may only transfer IP packets provided by the OTT content provider. Examples of OTT content providers include YOUTUBE, NETFLIX, and HULU, which provide audio and video via IP packets. OTT content providers may additionally or alternatively provide media guidance data described above. In addition to content and/or media guidance data, providers of OTT content can distribute media guidance applications (*e.g.*, web-based applications or cloud-based applications), or the content can be displayed by media guidance applications stored on the client device.

[0106] Media guidance data source 358 may make asset cataloging or recommendation applications available to users. Such applications may be downloaded from media guidance data source 368 to a client device, or may be accessed remotely by a user. These applications, as well as other applications, features and tools, may be provided to users on a subscription basis or may be selectively downloaded or used for an additional fee. In an embodiment, media guidance data source 368 may serve as a repository for media asset data developed by users and/or third-parties, and as a distribution source for this data and related applications.

[0107] Media guidance system 350 is intended to illustrate a number of approaches, or network configurations, by which client devices and sources of content and guidance data may communicate with each other for the purpose of accessing content and providing media guidance. The embodiments described herein may be applied in any one or a subset of these approaches, or in a system employing other approaches for delivering content and providing media guidance. The following four approaches provide specific illustrations of the generalized example of FIG. 3A.

[0108] In one approach, client devices may communicate with each other within a home network. Client devices can communicate with each other directly via short-range point-to-point communication schemes described above, via indirect paths through a hub or other similar device provided on a home network, or via communications network 364. Each of the multiple individuals in a single home may operate different client devices on the home network. As a result, it may be desirable for various media guidance information or settings to be communicated between the different client devices. For example, it may be desirable for users to maintain consistent media guidance application settings on different client devices within a home network, as described in greater detail in Ellis et al., U.S. Patent Application No. 11/179,360, filed July 11, 2005. Different types of client devices in a home network may also communicate with each other to transmit content. For example, a user may transmit content from user computer equipment to a portable video player or portable music player.

[0109] In a second approach, users may have multiple types of client devices by which they access content and obtain media guidance. For example, some users may have home networks that are accessed by in-home and mobile devices. Users may control in-home devices via a media guidance application implemented on a remote device. For example, users may access an online media guidance application on a website via a personal computer at their offices, or a mobile device such as

a PDA or web-enabled mobile telephone. The user may set various settings (*e.g.*, recordings, reminders, or other settings) on the online guidance application to control the user's in-home equipment. The online guide may control the user's equipment directly, or by communicating with a media guidance application on the user's in-home equipment. Various systems and methods for client devices communicating, where the client devices are in locations remote from each other, is discussed in, for example, Ellis et al., U.S. Patent No. 8,046,801, issued October 25, 2011, which is hereby incorporated by reference herein in its entirety.

[0110] In a third approach, users of client devices inside and outside a home can use their media guidance application to communicate directly with content source 366 to access content. Specifically, within a home, users of user television equipment 352 and user computer equipment 354 may access the media guidance application to navigate among and locate desirable content. Users may also access the media guidance application outside of the home using wireless user communications devices 356 to navigate among and locate desirable content.

[0111] In a fourth approach, client devices may operate in a cloud computing environment to access cloud services. In a cloud computing environment, various types of computing services for content sharing, storage or distribution (*e.g.*, video sharing sites or social networking sites) are provided by a collection of network-accessible computing and storage resources, referred to as "the cloud." For example, the cloud can include a collection of server computing devices, which may be located centrally or at distributed locations, that provide cloud-based services to various types of users and devices connected via a network such as the Internet via communications network 364. These cloud resources may include one or more content sources 366 and one or more media guidance data sources 358. In addition or in the alternative, the remote computing sites may include other client devices, such as user television equipment 352, user computer equipment 354, and wireless user communications device 356. For example, the other client devices may provide access to a stored copy of a video or a streamed video. In such embodiments, client devices may operate in a peer-to-peer manner without communicating with a central server.

[0112] The cloud provides access to services, such as content storage, content sharing, or social networking services, among other examples, as well as access to any content described above, for client devices. Services can be provided in the cloud through cloud computing service providers, or through other providers of online services. For example, the cloud-based services can include a content storage service, a content sharing site, a social networking site, or other services via which user-sourced content is distributed for viewing by others on connected devices. These cloud-based services may allow a client device to store content to the cloud and to receive content from the cloud rather than storing content locally and accessing locally-stored content.

[0113] A user may use various content capture devices, such as camcorders, digital cameras with video mode, audio recorders, mobile phones, and handheld computing devices, to record content. The user can upload content to a content storage service on the cloud either directly, for example, from

user computer equipment 354 or wireless user communications device 356 having content capture feature. Alternatively, the user can first transfer the content to a client device, such as user computer equipment 354. The client device storing the content uploads the content to the cloud using a data transmission service on communications network 364. In some embodiments, the client device itself is a cloud resource, and other client devices can access the content directly from the client device on which the user stored the content.

[0114] Cloud resources may be accessed by a client device using, for example, a web browser, a media guidance application, a desktop application, a mobile application, and/or any combination of access applications or the same. The client device may be a cloud client that relies on cloud computing for application delivery, or the client device may have some functionality without access to cloud resources. For example, some applications running on the client device may be cloud applications, *i.e.*, applications delivered as a service over the Internet, while other applications may be stored and run on the client device. In some embodiments, a client device may receive content from multiple cloud resources simultaneously. For example, a client device can stream audio from one cloud resource while downloading content from a second cloud resource. Or, a client device can download content from multiple cloud resources for more efficient downloading. In some embodiments, client devices can use cloud resources for processing operations such as the processing operations performed by processing circuitry described in relation to FIGS. 3A and 3B.

[0115] It is to be understood that while the invention has been described in conjunction with the various illustrative embodiments, the forgoing description is intended to illustrate and not limit the scope of the invention. While several embodiments have been provided in the present disclosure, it should be understood that the disclosed systems, components, and methods may be embodied in many other specific forms without departing from the scope of the present disclosure.

[0116] The intention is not to be limited to the details given herein or implemented in sub-combinations with one or more other features described herein. For example, a variety of systems and methods may be implemented based on the disclosure and still fall within the scope of the invention. Also, the various features described or illustrated above may be combined or integrated in other systems or certain features may be omitted, or not implemented.

[0117] Examples of changes, substitutions, and alterations are ascertainable by one skilled in the art and could be made without departing from the scope of the information disclosed herein. Certain particular aspects, advantages, and modifications are within the scope of the following claims. All references cited herein are incorporated by reference in their entirety and made part of this application.

What is claimed is:

1. A method of providing metadata to a client device based on inconsistent data records, comprising:

storing, with a processing device in a memory device of, first and second data records representative of a content element, the first data record including a first metadata field with a first value for a first attribute of the content element, the second data record including a second metadata field with a second value for the first attribute of the content element, and the data records being inconsistent in that the first value is different from the second value;

storing, with the processing device in the memory device, a confidence score for the first metadata field and a confidence score for the second metadata field;

receiving, at the processing device, a request from a client device for metadata descriptive of the common content element;

identifying, with the processing device, which of the first and second metadata fields has the greater confidence score; and

transmitting, from the processing device over a computer network to the client device for display, the value of the identified metadata field.

2. The method of claim 1, further comprising:

receiving, at the processing device, the first data record from a first cataloging system; and

receiving, at the processing device, the second data record from a second cataloging system different from the first cataloging system.

3. The method of claim 1, wherein a name of the first metadata field is the same as a name of the second metadata field.

4. The method of claim 1, further comprising:

retaining the value of the metadata field with the lesser confidence score in the memory device after selecting the value of the metadata field with the greater confidence score.

5. The method of claim 1, wherein the computer network comprises the Internet.

6. The method of claim 1, wherein the request from the client device is transmitted in response to receiving, at the client device, a request for a recommendation from a user device.

7. The method of claim 6, wherein the client device is configure to transmit the value of the identified metadata field to a user device for display as a recommendation for the content element.

8. The method of claim 1, wherein the first data record further includes a third metadata field with a third value for a second attribute of the content element and the second data record further includes a fourth metadata field with a fourth value for the second attribute of the content element, and the third value is different from the fourth value, and further comprising:

storing, with the processing device in the memory device, a confidence score for the third metadata field and a confidence score for the fourth metadata field;

identifying, with the processing device, which of the third and fourth metadata fields has the greater confidence score; and

transmitting, from the processing device to the client device for display, the value of the identified metadata field.

9. The method of claim 1, further comprising:

receiving, at the processing device from an administrator input device of a metadata cataloging system, an instruction to lock the first data record; and

in response to receiving the lock instruction, storing, with the processing device, a locked status for the first data record in the memory device, wherein the locked status indicates that data included in the first data record will not be changed by subsequent instructions to change the data.

10. The method of claim 1, further comprising:

receiving, at the processing device from an administrator input device of a metadata cataloging system, a new confidence score for the first metadata field; and

storing, with the processing device in the memory device, the new confidence score to replace the confidence score for the first metadata field.

11. A system for providing metadata to a client device based on inconsistent data records, comprising:

an input device;

a storage device;

an output device configured to transmit data to the client device over a computer network; and

a processing device in communication with the input device, the storage device, and the output device and configured to:

store, in the storage device, first and second data records representative of a content element, the first data record including a first metadata field with a first value for a first attribute of the content element, the second data record including a second metadata field with a second value for the first attribute of the content element, and the data records being inconsistent in that the first value is different from the second value;

store, in the memory device, a confidence score for the first metadata field and a confidence score for the second metadata field;

receive, via the input device, a request from a client device for metadata descriptive of the common content element;

identifying which of the first and second metadata fields has the greater confidence score; and

transmit, via the output device to the client device for display, the value of the identified metadata field.

12. The system of claim 11, wherein the processing device is further configured to:

receive, at the input device, the first data record from a first cataloging system; and

receive, at the input device, the second data record from a second cataloging system different from the first cataloging system.

13. The system of claim 11, wherein a name of the first metadata field is the same as a name of the second metadata field.

14. The system of claim 11, wherein the processing device is further configured to:

retain the value of the metadata field with the lesser confidence score in the memory device after selecting the value of the metadata field with the greater confidence score.

15. The system of claim 11, wherein the computer network comprises the Internet.

16. The system of claim 11, wherein the request from the client device is transmitted in response to receiving, at the client device, a request for a recommendation from a user device.

17. The system of claim 16, wherein the client device is configure to transmit the value of the identified metadata field to a user device for display as a recommendation for the content element.

18. The system of claim 11, wherein the first data record further includes a third metadata field with a third value for a second attribute of the content element and the second data record further includes a fourth metadata field with a fourth value for the second attribute of the content element, and the third value is different from the fourth value, and wherein the processing device is further configured to:

store, in the memory device, a confidence score for the third metadata field and a confidence score for the fourth metadata field;

identify which of the third and fourth metadata fields has the greater confidence score; and

transmit, to the client device for display, via the output device, t the value of the identified metadata field.

19. The system of claim 11, further comprising an administrator input device in communication with the processing device, wherein the processing device is further configured to:

receive, from the administrator input device, an instruction to lock the first data record; and

in response to receiving the lock instruction, store a locked status for the first data record in the memory device, wherein the locked status indicates that data included in the first data record will not be changed by subsequent instructions to change the data.

20. The system of claim 11, further comprising an administrator input device in communication with the processing device, wherein the processing device is further configured to:

receive, from the administrator input device, a new confidence score for the metadata field of the first data record; and

replace, in the memory device, the first confidence score for the metadata field of the first data record with the new confidence score.

21. A system for providing metadata to a client device based on inconsistent data records, comprising:

means for storing first and second data records representative of a content element, the first data record including a first metadata field with a first value for a first attribute of the content element, the second data record including a second metadata field with a second value for the first attribute of the content element, and the data records being inconsistent in that the first value is different from the second value;

means for storing a confidence score for the first metadata field and a confidence score for the second metadata field;

means for receiving a request from a client device for metadata descriptive of the common content element;

means for identifying which of the first and second metadata fields has the greater confidence score; and

means for transmitting, over a computer network to the client device for display, the value of the identified metadata field.

22. The system of claim 21, further comprising:

means for receiving the first data record from a first cataloging system; and

means for receiving the second data record from a second cataloging system different from the first cataloging system.

23. The system of claim 21, wherein a name of the first metadata field is the same as a name of the second metadata field.

24. The system of claim 21, further comprising:

means for retaining the value of the metadata field with the lesser confidence score after selecting the value of the metadata field with the greater confidence score.

25. The system of claim 21, wherein the computer network comprises the Internet.

26. The system of claim 21, wherein the request from the client device is transmitted in response to receiving, at the client device, a request for a recommendation from a user device.

27. The system of claim 26, wherein the client device is configure to transmit the value of the identified metadata field to a user device for display as a recommendation for the content element.

28. The system of claim 21, wherein the first data record further includes a third metadata field with a third value for a second attribute of the content element and the second data record further includes a fourth metadata field with a fourth value for the second attribute of the content element, and the third value is different from the fourth value, and further comprising:

means for storing a confidence score for the third metadata field and a confidence score for the fourth metadata field;

means for identifying which of the third and fourth metadata fields has the greater confidence score; and

means for transmitting, to the client device for display, the value of the identified metadata field.

29. The system of claim 21, further comprising:

means for receiving an instruction to lock the first data record; and

means for storing, in response to receiving the lock instruction, a locked status for the first data record in the memory device, wherein the locked status indicates that data included in the first data record will not be changed by subsequent instructions to change the data.

30. The system of claim 21, further comprising:

means for receiving a new confidence score for the metadata field of the first data record; and

means for replacing the first confidence score for the metadata field of the first data record with the new confidence score.

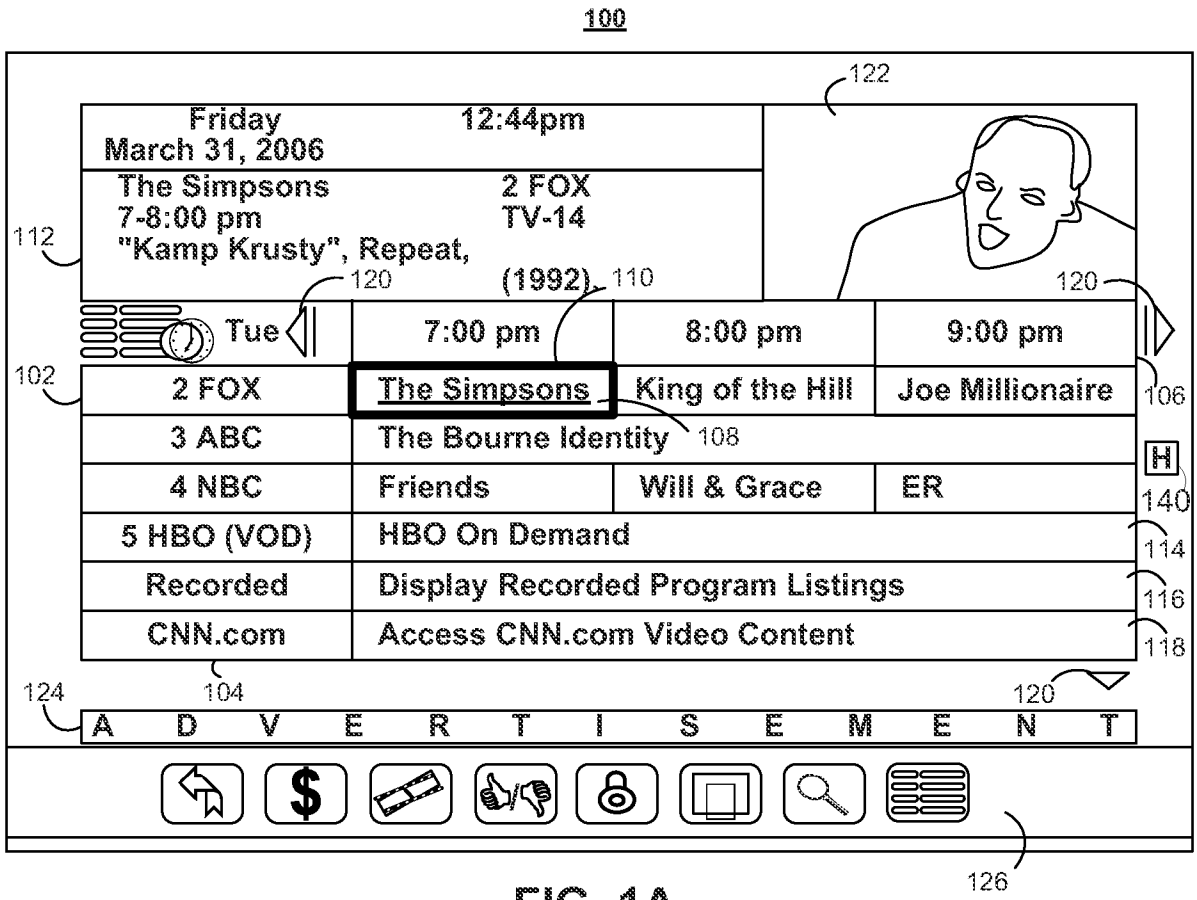

100
122
Friday
March 31, 2006
12:44pm
The Simpsons
7-8:00 pm
"Kamp Krusty", Repeat,
2 FOX
TV-14
(1992)
112
120
110
120
Tue
7:00 pm
8:00 pm
9:00 pm
102
106
2 FOX
The Simpsons
King of the Hill
Joe Millionaire
3 ABC
The Bourne Identity
108
H
4 NBC
Friends
Will & Grace
ER
140
5 HBO (VOD)
HBO On Demand
114
Recorded
Display Recorded Program Listings
116
CNN.com
Access CNN.com Video Content
118
124
104
120
A D V E R T I S E M E N T
126

FIG. 1A

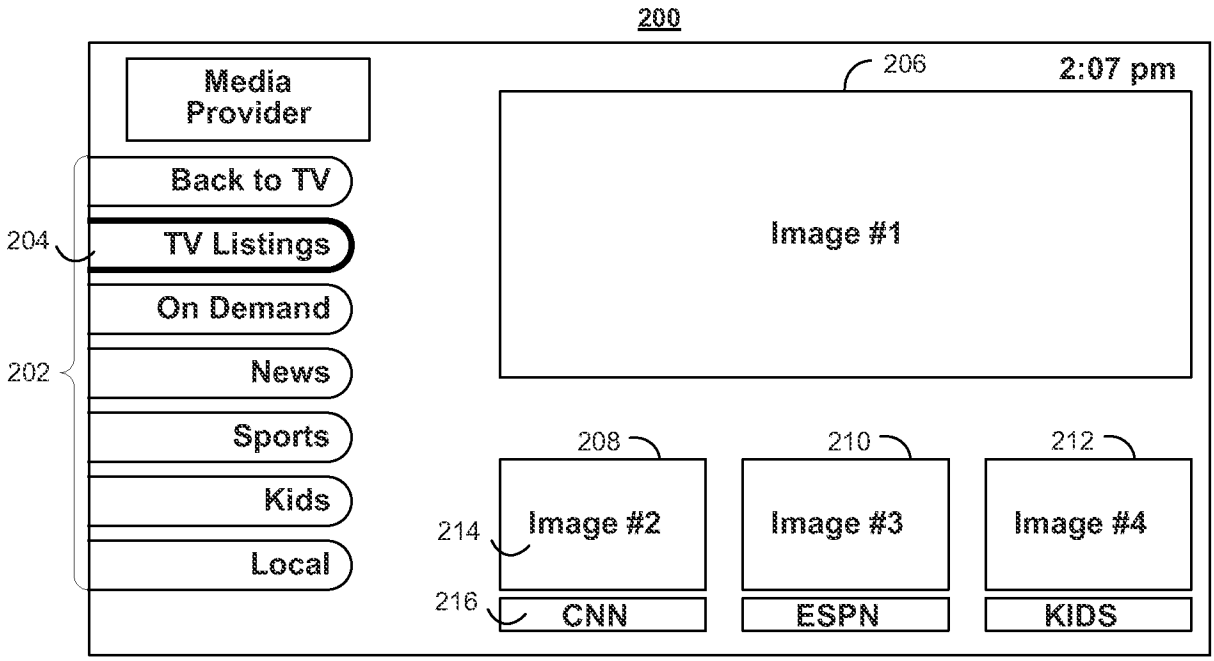

200
Media Provider
206
2:07 pm
Back to TV
TV Listings
On Demand
News
Sports
Kids
Local
204
202
Image #1
208
210
212
214
Image #2
Image #3
Image #4
216
CNN
ESPN
KIDS

FIG. 1B

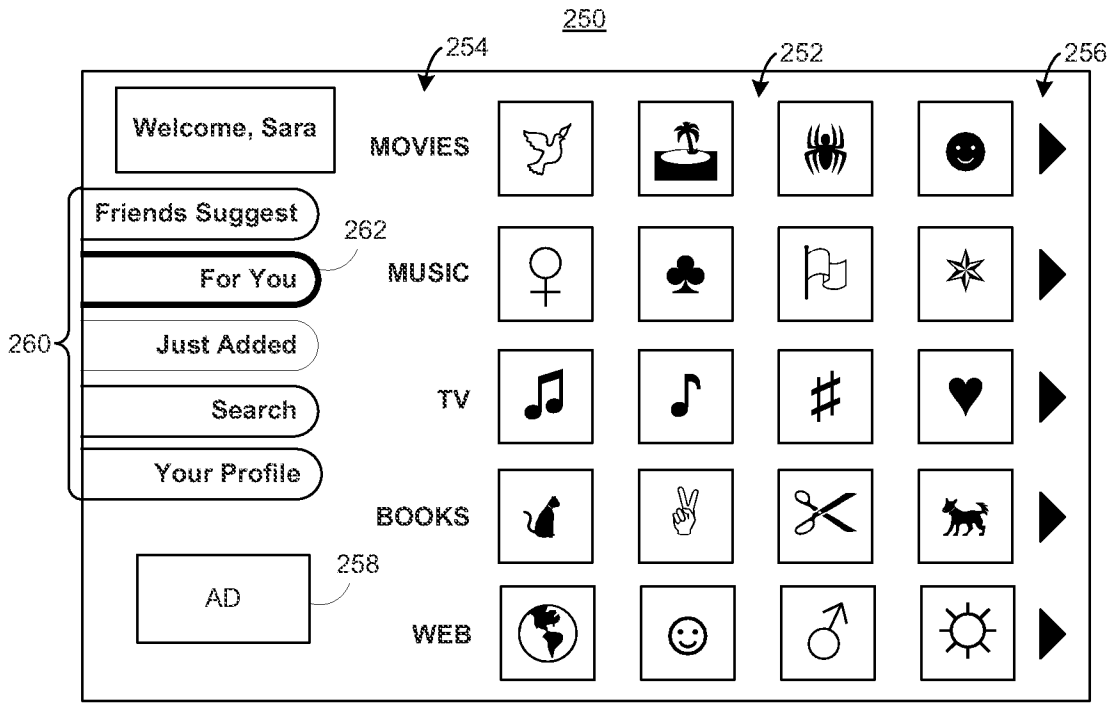

250
254
252
256
Welcome, Sara
MOVIES
Friends Suggest
262
For You
MUSIC
260
Just Added
Search
TV
Your Profile
BOOKS
258
AD
WEB

FIG. 2

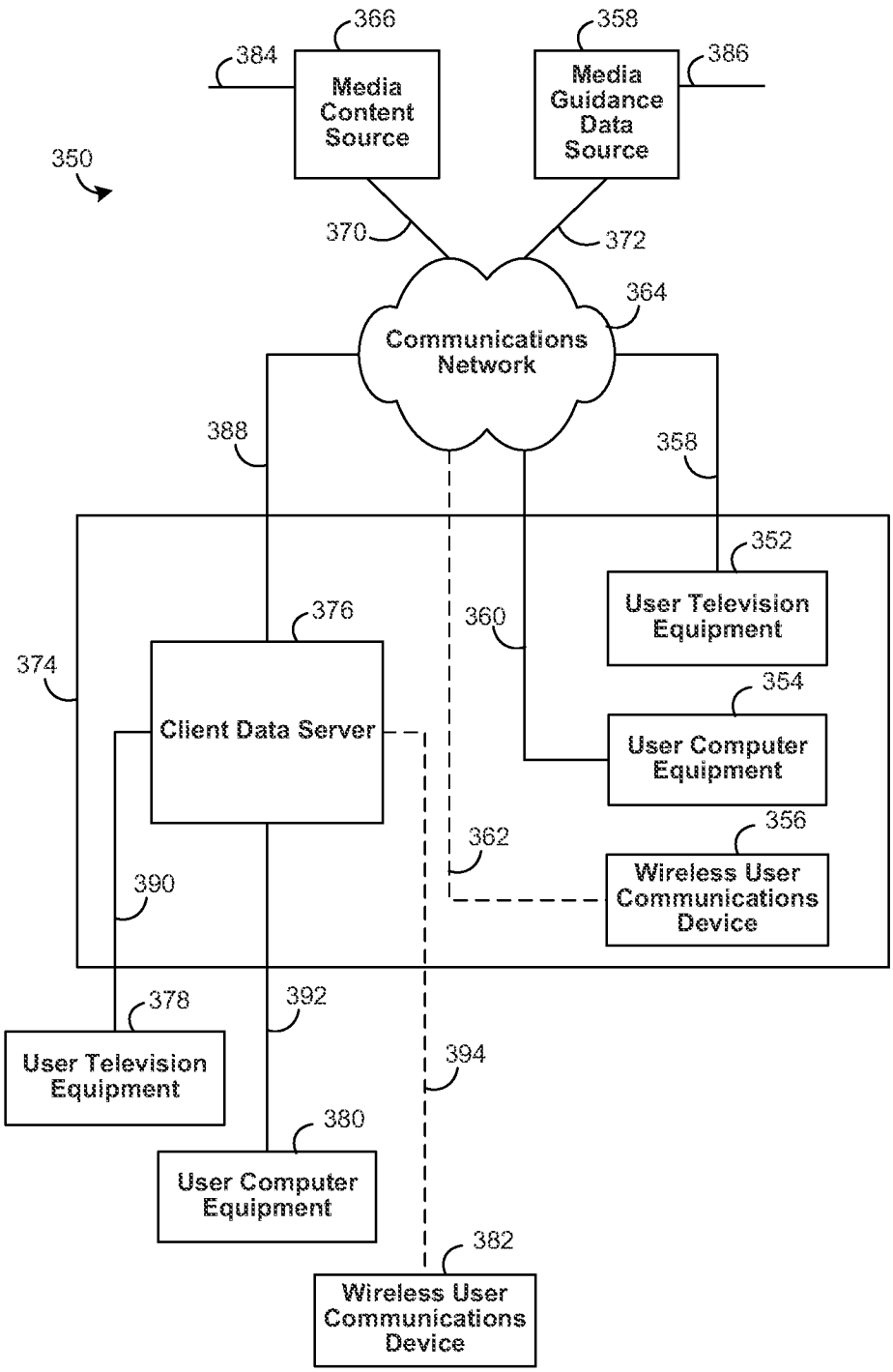

366
384
Media Content Source
358
386
Media Guidance Data Source
350
370
372
364
Communications Network
388
358
352
User Television Equipment
376
360
374
Client Data Server
354
User Computer Equipment
362
356
Wireless User Communications Device
390
378
392
User Television Equipment
394
380
User Computer Equipment
382
Wireless User Communications Device

FIG. 3A

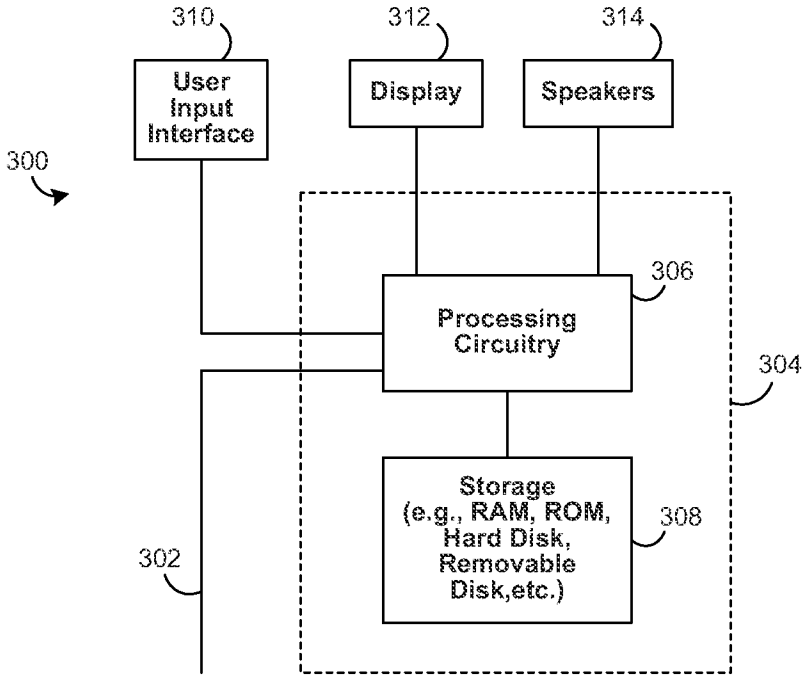

310
312
314
User Input Interface
Display
Speakers
300
306
Processing Circuitry
304
Storage (e.g., RAM, ROM, Hard Disk, Removable Disk,etc.)
308
302

FIG. 3B

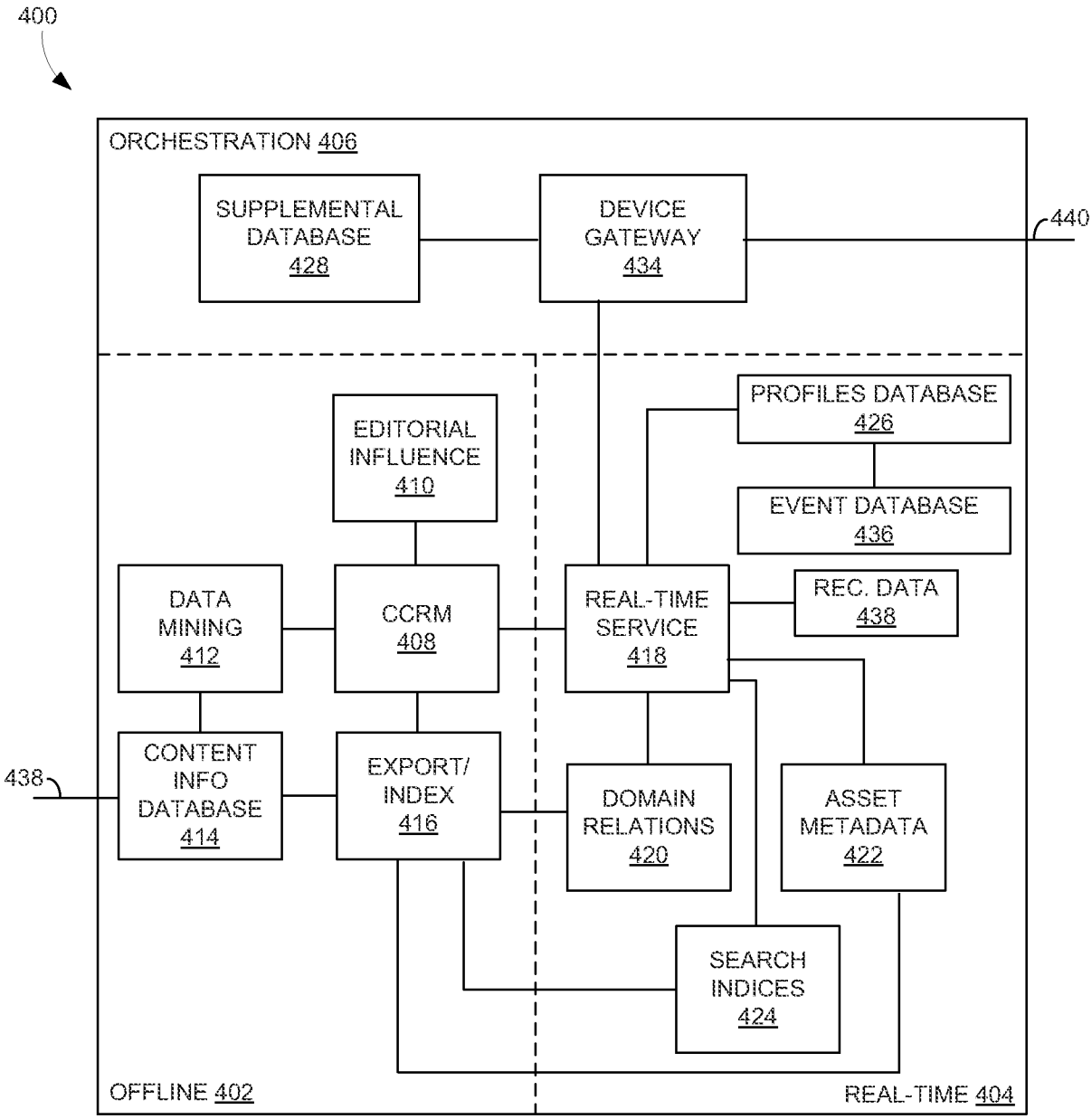

400
ORCHESTRATION 406
SUPPLEMENTAL DATABASE 428
DEVICE GATEWAY 434
440
PROFILES DATABASE 426
EDITORIAL INFLUENCE 410
EVENT DATABASE 436
DATA MINING 412
CCRM 408
REAL-TIME SERVICE 418
REC. DATA 438
438
CONTENT INFO DATABASE 414
EXPORT/ INDEX 416
DOMAIN RELATIONS 420
ASSET METADATA 422
SEARCH INDICES 424
OFFLINE 402
REAL-TIME 404

FIG. 4

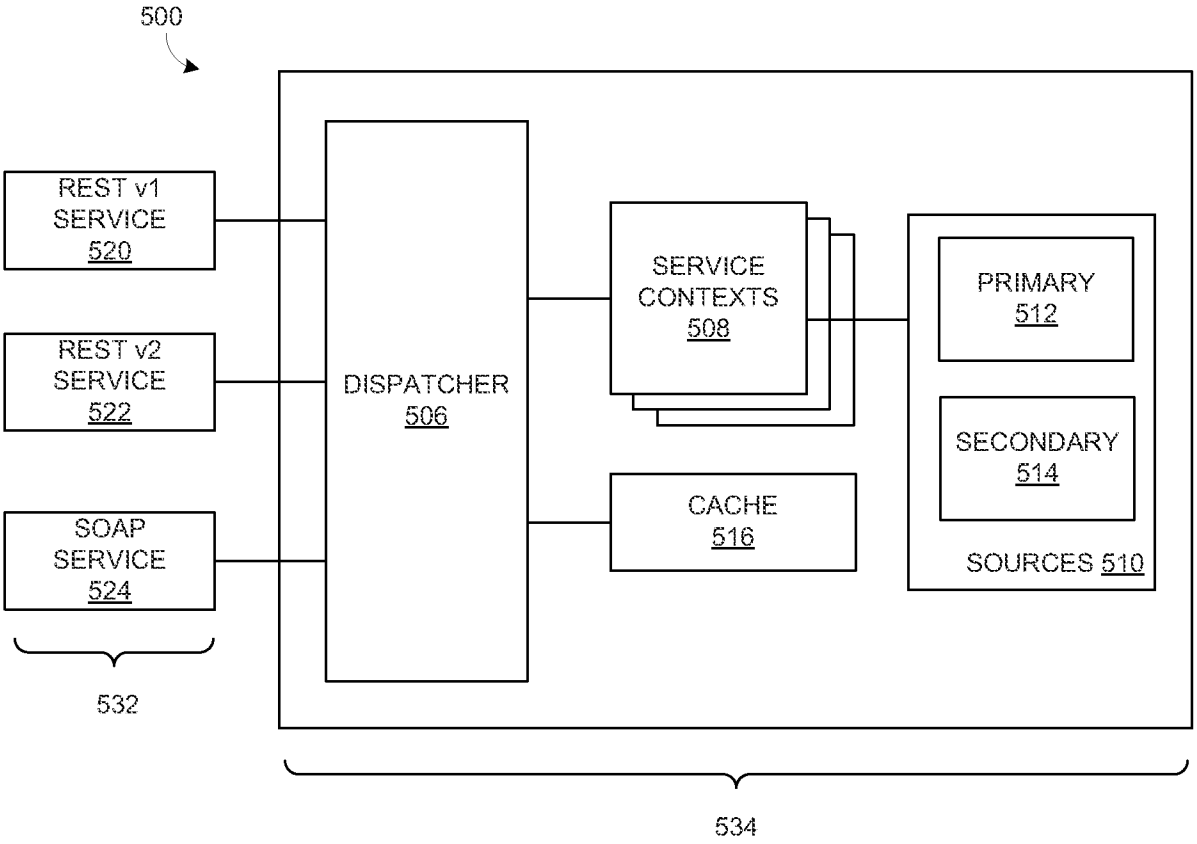

500
REST v1
SERVICE
520
REST v2
SERVICE
522
SOAP
SERVICE
524
532
DISPATCHER
506
SERVICE
CONTEXTS
508
CACHE
516
PRIMARY
512
SECONDARY
514
SOURCES 510
534

FIG. 5

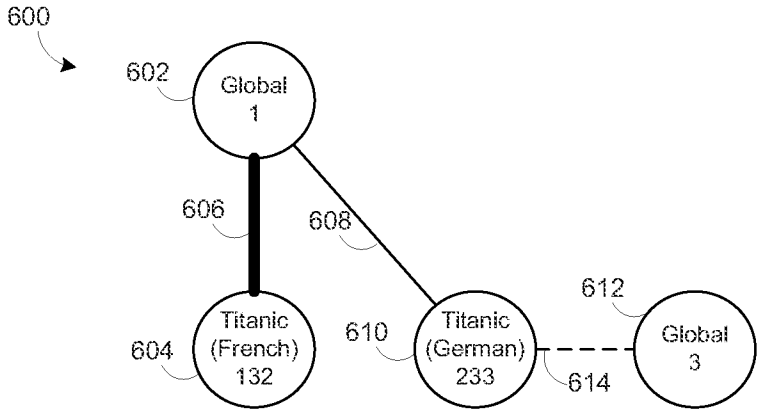


FIG. 6

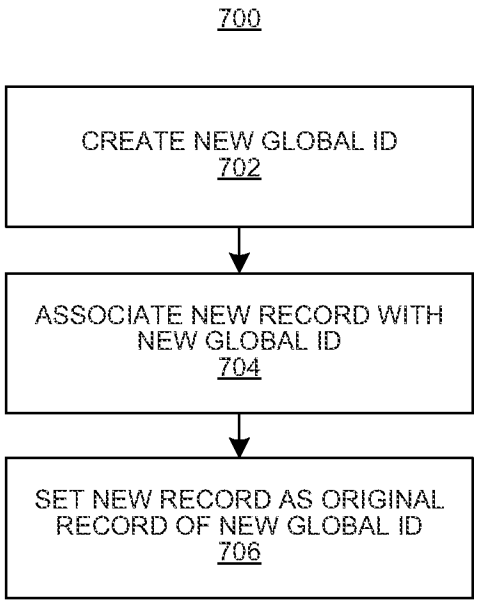


FIG. 7

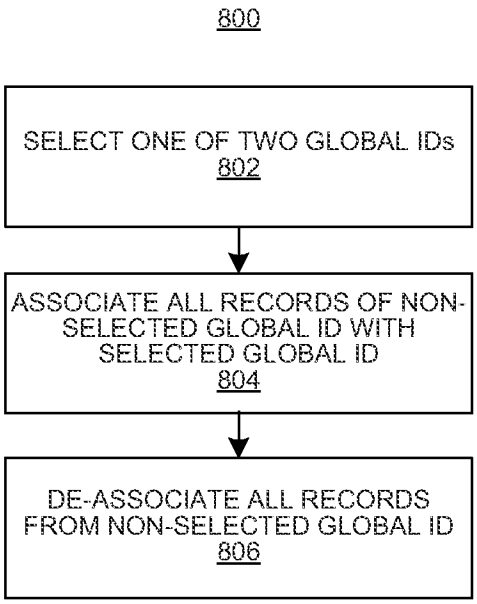


FIG. 8

FIG. 9A

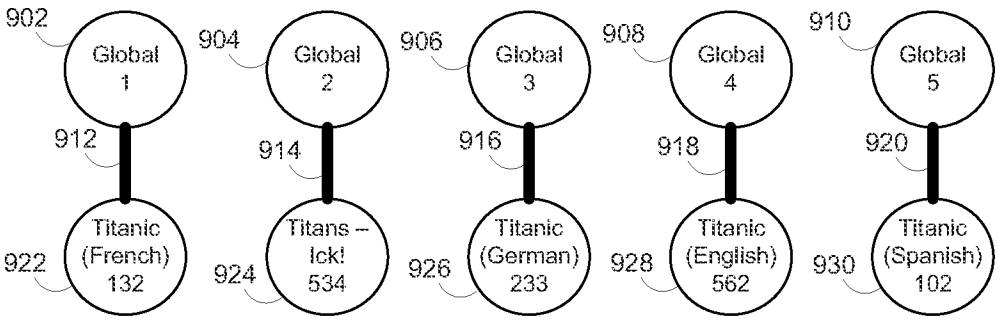


FIG. 9B

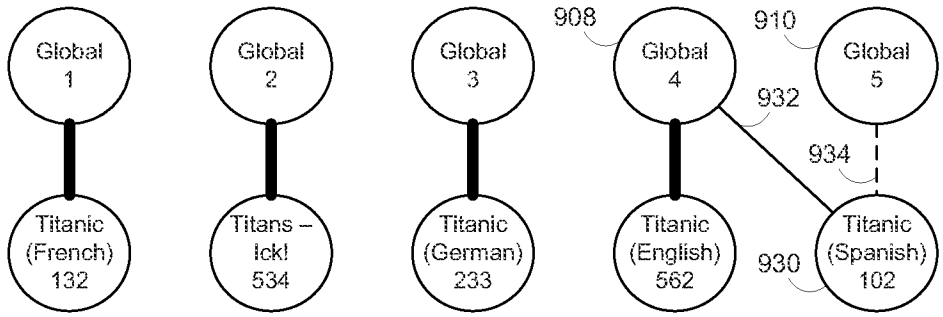


FIG. 9C

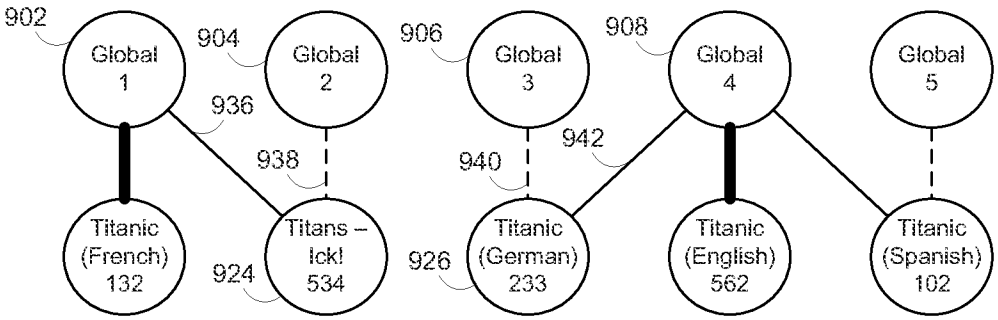
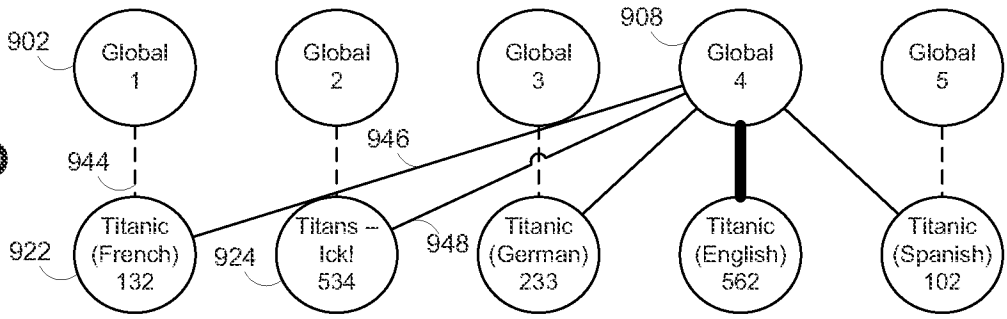


FIG. 9D



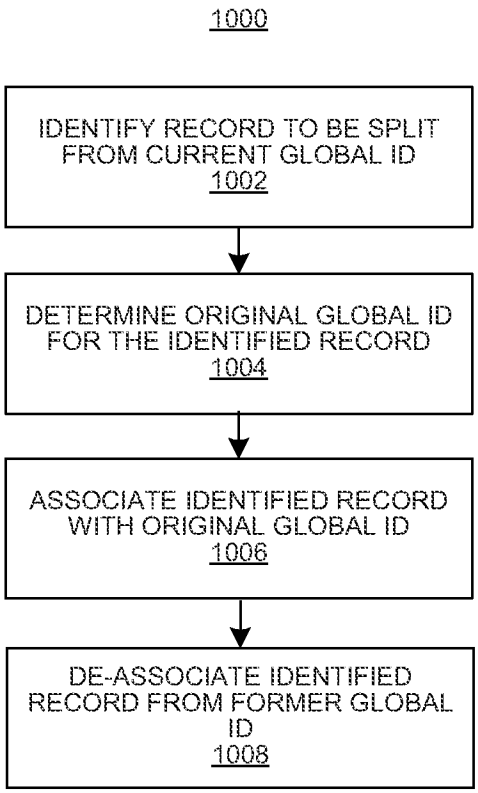


FIG. 10

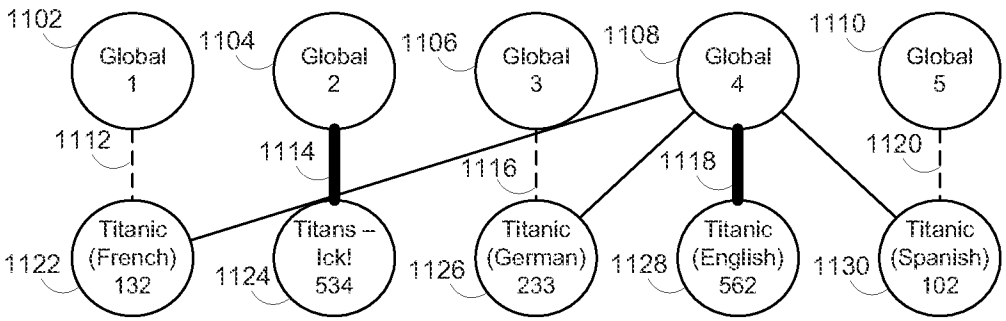


FIG. 11A

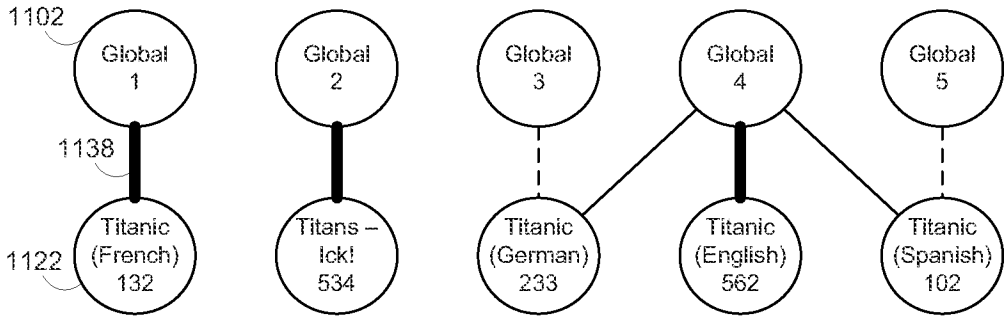


FIG. 11B

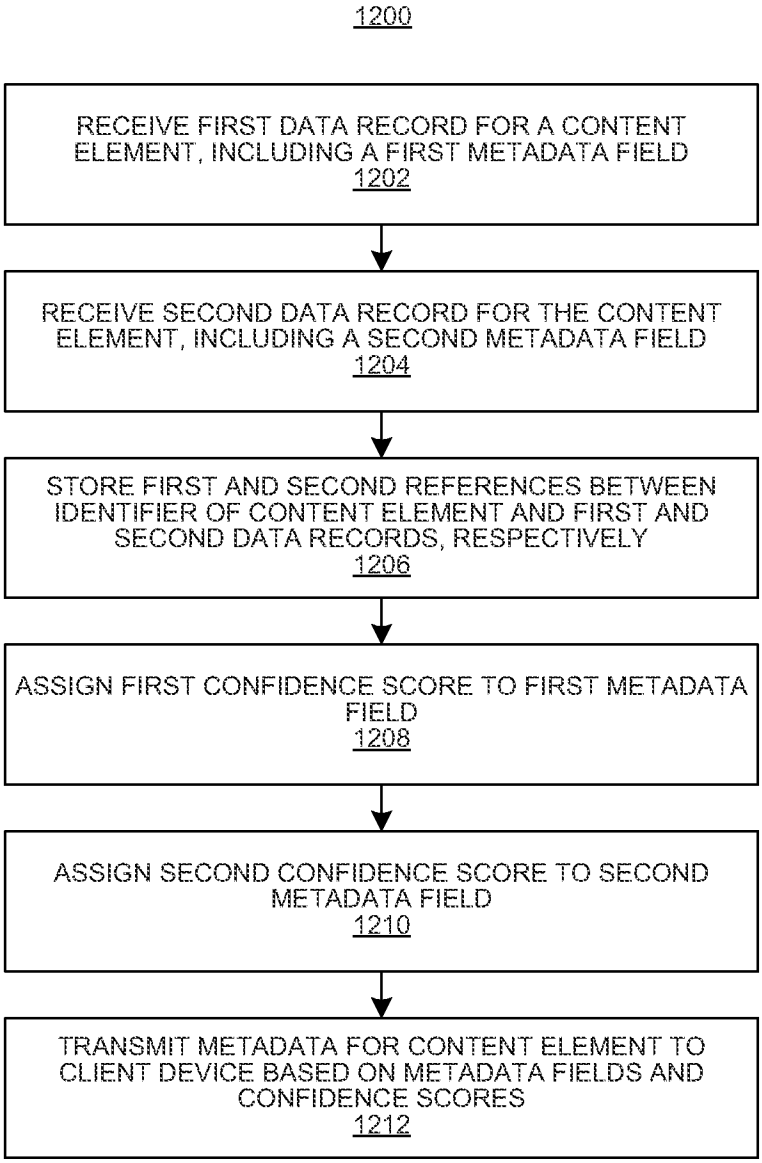

1200
RECEIVE FIRST DATA RECORD FOR A CONTENT ELEMENT, INCLUDING A FIRST METADATA FIELD
1202
RECEIVE SECOND DATA RECORD FOR THE CONTENT ELEMENT, INCLUDING A SECOND METADATA FIELD
1204
STORE FIRST AND SECOND REFERENCES BETWEEN IDENTIFIER OF CONTENT ELEMENT AND FIRST AND SECOND DATA RECORDS, RESPECTIVELY
1206
ASSIGN FIRST CONFIDENCE SCORE TO FIRST METADATA FIELD
1208
ASSIGN SECOND CONFIDENCE SCORE TO SECOND METADATA FIELD
1210
TRANSMIT METADATA FOR CONTENT ELEMENT TO CLIENT DEVICE BASED ON METADATA FIELDS AND CONFIDENCE SCORES
1212

FIG. 12

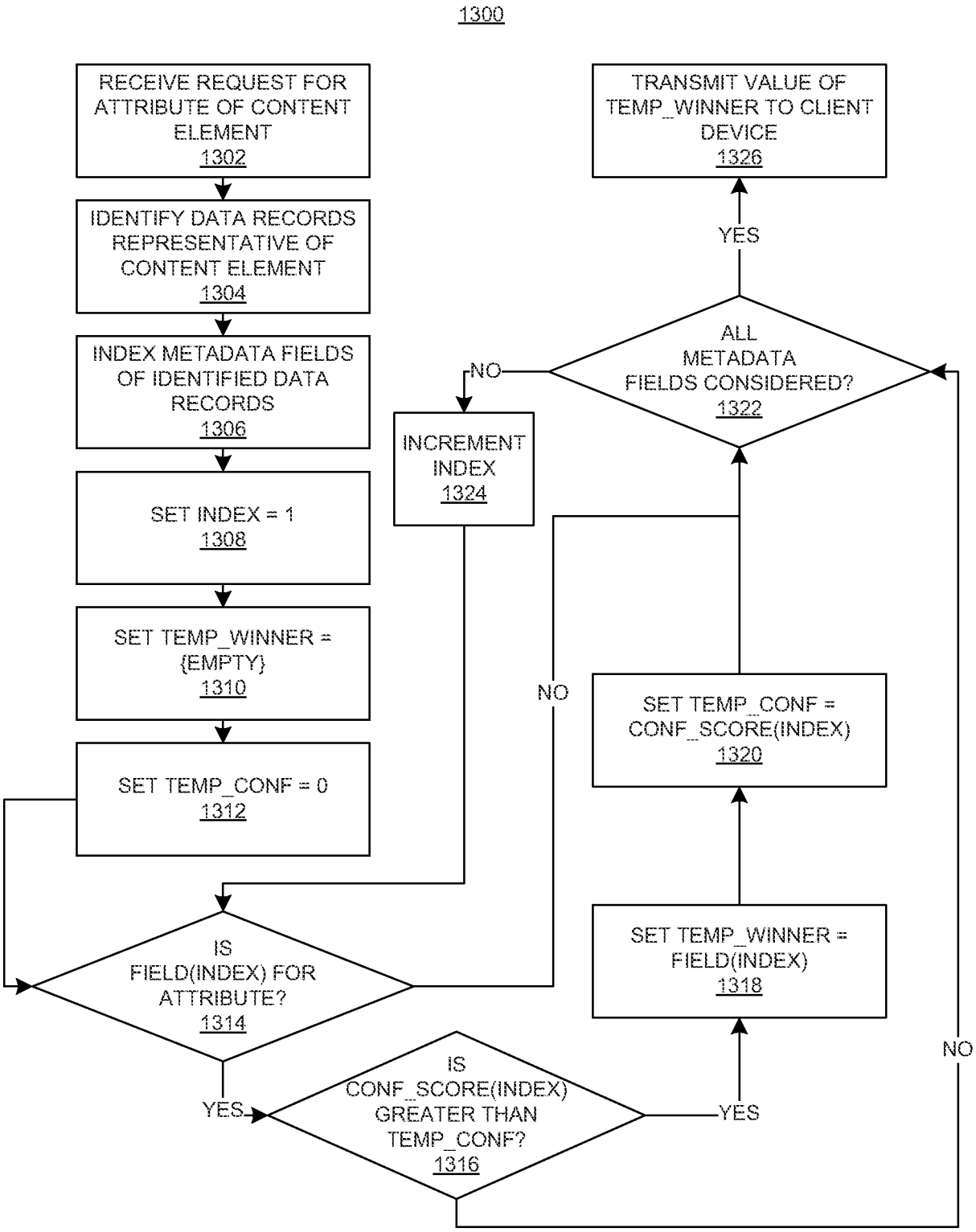

1300
RECEIVE REQUEST FOR ATTRIBUTE OF CONTENT ELEMENT 1302
IDENTIFY DATA RECORDS REPRESENTATIVE OF CONTENT ELEMENT 1304
INDEX METADATA FIELDS OF IDENTIFIED DATA RECORDS 1306
SET INDEX = 1 1308
SET TEMP_WINNER = {EMPTY} 1310
SET TEMP_CONF = 0 1312
IS FIELD(INDEX) FOR ATTRIBUTE? 1314
YES
IS CONF_SCORE(INDEX) GREATER THAN TEMP_CONF? 1316
YES
SET TEMP_WINNER = FIELD(INDEX) 1318
SET TEMP_CONF = CONF_SCORE(INDEX) 1320
ALL METADATA FIELDS CONSIDERED? 1322
NO
INCREMENT INDEX 1324
NO
NO
YES
TRANSMIT VALUE OF TEMP_WINNER TO CLIENT DEVICE 1326

FIG. 13

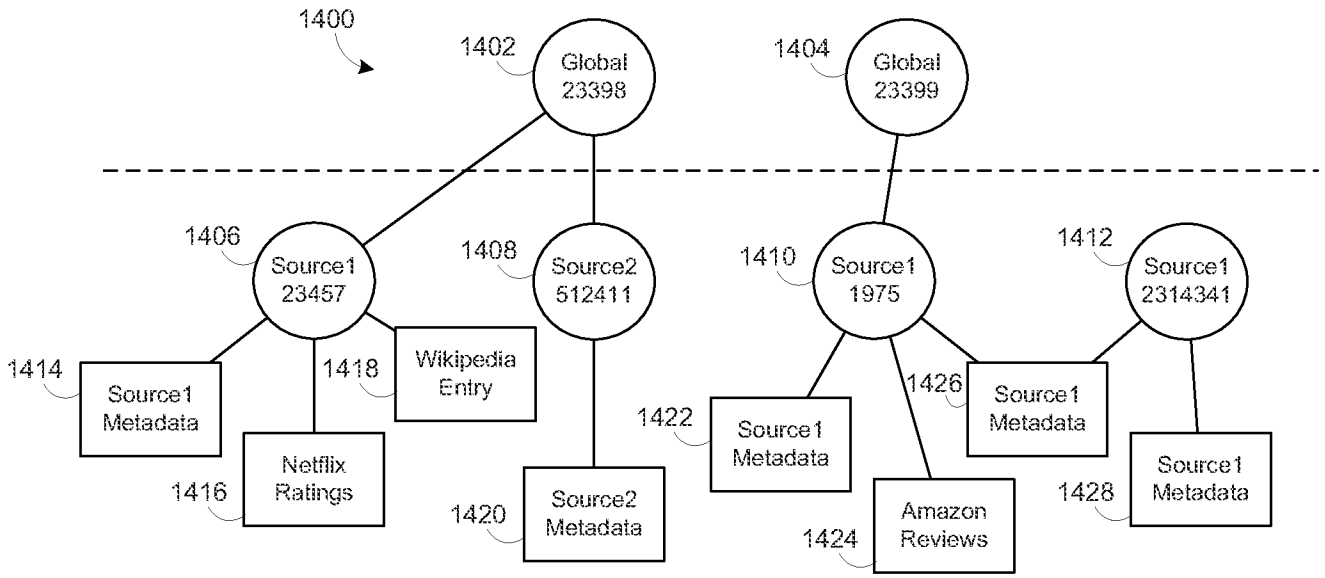


FIG. 14A

1430

1442 Top	1432 Field	1434 Record	1436 Source	1438 Value	1440 Confidence
	Name	Source1 512411	Source1 Metadata	Crash	0.54
	Year	Source1 512411	Source1 Metadata	1996	0.55
	Runtime	Source1 512411	Source1 Metadata	112 minutes	0.74
*	Name	Source2 23457	Source2 Metadata	Crash	0.95
*	Year	Source2 23457	Source2 Metadata	2004	0.93
*	MPAA	Source2 23457	Source2 Metadata	R	0.97
*	Plot	Source2 23457	Source2 Metadata	Issues of race and gender cause a group of strangers...	0.90
	Name	Source2 23457	Wikipedia	Crash	0.94
*	Runtime	Source2 23457	Wikipedia	111 minutes	0.86
	Plot	Source2 23457	Wikipedia	Set in Los Angeles, the film opens...	0.87
*	Rating	Source2 23457	Netflix Ratings	4.73	0.63

FIG. 14B

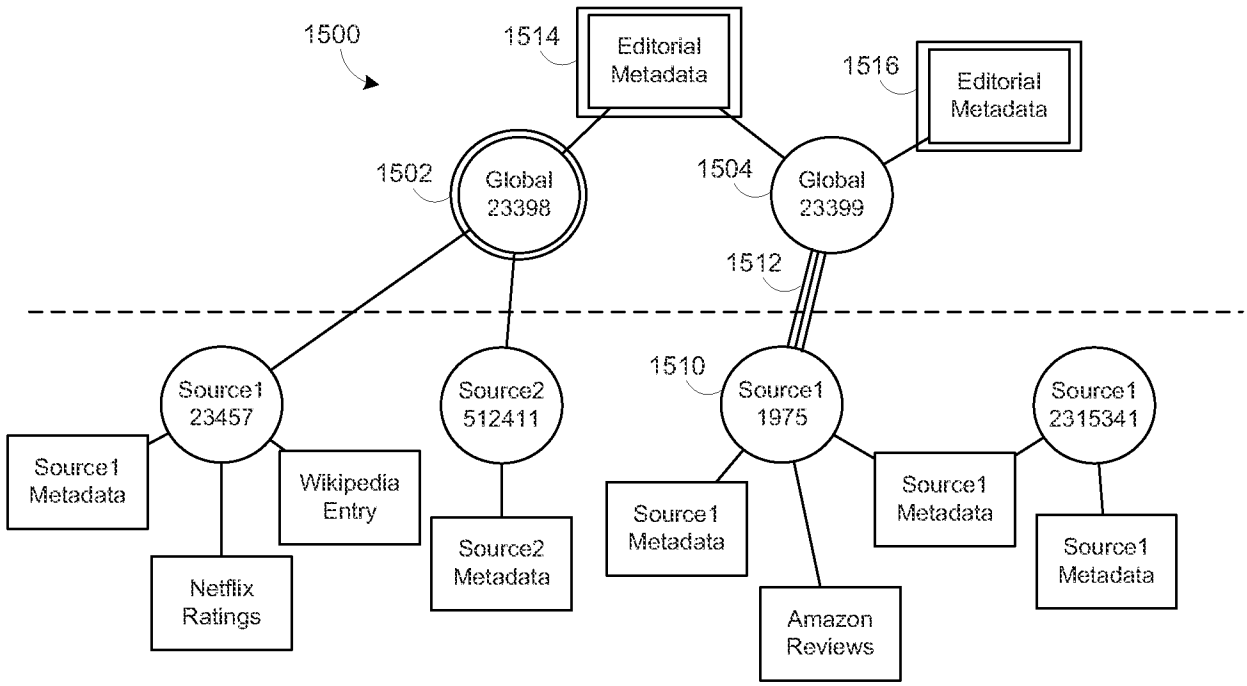

1500
1514
Editorial Metadata
1516
Editorial Metadata
1502
Global 23398
1504
Global 23399
1512
Source1 23457
Source2 512411
1510
Source1 1975
Source1 2315341
Source1 Metadata
Wikipedia Entry
Netflix Ratings
Source2 Metadata
Source1 Metadata
Amazon Reviews
Source1 Metadata
Source1 Metadata

FIG. 15

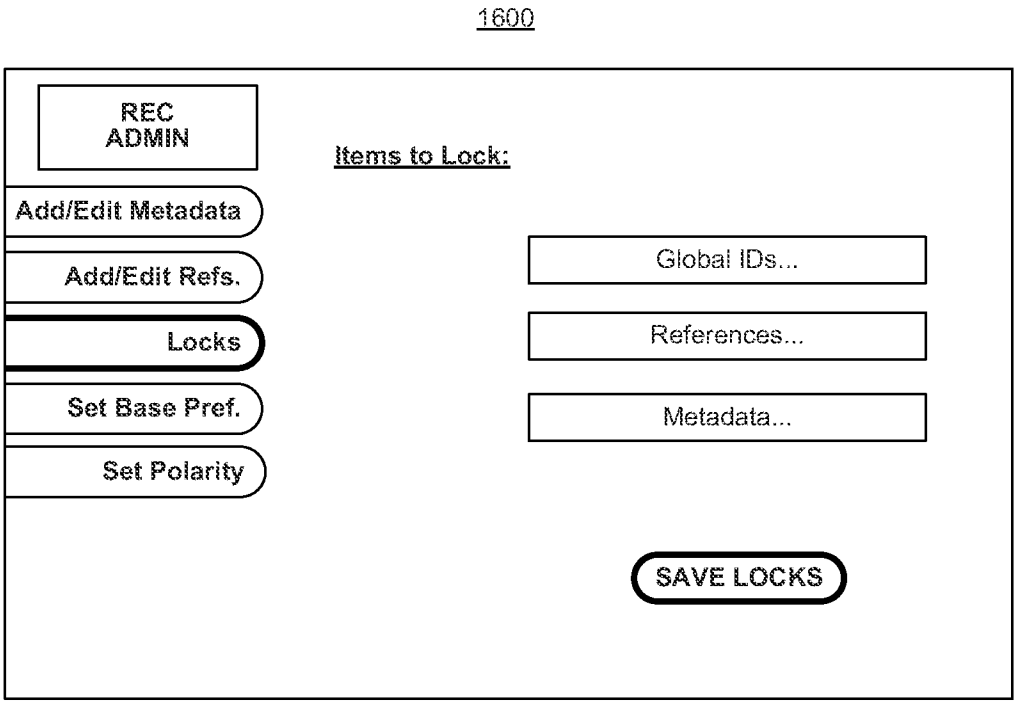

1600
REC ADMIN
Items to Lock:
Add/Edit Metadata
Add/Edit Refs.
Locks
Set Base Pref.
Set Polarity
Global IDs...
References...
Metadata...
SAVE LOCKS

FIG. 16

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2012/026544

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F17/30
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, INSPEC, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 1 349 080 A1 (THOMSON BRANDT GMBH [DE]) 1 October 2003 (2003-10-01) paragraph [0020] - paragraph [0032] -----	1-30

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
4 June 2012	13/06/2012

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Konak, Eyüp

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2012/026544

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 1349080 A1	01-10-2003	AU 2003212355 A1	08-10-2003
		CN 1643514 A	20-07-2005
		EP 1349080 A1	01-10-2003
		EP 1488346 A1	22-12-2004
		JP 2005521309 A	14-07-2005
		US 2005165816 A1	28-07-2005
		WO 03081459 A1	02-10-2003