

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
20 October 2005 (20.10.2005)

PCT

(10) International Publication Number
WO 2005/098598 A2

(51) International Patent Classification ⁷ :	G06F 7/00	60/602,956	18 August 2004 (18.08.2004)	US	
(21) International Application Number:	PCT/US2005/011015	60/602,897	18 August 2004 (18.08.2004)	US	
(22) International Filing Date:	1 April 2005 (01.04.2005)	60/602,925	18 August 2004 (18.08.2004)	US	
(25) Filing Language:	English	60/602,930	18 August 2004 (18.08.2004)	US	
(26) Publication Language:	English	60/602,896	18 August 2004 (18.08.2004)	US	
(30) Priority Data:		60/602,947	18 August 2004 (18.08.2004)	US	
60/558,527	1 April 2004 (01.04.2004)	US	60/602,898	18 August 2004 (18.08.2004)	US
60/559,265	1 April 2004 (01.04.2004)	US	60/603,466	19 August 2004 (19.08.2004)	US
60/558,893	1 April 2004 (01.04.2004)	US	60/603,082	19 August 2004 (19.08.2004)	US
60/558,499	1 April 2004 (01.04.2004)	US	60/603,081	19 August 2004 (19.08.2004)	US
60/559,279	1 April 2004 (01.04.2004)	US	60/603,498	20 August 2004 (20.08.2004)	US
60/558,760	1 April 2004 (01.04.2004)	US	60/603,358	20 August 2004 (20.08.2004)	US
60/558,789	1 April 2004 (01.04.2004)	US	60/604,103	23 August 2004 (23.08.2004)	US
60/558,867	1 April 2004 (01.04.2004)	US	60/604,100	23 August 2004 (23.08.2004)	US
60/558,969	1 April 2004 (01.04.2004)	US	60/604,102	23 August 2004 (23.08.2004)	US
60/558,717	1 April 2004 (01.04.2004)	US	60/604,098	23 August 2004 (23.08.2004)	US
60/558,791	1 April 2004 (01.04.2004)	US	60/605,105	27 August 2004 (27.08.2004)	US
60/559,277	1 April 2004 (01.04.2004)	US	60/605,229	27 August 2004 (27.08.2004)	US
60/559,278	1 April 2004 (01.04.2004)	US	60/613,589	27 September 2004 (27.09.2004)	US
60/559,226	1 April 2004 (01.04.2004)	US	60/613,454	27 September 2004 (27.09.2004)	US
60/558,370	1 April 2004 (01.04.2004)	US	60/613,634	27 September 2004 (27.09.2004)	US
60/558,968	1 April 2004 (01.04.2004)	US	60/613,400	27 September 2004 (27.09.2004)	US
60/558,892	1 April 2004 (01.04.2004)	US	60/613,633	27 September 2004 (27.09.2004)	US
60/559,125	2 April 2004 (02.04.2004)	US	60/613,455	27 September 2004 (27.09.2004)	US
60/559,087	2 April 2004 (02.04.2004)	US	60/613,341	27 September 2004 (27.09.2004)	US
60/559,033	2 April 2004 (02.04.2004)	US	60/613,602	27 September 2004 (27.09.2004)	US
60/559,127	2 April 2004 (02.04.2004)	US	60/613,460	27 September 2004 (27.09.2004)	US
60/559,131	2 April 2004 (02.04.2004)	US	60/613,361	27 September 2004 (27.09.2004)	US
60/558,909	2 April 2004 (02.04.2004)	US	60/613,461	27 September 2004 (27.09.2004)	US
60/559,766	6 April 2004 (06.04.2004)	US	60/613,339	27 September 2004 (27.09.2004)	US
60/561,768	12 April 2004 (12.04.2004)	US	60/613,456	27 September 2004 (27.09.2004)	US
60/563,520	19 April 2004 (19.04.2004)	US	60/613,242	27 September 2004 (27.09.2004)	US
60/563,485	19 April 2004 (19.04.2004)	US	60/613,340	27 September 2004 (27.09.2004)	US
60/564,846	23 April 2004 (23.04.2004)	US	60/613,628	27 September 2004 (27.09.2004)	US
60/564,688	23 April 2004 (23.04.2004)	US	60/613,243	27 September 2004 (27.09.2004)	US
60/566,667	30 April 2004 (30.04.2004)	US	60/613,632	27 September 2004 (27.09.2004)	US
60/571,381	14 May 2004 (14.05.2004)	US	60/615,378	1 October 2004 (01.10.2004)	US
60/571,560	14 May 2004 (14.05.2004)	US	60/615,112	1 October 2004 (01.10.2004)	US
60/571,715	17 May 2004 (17.05.2004)	US	60/615,538	1 October 2004 (01.10.2004)	US
60/589,203	19 July 2004 (19.07.2004)	US	60/617,122	7 October 2004 (07.10.2004)	US
60/589,201	19 July 2004 (19.07.2004)	US	60/622,906	28 October 2004 (28.10.2004)	US
60/589,202	19 July 2004 (19.07.2004)	US	11/004,637	3 December 2004 (03.12.2004)	US
60/598,821	2 August 2004 (02.08.2004)	US	60/633,453	6 December 2004 (06.12.2004)	US
			60/633,678	6 December 2004 (06.12.2004)	US
			60/633,452	6 December 2004 (06.12.2004)	US
			60/633,486	6 December 2004 (06.12.2004)	US
			60/634,739	9 December 2004 (09.12.2004)	US

[Continued on next page]

(54) Title: PUBLISHING TECHNIQUES FOR ADDING VALUE TO A RENDERED DOCUMENT

(57) Abstract: A system for processing a visual capture operation as described. The system receives an indication of a visual capture operation performed from a rendered document. The indication specifies both a text sequence capture as part of the capture operation and a supplemental marking captured as part of the capture operation. The system determines an action to perform in response to receiving the indication, based both upon the text sequence specified in the indication and the supplemental markings specified by the indication.

WO 2005/098598 A2



60/634,627	9 December 2004 (09.12.2004)	US
60/647,684	26 January 2005 (26.01.2005)	US
60/648,746	31 January 2005 (31.01.2005)	US
60/653,372	15 February 2005 (15.02.2005)	US
60/653,679	16 February 2005 (16.02.2005)	US
60/653,669	16 February 2005 (16.02.2005)	US
60/653,847	16 February 2005 (16.02.2005)	US
60/653,663	16 February 2005 (16.02.2005)	US
60/653,899	16 February 2005 (16.02.2005)	US
60/654,379	17 February 2005 (17.02.2005)	US
60/654,326	18 February 2005 (18.02.2005)	US
60/654,196	18 February 2005 (18.02.2005)	US
60/654,368	18 February 2005 (18.02.2005)	US
60/655,697	22 February 2005 (22.02.2005)	US
60/655,280	22 February 2005 (22.02.2005)	US
60/655,281	22 February 2005 (22.02.2005)	US
60/655,279	22 February 2005 (22.02.2005)	US
60/655,987	22 February 2005 (22.02.2005)	US
60/657,309	28 February 2005 (28.02.2005)	US

(71) **Applicant** (for all designated States except US): **EXBIBLIO** [NL/US]; 216 1st Avenue South, Suite 400, Seattle, wa 98104 (US).

(72) **Inventors; and**

(75) **Inventors/Applicants** (for US only): **KING, Martin T.** [US/US]; 17322 115th Ave. SW, Vashon Island, wa 98070 (US). **KUSHLER, Clifford A.** [US/US]; 20615 36th Place West, Lynnwood, wa 98036-5094 (US). **STAFFORD-FRASER, James Quentin** [GB/GB]; 20 Marlowe Road, Cambridge Cambridgeshire CB3 9JW (GB). **GROVER, Dale Lawrence** [US/US]; 511 Little Lake Drive, Ann Arbor, MI 48103 (US).

(74) **Agents:** **LAWRENZ, Steven D.** et al.; Perkins Coie LLP, Patents-SEA, P.O. Box 1247, Seattle, WA 98111-1247 (US).

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

(84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

PUBLISHING TECHNIQUES FOR ADDING VALUE TO A
RENDERED DOCUMENT

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application is a Continuation-In-Part of U.S. Patent Application No. 11/004,637 filed on December 3, 2004, which is hereby incorporated by reference in its entirety.

[0002] This application is related to, and incorporates by reference in their entirety, the following U.S. Patent Applications, filed concurrently herewith: U.S. Patent Application No. _____, entitled METHODS AND SYSTEMS FOR INITIATING APPLICATION PROCESSES BY DATA CAPTURE FROM RENDERED DOCUMENTS (Attorney Docket No. 435188104US1), U.S. Patent Application No. _____, entitled DETERMINING ACTIONS INVOLVING CAPTURED INFORMATION AND ELECTRONIC CONTENT ASSOCIATED WITH RENDERED DOCUMENTS (Attorney Docket No. 435188075US1), U.S. Patent Application No. _____, entitled CONTENT ACCESS WITH HANDHELD DOCUMENT DATA CAPTURE DEVICES (Attorney Docket No. 435188018US1), U.S. Patent Application No. _____, entitled SEARCH ENGINES AND SYSTEMS WITH HANDHELD DOCUMENT DATA CAPTURE DEVICES (Attorney Docket No. 435188021US1), U.S. Patent Application No. _____, entitled TRIGGERING ACTIONS IN RESPONSE TO OPTICALLY OR ACOUSTICALLY CAPTURING KEYWORDS FROM A RENDERED DOCUMENT (Attorney Docket No. 435188003US1), U.S. Patent Application No. _____, entitled SEARCHING AND ACCESSING DOCUMENTS ON PRIVATE NETWORKS FOR USE WITH CAPTURES FROM RENDERED DOCUMENTS (Attorney Docket No. 435188001US1), U.S. Patent Application No. _____, entitled INFORMATION GATHERING SYSTEM AND METHOD (Attorney Docket No. 435188005US1), U.S. Patent Application No. _____, entitled DOCUMENT ENHANCEMENT SYSTEM AND METHOD (Attorney Docket No. 43518.8009US1), U.S. Patent Application No. _____, entitled

ARCHIVE OF TEXT CAPTURES FROM RENDERED DOCUMENTS (Attorney Docket No. 435188116US), U.S. Patent Application No. _____, entitled ADDING INFORMATION OR FUNCTIONALITY TO A RENDERED DOCUMENT VIA ASSOCIATION WITH AN ELECTRONIC COUNTERPART (Attorney Docket No. 435188118US), U.S. Patent Application No. _____, entitled AGGREGATE ANALYSIS OF TEXT CAPTURES PERFORMED BY MULTIPLE USERS FROM RENDERED DOCUMENTS (Attorney Docket No. 435188122US), U.S. Patent Application No. _____, entitled ESTABLISHING AN INTERACTIVE ENVIRONMENT FOR RENDERED DOCUMENTS (Attorney Docket No. 435188123US), U.S. Patent Application No. _____, entitled DATA CAPTURE FROM RENDERED DOCUMENTS USING HANDHELD DEVICE (Attorney Docket No. 435188117US), and U.S. Patent Application No. _____, entitled CAPTURING TEXT FROM RENDERED DOCUMENTS USING SUPPLEMENTAL INFORMATION (Attorney Docket No. 435188120US).

[0003] This application claims priority to, and incorporates by reference in their entirety, the following U.S. Provisional Patent Applications: Application No. 60/559,226 filed on April 1, 2004, Application No. 60/558,893 filed on April 1, 2004, Application No. 60/558,968 filed on April 1, 2004, Application No. 60/558,867 filed on April 1, 2004, Application No. 60/559,278 filed on April 1, 2004, Application No. 60/559,279 filed on April 1, 2004, Application No. 60/559,265 filed on April 1, 2004, Application No. 60/559,277 filed on April 1, 2004, Application No. 60/558,969 filed on April 1, 2004, Application No. 60/558,892 filed on April 1, 2004, Application No. 60/558,760 filed on April 1, 2004, Application No. 60/558,717 filed on April 1, 2004, Application No. 60/558,499 filed on April 1, 2004, Application No. 60/558,370 filed on April 1, 2004, Application No. 60/558,789 filed on April 1, 2004, Application No. 60/558,791 filed on April 1, 2004, Application No. 60/558,527 filed on April 1, 2004, Application No. 60/559,125 filed on April 2, 2004, Application No. 60/558,909 filed on April 2, 2004, Application No. 60/559,033 filed on April 2, 2004, Application No. 60/559,127 filed on April 2, 2004, Application No. 60/559,087 filed on April 2, 2004, Application No. 60/559,131 filed on April 2,

2004, Application No. 60/559,766 filed on April 6, 2004, Application No. 60/561,768 filed on April 12, 2004, Application No. 60/563,520 filed on April 19, 2004, Application No. 60/563,485 filed on April 19, 2004, Application No. 60/564,688 filed on April 23, 2004, Application No. 60/564,846 filed on April 23, 2004, Application No. 60/556,667 filed on April 30, 2004, Application No. 60/571,381 filed on May 14, 2004, Application No. 60/571,560 filed on May 14, 2004, Application No. 60/571,715 filed on May 17, 2004, Application No. 60/589,203 filed on July 19, 2004, Application No. 60/589,201 filed on July 19, 2004, Application No. 60/589,202 filed on July 19, 2004, Application No. 60/598,821 filed on August 2, 2004, Application No. 60/602,956 filed on August 18, 2004, Application No. 60/602,925 filed on August 18, 2004, Application No. 60/602,947 filed on August 18, 2004, Application No. 60/602,897 filed on August 18, 2004, Application No. 60/602,896 filed on August 18, 2004, Application No. 60/602,930 filed on August 18, 2004, Application No. 60/602,898 filed on August 18, 2004, Application No. 60/603,466 filed on August 19, 2004, Application No. 60/603,082 filed on August 19, 2004, Application No. 60/603,081 filed on August 19, 2004, Application No. 60/603,498 filed on August 20, 2004, Application No. 60/603,358 filed on August 20, 2004, Application No. 60/604,103 filed on August 23, 2004, Application No. 60/604,098 filed on August 23, 2004, Application No. 60/604,100 filed on August 23, 2004, Application No. 60/604,102 filed on August 23, 2004, Application No. 60/605,229 filed on August 27, 2004, Application No. 60/605,105 filed on August 27, 2004, Application No. 60/613,243 filed on September 27, 2004, Application No. 60/613,628 filed on September 27, 2004, Application No. 60/613,632 filed on September 27, 2004, Application No. 60/613,589 filed on September 27, 2004, Application No. 60/613,242 filed on September 27, 2004, Application No. 60/613,602 filed on September 27, 2004, Application No. 60/613,340 filed on September 27, 2004, Application No. 60/613,634 filed on September 27, 2004, Application No. 60/613,461 filed on September 27, 2004, Application No. 60/613,455 filed on September 27, 2004, Application No. 60/613,460 filed on September 27, 2004, Application No. 60/613,400 filed on September 27, 2004, Application No. 60/613,456 filed on September 27, 2004, Application

No. 60/613,341 filed on September 27, 2004, Application No. 60/613,361 filed on September 27, 2004, Application No. 60/613,454 filed on September 27, 2004, Application No. 60/613,339 filed on September 27, 2004, Application No. 60/613,633 filed on September 27, 2004, Application No. 60/615,378 filed on October 1, 2004, Application No. 60/615,112 filed on October 1, 2004, Application No. 60/615,538 filed on October 1, 2004, Application No. 60/617,122 filed on October 7, 2004, Application No. 60/622,906 filed on October 28, 2004, Application No. 60/633,452 filed on December 6, 2004, Application No. 60/633,678 filed on December 6, 2004, Application No. 60/633,486 filed on December 6, 2004, Application No. 60/633,453 filed on December 6, 2004, Application No. 60/634,627 filed on December 9, 2004, Application No. 60/634,739 filed on December 9, 2004, Application No. 60/647,684 filed on January 26, 2005, Application No. 60/648,746 filed on January 31, 2005, Application No. 60/653,372 filed on February 15, 2005, Application No. 60/653,663 filed on February 16, 2005, Application No. 60/653,669 filed on February 16, 2005, Application No. 60/653,899 filed on February 16, 2005, Application No. 60/653,679 filed on February 16, 2005, Application No. 60/653,847 filed on February 16, 2005, Application No. 60/654,379 filed on February 17, 2005, Application No. 60/654,368 filed on February 18, 2005, Application No. 60/654,326 filed on February 18, 2005, Application No. 60/654,196 filed on February 18, 2005, Application No. 60/655,279 filed on February 22, 2005, Application No. 60/655,280 filed on February 22, 2005, Application No. 60/655,987 filed on February 22, 2005, Application No. 60/655,697 filed on February 22, 2005, Application No. 60/655,281 filed on February 22, 2005, and Application No. 60/657,309 filed on February 28, 2005.

TECHNICAL FIELD

[0004] The described technology is directed to the field of document processing.

BACKGROUND

[0005] The concept of interactive paper leads to new models of publishing books, documents, and periodicals. Any printed or rendered document can have associated supplemental materials. These materials may be considered an interactive environment that can be accessed by a user of the described system. In recent years many books have been published with an accompanying CD-ROM or associated web page or website, but for the rendered document and the associated digital materials to be truly mutually beneficial, they need to be more closely coupled than has been possible in the past.

[0006] As an example, some books currently have an associated web site on which errata are listed after the book has been published. The process for the reader of discovering errata associated with the particular page he is currently reading, however, is usually inconvenient. It involves going to a computer, turning it on, starting a web browser, typing in the URL of the book's web site, selecting the errata page, and checking the notes to see what applies to the page in question. The user is hardly likely to do this for every page.

[0007] In light of these limitations of the prior art, a system which provides the ability to create for a rendered document a richly interactive electronic environment (whether CD-ROM-based, web-based, or otherwise), and to allow a reader of the rendered document to easily move back and forth between a location in the rendered document and associated materials would have particular utility.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] Figure 1 is a data flow diagram that illustrates the flow of information in one embodiment of the core system.

[0009] Figure 2 is a component diagram of components included in a typical implementation of the system in the context of a typical operating environment.

[0010] Figure 3 is a block diagram of an embodiment of a scanner.

[0011] Figure 4 illustrates the interaction between a user, paper document, and the system in one embodiment.

[0012] Figure 5 is a flow diagram showing steps typically performed by the system to perform an action in response to a text capture.

[0013] Figure 6 is a flow diagram illustrating the steps taken by the system to find appropriate markup for a document in one embodiment.

[0014] Figure 7 is an example document illustrating several forms of supplemental markings used in some embodiments.

[0015] Figure 8 illustrates a printed menu recognized by the system in some embodiments.

[0016] Figure 9 illustrates supplemental markings using text styles and other system-readable information.

[0017] Figure 10 is a data structure used in some embodiments for storing the known-document registry.

DETAILED DESCRIPTION

OVERVIEW

[0018] In a world where a great deal of attention is being paid to the new capabilities of digital documents and electronic displays, printed documents continue to be popular and are produced in vast numbers. A system that provides a way for authors and publishers to add value to those printed documents by endowing them with an electronic life which is easily accessible to the reader ("the system") is described.

PART I – INTRODUCTION

1. Nature of the System

[0019] For every paper document that has an electronic counterpart, there exists a discrete amount of information in the paper document that can identify the electronic counterpart. In some embodiments, the system uses a sample of text captured from a paper document, for example using a handheld scanner, to identify and locate an electronic counterpart of the

document. In most cases, the amount of text needed by the facility is very small in that a few words of text from a document can often function as an identifier for the paper document and as a link to its electronic counterpart. In addition, the system may use those few words to identify not only the document, but also a location within the document.

[0020] Thus, paper documents and their digital counterparts can be associated in many useful ways using the system discussed herein.

1.1. A quick overview of the future

[0021] Once the system has associated a piece of text in a paper document with a particular digital entity has been established, the system is able to build a huge amount of functionality on that association.

[0022] It is increasingly the case that most paper documents have an electronic counterpart that is accessible on the World Wide Web or from some other online database or document corpus, or can be made accessible, such as in response to the payment of a fee or subscription. At the simplest level, then, when a user scans a few words in a paper document, the system can retrieve that electronic document or some part of it, or display it, email it to somebody, purchase it, print it or post it to a web page. As additional examples, scanning a few words of a book that a person is reading over breakfast could cause the audio-book version in the person's car to begin reading from that point when s/he starts driving to work, or scanning the serial number on a printer cartridge could begin the process of ordering a replacement.

[0023] The system implements these and many other examples of "paper/digital integration" without requiring changes to the current processes of writing, printing and publishing documents, giving such conventional rendered documents a whole new layer of digital functionality.

1.2. Terminology

[0024] A typical use of the system begins with using an optical scanner to scan text from a paper document, but it is important to note that other methods of capture from other types of document are equally applicable. The

system is therefore sometimes described as *scanning* or *capturing* text from a *rendered* document, where those terms are defined as follows:

[0025] A *rendered document* is a printed document or a document shown on a display or monitor. It is a document that is perceptible to a human, whether in permanent form or on a transitory display.

[0026] *Scanning or capturing* is the process of systematic examination to obtain information from a rendered document. The process may involve optical capture using a scanner or camera (for example a camera in a cellphone), or it may involve reading aloud from the document into an audio capture device or typing it on a keypad or keyboard. For more examples, see Section 15.

2. Introduction to the System

[0027] This section describes some of the devices, processes and systems that constitute a system for paper/digital integration. In various embodiments, the system builds a wide variety of services and applications on this underlying core that provides the basic functionality.

2.1. The processes

[0028] Figure 1 is a data flow diagram that illustrates the flow of information in one embodiment of the core system. Other embodiments may not use all of the stages or elements illustrated here, while some will use many more.

[0029] Text from a rendered document is captured 100, typically in optical form by an optical scanner or audio form by a voice recorder, and this image or sound data is then processed 102, for example to remove artifacts of the capture process or to improve the signal-to-noise ratio. A recognition process 104 such as OCR, speech recognition, or autocorrelation then converts the data into a signature, comprised in some embodiments of text, text offsets, or other symbols. Alternatively, the system performs an alternate form of extracting document signature from the rendered document. The signature represents a set of possible text transcriptions in some embodiments. This process may be influenced by feedback from other

stages, for example, if the search process and context analysis 110 have identified some candidate documents from which the capture may originate, thus narrowing the possible interpretations of the original capture.

[0030] A post-processing 106 stage may take the output of the recognition process and filter it or perform such other operations upon it as may be useful. Depending upon the embodiment implemented, it may be possible at this stage to deduce some direct actions 107 to be taken immediately without reference to the later stages, such as where a phrase or symbol has been captured which contains sufficient information in itself to convey the user's intent. In these cases no digital counterpart document need be referenced, or even known to the system.

[0031] Typically, however, the next stage will be to construct a query 108 or a set of queries for use in searching. Some aspects of the query construction may depend on the search process used and so cannot be performed until the next stage, but there will typically be some operations, such as the removal of obviously misrecognized or irrelevant characters, which can be performed in advance.

[0032] The query or queries are then passed to the search and context analysis stage 110. Here, the system optionally attempts to identify the document from which the original data was captured. To do so, the system typically uses search indices and search engines 112, knowledge about the user 114 and knowledge about the user's context or the context in which the capture occurred 116. Search engine 112 may employ and/or index information specifically about rendered documents, about their digital counterpart documents, and about documents that have a web (internet) presence). It may write to, as well as read from, many of these sources and, as has been mentioned, it may feed information into other stages of the process, for example by giving the recognition system 104 information about the language, font, rendering and likely next words based on its knowledge of the candidate documents.

[0033] In some circumstances the next stage will be to retrieve 120 a copy of the document or documents that have been identified. The sources of

the documents 124 may be directly accessible, for example from a local filing system or database or a web server, or they may need to be contacted via some access service 122 which might enforce authentication, security or payment or may provide other services such as conversion of the document into a desired format.

[0034] Applications of the system may take advantage of the association of extra functionality or data with part or all of a document. For example, advertising applications discussed in Section 10.4 may use an association of particular advertising messages or subjects with portions of a document. This extra associated functionality or data can be thought of as one or more overlays on the document, and is referred to herein as "markup." The next stage of the process 130, then, is to identify any markup relevant to the captured data. Such markup may be provided by the user, the originator, or publisher of the document, or some other party, and may be directly accessible from some source 132 or may be generated by some service 134. In various embodiments, markup can be associated with, and apply to, a rendered document and/or the digital counterpart to a rendered document, or to groups of either or both of these documents.

[0035] Lastly, as a result of the earlier stages, some actions may be taken 140. These may be default actions such as simply recording the information found, they may be dependent on the data or document, or they may be derived from the markup analysis. Sometimes the action will simply be to pass the data to another system. In some cases the various possible actions appropriate to a capture at a specific point in a rendered document will be presented to the user as a menu on an associated display, for example on a local display 332, on a computer display 212 or a mobile phone or PDA display 216. If the user doesn't respond to the menu, the default actions can be taken.

2.2. The components

[0036] Figure 2 is a component diagram of components included in a typical implementation of the system in the context of a typical operating environment. As illustrated, the operating environment includes one or more

optical scanning capture devices 202 or voice capture devices 204. In some embodiments, the same device performs both functions. Each capture device is able to communicate with other parts of the system such as a computer 212 and a mobile station 216 (e.g., a mobile phone or PDA) using either a direct wired or wireless connection, or through the network 220, with which it can communicate using a wired or wireless connection, the latter typically involving a wireless base station 214. In some embodiments, the capture device is integrated in the mobile station, and optionally shares some of the audio and/or optical components used in the device for voice communications and picture-taking.

[0037] Computer 212 may include a memory containing computer executable instructions for processing an order from scanning devices 202 and 204. As an example, an order can include an identifier (such as a serial number of the scanning device 202/204 or an identifier that partially or uniquely identifies the user of the scanner), scanning context information (e.g., time of scan, location of scan, etc.) and/or scanned information (such as a text string) that is used to uniquely identify the document being scanned. In alternative embodiments, the operating environment may include more or less components.

[0038] Also available on the network 220 are search engines 232, document sources 234, user account services 236, markup services 238 and other network services 239. The network 220 may be a corporate intranet, the public Internet, a mobile phone network or some other network, or any interconnection of the above.

[0039] Regardless of the manner by which the devices are coupled to each other, they may all may be operable in accordance with well-known commercial transaction and communication protocols (e.g., Internet Protocol (IP)). In various embodiments, the functions and capabilities of scanning device 202, computer 212, and mobile station 216 may be wholly or partially integrated into one device. Thus, the terms scanning device, computer, and mobile station can refer to the same device depending upon whether the device incorporates functions or capabilities of the scanning device 202,

computer 212 and mobile station 216. In addition, some or all of the functions of the search engines 232, document sources 234, user account services 236, markup services 238 and other network services 239 may be implemented on any of the devices and/or other devices not shown.

2.3. The capture device

[0040] As described above, the capture device may capture text using an optical scanner that captures image data from the rendered document, or using an audio recording device that captures a user's spoken reading of the text, or other methods. Some embodiments of the capture device may also capture images, graphical symbols and icons, etc., including machine readable codes such as barcodes. The device may be exceedingly simple, consisting of little more than the transducer, some storage, and a data interface, relying on other functionality residing elsewhere in the system, or it may be a more full-featured device. For illustration, this section describes a device based around an optical scanner and with a reasonable number of features.

[0041] Scanners are well known devices that capture and digitize images. An offshoot of the photocopier industry, the first scanners were relatively large devices that captured an entire document page at once. Recently, portable optical scanners have been introduced in convenient form factors, such as a pen-shaped handheld device.

[0042] In some embodiments, the portable scanner is used to scan text, graphics, or symbols from rendered documents. The portable scanner has a scanning element that captures text, symbols, graphics, etc, from rendered documents. In addition to documents that have been printed on paper, in some embodiments, rendered documents include documents that have been displayed on a screen such as a CRT monitor or LCD display.

[0043] Figure 3 is a block diagram of an embodiment of a scanner 302. The scanner 302 comprises an optical scanning head 308 to scan information from rendered documents and convert it to machine-compatible data, and an optical path 306, typically a lens, an aperture or an image conduit to convey the image from the rendered document to the scanning head. The scanning

head 308 may incorporate a Charge-Coupled Device (CCD), a Complementary Metal Oxide Semiconductor (CMOS) imaging device, or an optical sensor of another type.

[0044] A microphone 310 and associated circuitry convert the sound of the environment (including spoken words) into machine-compatible signals, and other input facilities exist in the form of buttons, scroll-wheels or other tactile sensors such as touch-pads 314.

[0045] Feedback to the user is possible through a visual display or indicator lights 332, through a loudspeaker or other audio transducer 334 and through a vibrate module 336.

[0046] The scanner 302 comprises logic 326 to interact with the various other components, possibly processing the received signals into different formats and/or interpretations. Logic 326 may be operable to read and write data and program instructions stored in associated storage 330 such as RAM, ROM, flash, or other suitable memory. It may read a time signal from the clock unit 328. The scanner 302 also includes an interface 316 to communicate scanned information and other signals to a network and/or an associated computing device. In some embodiments, the scanner 302 may have an on-board power supply 332. In other embodiments, the scanner 302 may be powered from a tethered connection to another device, such as a Universal Serial Bus (USB) connection.

[0047] As an example of one use of scanner 302, a reader may scan some text from a newspaper article with scanner 302. The text is scanned as a bit-mapped image via the scanning head 308. Logic 326 causes the bit-mapped image to be stored in memory 330 with an associated time-stamp read from the clock unit 328. Logic 326 may also perform optical character recognition (OCR) or other post-scan processing on the bit-mapped image to convert it to text. Logic 326 may optionally extract a signature from the image, for example by performing a convolution-like process to locate repeating occurrences of characters, symbols or objects, and determine the distance or number of other characters, symbols, or objects between these repeated elements. The reader may then upload the bit-mapped image (or

text or other signature, if post-scan processing has been performed by logic 326) to an associated computer via interface 316.

[0048] As an example of another use of scanner 302, a reader may capture some text from an article as an audio file by using microphone 310 as an acoustic capture port. Logic 326 causes audio file to be stored in memory 328. Logic 326 may also perform voice recognition or other post-scan processing on the audio file to convert it to text. As above, the reader may then upload the audio file (or text produced by post-scan processing performed by logic 326) to an associated computer via interface 316.

PART II – OVERVIEW OF THE AREAS OF THE CORE SYSTEM

[0049] As paper-digital integration becomes more common, there are many aspects of existing technologies that can be changed to take better advantage of this integration, or to enable it to be implemented more effectively. This section highlights some of those issues.

3. Search

[0050] Searching a corpus of documents, even so large a corpus as the World Wide Web, has become commonplace for ordinary users, who use a keyboard to construct a search query which is sent to a search engine. This section and the next discuss the aspects of both the construction of a query originated by a capture from a rendered document, and the search engine that handles such a query.

3.1. Scan/Speak/Type as search query

[0051] Use of the described system typically starts with a few words being captured from a rendered document using any of several methods, including those mentioned in Section 1.2 above. Where the input needs some interpretation to convert it to text, for example in the case of OCR or speech input, there may be end-to-end feedback in the system so that the document corpus can be used to enhance the recognition process. End-to-end feedback can be applied by performing an approximation of the recognition or interpretation, identifying a set of one or more candidate matching documents, and then using information from the possible matches in the candidate

documents to further refine or restrict the recognition or interpretation. Candidate documents can be weighted according to their probable relevance (for example, based on the number of other users who have scanned in these documents, or their popularity on the Internet), and these weights can be applied in this iterative recognition process.

3.2. Short Phrase Searching

[0052] Because the selective power of a search query based on a few words is greatly enhanced when the relative positions of these words are known, only a small amount of text need be captured for the system to identify the text's location in a corpus. Most commonly, the input text will be a contiguous sequence of words, such as a short phrase.

3.2.1. Finding document and location in document from short capture

[0053] In addition to locating the document from which a phrase originates, the system can identify the location in that document and can take action based on this knowledge.

3.2.2. Other methods of finding location

[0054] The system may also employ other methods of discovering the document and location, such as by using watermarks or other special markings on the rendered document.

3.3. Incorporation of other factors in search query

[0055] In addition to the captured text, other factors (i.e., information about user identity, profile, and context) may form part of the search query, such as the time of the capture, the identity and geographical location of the user, knowledge of the user's habits and recent activities, etc.

[0056] The document identity and other information related to previous captures, especially if they were quite recent, may form part of a search query.

[0057] The identity of the user may be determined from a unique identifier associated with a capturing device, and/or biometric or other supplemental information (speech patterns, fingerprints, etc.).

3.4. Knowledge of nature of unreliability in search query (OCR errors etc)

[0058] The search query can be constructed taking into account the types of errors likely to occur in the particular capture method used. One example of this is an indication of suspected errors in the recognition of specific characters; in this instance a search engine may treat these characters as wildcards, or assign them a lower priority.

3.5. Local caching of index for performance/offline use

[0059] Sometimes the capturing device may not be in communication with the search engine or corpus at the time of the data capture. For this reason, information helpful to the offline use of the device may be downloaded to the device in advance, or to some entity with which the device can communicate. In some cases, all or a substantial part of an index associated with a corpus may be downloaded. This topic is discussed further in Section 15.3.

3.6. Queries, in whatever form, may be recorded and acted on later

[0060] If there are likely to be delays or cost associated with communicating a query or receiving the results, this pre-loaded information can improve the performance of the local device, reduce communication costs, and provide helpful and timely user feedback.

[0061] In the situation where no communication is available (the local device is "offline"), the queries may be saved and transmitted to the rest of the system at such a time as communication is restored.

[0062] In these cases it may be important to transmit a timestamp with each query. The time of the capture can be a significant factor in the interpretation of the query. For example, Section 13.1 discusses the importance of the time of capture in relation to earlier captures. It is important to note that the time of capture will not always be the same as the time that the query is executed.

3.7. Parallel searching

[0063] For performance reasons, multiple queries may be launched in response to a single capture, either in sequence or in parallel. Several queries may be sent in response to a single capture, for example as new words are added to the capture, or to query multiple search engines in parallel.

[0064] For example, in some embodiments, the system sends queries to a special index for the current document, to a search engine on a local machine, to a search engine on the corporate network, and to remote search engines on the Internet.

[0065] The results of particular searches may be given higher priority than those from others.

[0066] The response to a given query may indicate that other pending queries are superfluous; these may be cancelled before completion.

4. Paper and Search Engines

[0067] Often it is desirable for a search engine that handles traditional online queries also to handle those originating from rendered documents. Conventional search engines may be enhanced or modified in a number of ways to make them more suitable for use with the described system.

[0068] The search engine and/or other components of the system may create and maintain indices that have different or extra features. The system may modify an incoming paper-originated query or change the way the query is handled in the resulting search, thus distinguishing these paper-originated queries from those coming from queries typed into web browsers and other sources. And the system may take different actions or offer different options when the results are returned by the searches originated from paper as compared to those from other sources. Each of these approaches is discussed below.

4.1. Indexing

[0069] Often, the same index can be searched using either paper-originated or traditional queries, but the index may be enhanced for use in the current system in a variety of ways.

4.1.1. Knowledge about the paper form

[0070] Extra fields can be added to such an index that will help in the case of a paper-based search.

Index entry indicating document availability in paper form

[0071] The first example is a field indicating that the document is known to exist or be distributed in paper form. The system may give such documents higher priority if the query comes from paper.

Knowledge of popularity paper form

[0072] In this example statistical data concerning the popularity of paper documents (and, optionally, concerning sub-regions within these documents) – for example the amount of scanning activity, circulation numbers provided by the publisher or other sources, etc – is used to give such documents higher priority, to boost the priority of digital counterpart documents (for example, for browser-based queries or web searches), etc.

Knowledge of rendered format

[0073] Another important example may be recording information about the layout of a specific rendering of a document.

[0074] For a particular edition of a book, for example, the index may include information about where the line breaks and page breaks occur, which fonts were used, any unusual capitalization.

[0075] The index may also include information about the proximity of other items on the page, such as images, text boxes, tables and advertisements.

Use of semantic information in original

[0076] Lastly, semantic information that can be deduced from the source markup but is not apparent in the paper document, such as the fact that a

particular piece of text refers to an item offered for sale, or that a certain paragraph contains program code, may also be recorded in the index.

4.1.2. Indexing in the knowledge of the capture method

[0077] A second factor that may modify the nature of the index is the knowledge of the type of capture likely to be used. A search initiated by an optical scan may benefit if the index takes into account characters that are easily confused in the OCR process, or includes some knowledge of the fonts used in the document. Similarly, if the query is from speech recognition, an index based on similar-sounding phonemes may be much more efficiently searched. An additional factor that may affect the use of the index in the described model is the importance of iterative feedback during the recognition process. If the search engine is able to provide feedback from the index as the text is being captured, it can greatly increase the accuracy of the capture.

Indexing using offsets

[0078] If the index is likely to be searched using the offset-based/autocorrelation OCR methods described in Section 9, in some embodiments, the system stores the appropriate offset or signature information in an index.

4.1.3. Multiple indices

[0079] Lastly, in the described system, it may be common to conduct searches on many indices. Indices may be maintained on several machines on a corporate network. Partial indices may be downloaded to the capture device, or to a machine close to the capture device. Separate indices may be created for users or groups of users with particular interests, habits or permissions. An index may exist for each filesystem, each directory, even each file on a user's hard disk. Indexes are published and subscribed to by users and by systems. It will be important, then, to construct indices that can be distributed, updated, merged and separated efficiently.

4.2. Handling the queries

4.2.1. Knowing the capture is from paper

[0080] A search engine may take different actions when it recognizes that a search query originated from a paper document. The engine might handle the query in a way that is more tolerant to the types of errors likely to appear in certain capture methods, for example.

[0081] It may be able to deduce this from some indicator included in the query (for example a flag indicating the nature of the capture), or it may deduce this from the query itself (for example, it may recognize errors or uncertainties typical of the OCR process).

[0082] Alternatively, queries from a capture device can reach the engine by a different channel or port or type of connection than those from other sources, and can be distinguished in that way. For example, some embodiments of the system will route queries to the search engine by way of a dedicated gateway. Thus, the search engine knows that all queries passing through the dedicated gateway were originated from a paper document.

4.2.2. Use of context

[0083] Section 13 below describes a variety of different factors which are external to the captured text itself, yet which can be a significant aid in identifying a document. These include such things as the history of recent scans, the longer-term reading habits of a particular user, the geographic location of a user and the user's recent use of particular electronic documents. Such factors are referred to herein as "context."

[0084] Some of the context may be handled by the search engine itself, and be reflected in the search results. For example, the search engine may keep track of a user's scanning history, and may also cross-reference this scanning history to conventional keyboard-based queries. In such cases, the search engine maintains and uses more state information about each individual user than do most conventional search engines, and each interaction with a search engine may be considered to extend over several searches and a longer period of time than is typical today.

[0085] Some of the context may be transmitted to the search engine in the search query (Section 3.3), and may possibly be stored at the engine so as to play a part in future queries. Lastly, some of the context will best be handled elsewhere, and so becomes a filter or secondary search applied to the results from the search engine.

Data-stream input to search

[0086] An important input into the search process is the broader context of how the community of users is interacting with the rendered version of the document – for example, which documents are most widely read and by whom. There are analogies with a web search returning the pages that are most frequently linked to, or those that are most frequently selected from past search results. For further discussion of this topic, see Sections 13.4 and 14.2.

4.2.3. Document sub-regions

[0087] The described system can emit and use not only information about documents as a whole, but also information about sub-regions of documents, even down to individual words. Many existing search engines concentrate simply on locating a document or file that is relevant to a particular query. Those that can work on a finer grain and identify a location *within* a document will provide a significant benefit for the described system.

4.3. Returning the results

[0088] The search engine may use some of the further information it now maintains to affect the results returned.

[0089] The system may also return certain documents to which the user has access only as a result of being in possession of the paper copy (Section 7.4).

[0090] The search engine may also offer new actions or options appropriate to the described system, beyond simple retrieval of the text.

5. Markup, Annotations and Metadata

[0091] In addition to performing the capture-search-retrieve process, the described system also associates extra functionality with a document, and in

particular with specific locations or segments of text within a document. This extra functionality is often, though not exclusively, associated with the rendered document by being associated with its electronic counterpart. As an example, hyperlinks in a web page could have the same functionality when a printout of that web page is scanned. In some cases, the functionality is not defined in the electronic document, but is stored or generated elsewhere.

[0092] This layer of added functionality is referred to herein as "markup."

5.1. Overlays, static and dynamic

[0093] One way to think of the markup is as an "overlay" on the document, which provides further information about—and may specify actions associated with—the document or some portion of it. The markup may include human-readable content, but is often invisible to a user and/or intended for machine use. Examples include options to be displayed in a popup-menu on a nearby display when a user captures text from a particular area in a rendered document, or audio samples that illustrate the pronunciation of a particular phrase.

5.1.1. Several layers, possibly from several sources

[0094] Any document may have multiple overlays simultaneously, and these may be sourced from a variety of locations. Markup data may be created or supplied by the author of the document, or by the user, or by some other party.

[0095] Markup data may be attached to the electronic document or embedded in it. It may be found in a conventional location (for example, in the same place as the document but with a different filename suffix). Markup data may be included in the search results of the query that located the original document, or may be found by a separate query to the same or another search engine. Markup data may be found using the original captured text and other capture information or contextual information, or it may be found using already-deduced information about the document and location of the capture. Markup data may be found in a location specified in the document, even if the markup itself is not included in the document.

[0096] The markup may be largely static and specific to the document, similar to the way links on a traditional html web page are often embedded as static data within the html document, but markup may also be dynamically generated and/or applied to a large number of documents. An example of dynamic markup is information attached to a document that includes the up-to-date share price of companies mentioned in that document. An example of broadly applied markup is translation information that is automatically available on multiple documents or sections of documents in a particular language.

5.1.2. Personal "plug-in" layers

[0097] Users may also install, or subscribe to particular sources of, markup data, thus personalizing the system's response to particular captures.

5.2. Keywords and phrases, trademarks and logos

[0098] Some elements in documents may have particular "markup" or functionality associated with them based on their own characteristics rather than their location in a particular document. Examples include special marks that are printed in the document purely for the purpose of being scanned, as well as logos and trademarks that can link the user to further information about the organization concerned. The same applies to "keywords" or "key phrases" in the text. Organizations might register particular phrases with which they are associated, or with which they would like to be associated, and attach certain markup to them that would be available wherever that phrase was scanned.

[0099] Any word, phrase, etc. may have associated markup. For example, the system may add certain items to a pop-up menu (e.g., a link to an online bookstore) whenever the user captures the word "book," or the title of a book, or a topic related to books. In some embodiments, of the system, digital counterpart documents or indices are consulted to determine whether a capture occurred near the word "book," or the title of a book, or a topic related to books – and the system behavior is modified in accordance with this proximity to keyword elements. In the preceding example, note that markup

enables data captured from non-commercial text or documents to trigger a commercial transaction.

5.3. User-supplied content

5.3.1. User comments and annotations, including multimedia

[00100] Annotations are another type of electronic information that may be associated with a document. For example, a user can attach an audio file of his/her thoughts about a particular document for later retrieval as voice annotations. As another example of a multimedia annotation, a user may attach photographs of places referred to in the document. The user generally supplies annotations for the document but the system can associate annotations from other sources (for example, other users in a work group may share annotations).

5.3.2. Notes from proof-reading

[00101] An important example of user-sourced markup is the annotation of paper documents as part of a proofreading, editing or reviewing process.

5.4. Third-party content

[00102] As mentioned earlier, markup data may often be supplied by third parties, such as by other readers of the document. Online discussions and reviews are a good example, as are community-managed information relating to particular works, volunteer-contributed translations and explanations.

[00103] Another example of third-party markup is that provided by advertisers.

5.5. Dynamic markup based on other users' data streams

[00104] By analyzing the data captured from documents by several or all users of the system, markup can be generated based on the activities and interests of a community. An example might be an online bookstore that creates markup or annotations that tell the user, in effect, "People who enjoyed this book also enjoyed....." The markup may be less anonymous, and may tell the user which of the people in his/her contact list have also read this document recently. Other examples of datastream analysis are included in Section 14.

5.6. Markup based on external events and data sources

[00105] Markup will often be based on external events and data sources, such as input from a corporate database, information from the public Internet, or statistics gathered by the local operating system.

[00106] Data sources may also be more local, and in particular may provide information about the user's context – his/her identity, location and activities. For example, the system might communicate with the user's mobile phone and offer a markup layer that gives the user the option to send a document to somebody that the user has recently spoken to on the phone.

6. Authentication, Personalization and Security

[00107] In many situations, the identity of the user will be known. Sometimes this will be an "anonymous identity," where the user is identified only by the serial number of the capture device, for example. Typically, however, it is expected that the system will have a much more detailed knowledge of the user, which can be used for personalizing the system and to allow activities and transactions to be performed in the user's name.

6.1. User history and "Life Library"

[00108] One of the simplest and yet most useful functions that the system can perform is to keep a record for a user of the text that s/he has captured and any further information related to that capture, including the details of any documents found, the location within that document and any actions taken as a result.

[00109] This stored history is beneficial for both the user and the system.

6.1.1. For the user

[00110] The user can be presented with a "Life Library," a record of everything s/he has read and captured. This may be simply for personal interest, but may be used, for example, in a library by an academic who is gathering material for the bibliography of his next paper.

[00111] In some circumstances, the user may wish to make the library public, such as by publishing it on the web in a similar manner to a weblog, so that others may see what s/he is reading and finds of interest.

[00112] Lastly, in situations where the user captures some text and the system cannot immediately act upon the capture (for example, because an electronic version of the document is not yet available) the capture can be stored in the library and can be processed later, either automatically or in response to a user request. A user can also subscribe to new markup services and apply them to previously captured scans.

6.1.2. For the system

[00113] A record of a user's past captures is also useful for the system. Many aspects of the system operation can be enhanced by knowing the user's reading habits and history. The simplest example is that any scan made by a user is more likely to come from a document that the user has scanned in the recent past, and in particular if the previous scan was within the last few minutes it is very likely to be from the same document. Similarly, it is more likely that a document is being read in start-to-finish order. Thus, for English documents, it is also more likely that later scans will occur farther down in the document. Such factors can help the system establish the location of the capture in cases of ambiguity, and can also reduce the amount of text that needs to be captured.

6.2. Scanner as payment, identity and authentication device

[00114] Because the capture process generally begins with a device of some sort, typically an optical scanner or voice recorder, this device may be used as a key that identifies the user and authorizes certain actions.

6.2.1. Associate scanner with phone or other account

[00115] The device may be embedded in a mobile phone or in some other way associated with a mobile phone account. For example, a scanner may be associated with a mobile phone account by inserting a SIM card associated with the account into the scanner. Similarly, the device may be embedded in a credit card or other payment card, or have the facility for such a card to be connected to it. The device may therefore be used as a payment token, and financial transactions may be initiated by the capture from the rendered document.

6.2.2. Using scanner input for authentication

[00116] The scanner may also be associated with a particular user or account through the process of scanning some token, symbol or text associated with that user or account. In addition, scanner may be used for biometric identification, for example by scanning the fingerprint of the user. In the case of an audio-based capture device, the system may identify the user by matching the voice pattern of the user or by requiring the user to speak a certain password or phrase.

[00117] For example, where a user scans a quote from a book and is offered the option to buy the book from an online retailer, the user can select this option, and is then prompted to scan his/her fingerprint to confirm the transaction.

[00118] See also Sections 15.5 and 15.6.

6.2.3. Secure Scanning Device

[00119] When the capture device is used to identify and authenticate the user, and to initiate transactions on behalf of the user, it is important that communications between the device and other parts of the system are secure. It is also important to guard against such situations as another device impersonating a scanner, and so-called "man in the middle" attacks where communications between the device and other components are intercepted.

[00120] Techniques for providing such security are well understood in the art; in various embodiments, the hardware and software in the device and elsewhere in the system are configured to implement such techniques.

7. Publishing Models and Elements

[00121] An advantage of the described system is that there is no need to alter the traditional processes of creating, printing or publishing documents in order to gain many of the system's benefits. There are reasons, though, that the creators or publishers of a document – hereafter simply referred to as the "publishers" – may wish to create functionality to support the described system.

[00122] This section is primarily concerned with the published documents themselves. For information about other related commercial transactions, such as advertising, see Section 10 entitled "P-Commerce."

7.1. Electronic companions to printed documents

[00123] The system allows for printed documents to have an associated electronic presence. Conventionally publishers often ship a CD-ROM with a book that contains further digital information, tutorial movies and other multimedia data, sample code or documents, or further reference materials. In addition, some publishers maintain web sites associated with particular publications which provide such materials, as well as information which may be updated after the time of publishing, such as errata, further comments, updated reference materials, bibliographies and further sources of relevant data, and translations into other languages. Online forums allow readers to contribute their comments about the publication.

[00124] The described system allows such materials to be much more closely tied to the rendered document than ever before, and allows the discovery of and interaction with them to be much easier for the user. By capturing a portion of text from the document, the system can automatically connect the user to digital materials associated with the document, and more particularly associated with that specific part of the document. Similarly, the user can be connected to online communities that discuss that section of the text, or to annotations and commentaries by other readers. In the past, such information would typically need to be found by searching for a particular page number or chapter.

[00125] An example application of this is in the area of academic textbooks (Section 17.5).

7.2. "Subscriptions" to printed documents

[00126] Some publishers may have mailing lists to which readers can subscribe if they wish to be notified of new relevant matter or when a new edition of the book is published. With the described system, the user can register an interest in particular documents or parts of documents more easily, in some cases even before the publisher has considered providing any such

functionality. The reader's interest can be fed to the publisher, possibly affecting their decision about when and where to provide updates, further information, new editions or even completely new publications on topics that have proved to be of interest in existing books.

7.3. Printed marks with special meaning or containing special data

[00127] Many aspects of the system are enabled simply through the use of the text already existing in a document. If the document is produced in the knowledge that it may be used in conjunction with the system, however, extra functionality can be added by printing extra information in the form of special marks, which may be used to identify the text or a required action more closely, or otherwise enhance the document's interaction with the system. The simplest and most important example is an indication to the reader that the document is definitely accessible through the system. A special icon might be used, for example, to indicate that this document has an online discussion forum associated with it.

[00128] Such symbols may be intended purely for the reader, or they may be recognized by the system when scanned and used to initiate some action. Sufficient data may be encoded in the symbol to identify more than just the symbol: it may also store information, for example about the document, edition, and location of the symbol, which could be recognized and read by the system.

7.4. Authorization through possession of the paper document

[00129] There are some situations where possession of or access to the printed document would entitle the user to certain privileges, for example, the access to an electronic copy of the document or to additional materials. With the described system, such privileges could be granted simply as a result of the user capturing portions of text from the document, or scanning specially printed symbols. In cases where the system needed to ensure that the user was in possession of the entire document, it might prompt the user to scan particular items or phrases from particular pages, e.g. "the second line of page 46."

7.5. Documents which expire

[00130] If the printed document is a gateway to extra materials and functionality, access to such features can also be time-limited. After the expiry date, a user may be required to pay a fee or obtain a newer version of the document to access the features again. The paper document will, of course, still be usable, but will lose some of its enhanced electronic functionality. This may be desirable, for example, because there is profit for the publisher in receiving fees for access to electronic materials, or in requiring the user to purchase new editions from time to time, or because there are disadvantages associated with outdated versions of the printed document remaining in circulation. Coupons are an example of a type of commercial document that can have an expiration date.

7.6. Popularity analysis and publishing decisions

[00131] Section 10.5 discusses the use of the system's statistics to influence compensation of authors and pricing of advertisements.

[00132] In some embodiments, the system deduces the popularity of a publication from the activity in the electronic community associated with it as well as from the use of the paper document. These factors may help publishers to make decisions about what they will publish in future. If a chapter in an existing book, for example, turns out to be exceedingly popular, it may be worth expanding into a separate publication.

8. Document Access Services

[00133] An important aspect of the described system is the ability to provide to a user who has access to a rendered copy of a document access to an electronic version of that document. In some cases, a document is freely available on a public network or a private network to which the user has access. The system uses the captured text to identify, locate and retrieve the document, in some cases displaying it on the user's screen or depositing it in their email inbox.

[00134] In some cases, a document will be available in electronic form, but for a variety of reasons may not be accessible to the user. There may not be sufficient connectivity to retrieve the document, the user may not be entitled to

retrieve it, there may be a cost associated with gaining access to it, or the document may have been withdrawn and possibly replaced by a new version, to name just a few possibilities. The system typically provides feedback to the user about these situations.

[00135] As mentioned in Section 7.4, the degree or nature of the access granted to a particular user may be different if it is known that the user already has access to a printed copy of the document.

8.1. Authenticated document access

[00136] Access to the document may be restricted to specific users, or to those meeting particular criteria, or may only be available in certain circumstances, for example when the user is connected to a secure network. Section 6 describes some of the ways in which the credentials of a user and scanner may be established.

8.2. Document purchase – copyright-owner compensation

[00137] Documents that are not freely available to the general public may still be accessible on payment of a fee, often as compensation to the publisher or copyright-holder. The system may implement payment facilities directly or may make use of other payment methods associated with the user, including those described in Section 6.2.

8.3. Document escrow and proactive retrieval

[00138] Electronic documents are often transient; the digital source version of a rendered document may be available now but inaccessible in future. The system may retrieve and store the existing version on behalf of the user, even if the user has not requested it, thus guaranteeing its availability should the user request it in future. This also makes it available for the system's use, for example for searching as part of the process of identifying future captures.

[00139] In the event that payment is required for access to the document, a trusted "document escrow" service can retrieve the document on behalf of the user, such as upon payment of a modest fee, with the assurance that the

copyright holder will be fully compensated in future if the user should ever request the document from the service.

[00140] Variations on this theme can be implemented if the document is not available in electronic form at the time of capture. The user can authorize the service to submit a request for or make a payment for the document on his/her behalf if the electronic document should become available at a later date.

8.4. Association with other subscriptions and accounts

[00141] Sometimes payment may be waived, reduced or satisfied based on the user's existing association with another account or subscription. Subscribers to the printed version of a newspaper might automatically be entitled to retrieve the electronic version, for example.

[00142] In other cases, the association may not be quite so direct: a user may be granted access based on an account established by their employer, or based on their scanning of a printed copy owned by a friend who is a subscriber.

8.5. Replacing photocopying with scan-and-print

[00143] The process of capturing text from a paper document, identifying an electronic original, and printing that original, or some portion of that original associated with the capture, forms an alternative to traditional photocopying with many advantages:

[00144] • the paper document need not be in the same location as the final printout, and in any case need not be there at the same time

[00145] • the wear and damage caused to documents by the photocopying process, especially to old, fragile and valuable documents, can be avoided

[00146] • the quality of the copy is typically be much higher

[00147] • records may be kept about which documents or portions of documents are the most frequently copied

[00148] • payment may be made to the copyright owner as part of the process

[00149] • unauthorized copying may be prohibited

8.6. Locating valuable originals from photocopies

[00150] When documents are particularly valuable, as in the case of legal instruments or documents that have historical or other particular significance, people may typically work from copies of those documents, often for many years, while the originals are kept in a safe location.

[00151] The described system could be coupled to a database which records the location of an original document, for example in an archiving warehouse, making it easy for somebody with access to a copy to locate the archived original paper document.

9. Text Recognition Technologies

[00152] Optical Character Recognition (OCR) technologies have traditionally focused on images that include a large amount of text, for example from a flat-bed scanner capturing a whole page. OCR technologies often need substantial training and correcting by the user to produce useful text. OCR technologies often require substantial processing power on the machine doing the OCR, and, while many systems use a dictionary, they are generally expected to operate on an effectively infinite vocabulary.

[00153] All of the above traditional characteristics may be improved upon in the described system.

[00154] While this section focuses on OCR, many of the issues discussed map directly onto other recognition technologies, in particular speech recognition. As mentioned in Section 3.1, the process of capturing from paper may be achieved by a user reading the text aloud into a device which captures audio. Those skilled in the art will appreciate that principles discussed here with respect to images, fonts, and text fragments often also apply to audio samples, user speech models and phonemes.

9.1. Optimization for appropriate devices

[00155] A scanning device for use with the described system will often be small, portable, and low power. The scanning device may capture only a few words at a time, and in some implementations does not even capture a whole character at once, but rather a horizontal slice through the text, many such slices being stitched together to form a recognizable signal from which the text may be deduced. The scanning device may also have very limited processing power or storage so, while in some embodiments it may perform all of the OCR process itself, many embodiments will depend on a connection to a more powerful device, possibly at a later time, to convert the captured signals into text. Lastly, it may have very limited facilities for user interaction, so may need to defer any requests for user input until later, or operate in a "best-guess" mode to a greater degree than is common now.

9.2. "Uncertain" OCR

[00156] The primary new characteristic of OCR within the described system is the fact that it will, in general, examine images of text which exists elsewhere and which may be retrieved in digital form. An exact transcription of the text is therefore not always required from the OCR engine. The OCR system may output a set or a matrix of possible matches, in some cases including probability weightings, which can still be used to search for the digital original.

9.3. Iterative OCR – guess, disambiguate, guess...

[00157] If the device performing the recognition is able to contact the document index at the time of processing, then the OCR process can be informed by the contents of the document corpus as it progresses, potentially offering substantially greater recognition accuracy.

[00158] Such a connection will also allow the device to inform the user when sufficient text has been captured to identify the digital source.

9.4. Using knowledge of likely rendering

[00159] When the system has knowledge of aspects of the likely printed rendering of a document—such as the font typeface used in printing, or the

layout of the page, or which sections are in italics—this too can help in the recognition process. (Section 4.1.1)

9.5. Font caching – determine font on host, download to client

[00160] As candidate source texts in the document corpus are identified, the font, or a rendering of it, may be downloaded to the device to help with the recognition.

9.6. Autocorrelation and character offsets

[00161] While component characters of a text fragment may be the most recognized way to represent a fragment of text that may be used as a document signature, other representations of the text may work sufficiently well that the actual text of a text fragment need not be used when attempting to locate the text fragment in a digital document and/or database, or when disambiguating the representation of a text fragment into a readable form. Other representations of text fragments may provide benefits that actual text representations lack. For example, optical character recognition of text fragments is often prone to errors, unlike other representations of captured text fragments that may be used to search for and/or recreate a text fragment without resorting to optical character recognition for the entire fragment. Such methods may be more appropriate for some devices used with the current system.

[00162] Those of ordinary skill in the art and others will appreciate that there are many ways of describing the appearance of text fragments. Such characterizations of text fragments may include, but are not limited to, word lengths, relative word lengths, character heights, character widths, character shapes, character frequencies, token frequencies, and the like. In some embodiments, the offsets between matching text tokens (i.e., the number of intervening tokens plus one) are used to characterize fragments of text.

[00163] Conventional OCR uses knowledge about fonts, letter structure and shape to attempt to determine characters in scanned text. Embodiments of the present invention are different; they employ a variety of methods that use the rendered text itself to assist in the recognition process. These embodiments use characters (or tokens) to "recognize each other." One way

to refer to such self-recognition is "template matching," and is similar to "convolution." To perform such self-recognition, the system slides a copy of the text horizontally over itself and notes matching regions of the text images. Prior template matching and convolution techniques encompass a variety of related techniques. These techniques to tokenize and/or recognize characters/tokens will be collectively referred to herein as "autocorrelation," as the text is used to correlate with its own component parts when matching characters/tokens.

[00164] When autocorrelating, complete connected regions that match are of interest. This occurs when characters (or groups of characters) overlay other instances of the same character (or group). Complete connected regions that match automatically provide tokenizing of the text into component tokens. As the two copies of the text are slid past each other, the regions where perfect matching occurs (i.e., all pixels in a vertical slice are matched) are noted. When a character/token matches itself, the horizontal extent of this matching (e.g., the connected matching portion of the text) also matches.

[00165] Note that at this stage there is no need to determine the actual identity of each token (i.e., the particular letter, digit or symbol, or group of these, that corresponds to the token image), only the offset to the next occurrence of the same token in the scanned text. The offset number is the distance (number of tokens) to the next occurrence of the same token. If the token is unique within the text string, the offset is zero (0). The sequence of token offsets thus generated is a signature that can be used to identify the scanned text.

[00166] In some embodiments, the token offsets determined for a string of scanned tokens are compared to an index that indexes a corpus of electronic documents based upon the token offsets of their contents (Section 4.1.2). In other embodiments, the token offsets determined for a string of scanned tokens are converted to text, and compared to a more conventional index that indexes a corpus of electronic documents based upon their contents

[00167] As has been noted earlier, a similar token-correlation process may be applied to speech fragments when the capture process consists of audio samples of spoken words.

9.7. Font/character "self-recognition"

[00168] Conventional template-matching OCR compares scanned images to a library of character images. In essence, the alphabet is stored for each font and newly scanned images are compared to the stored images to find matching characters. The process generally has an initial delay until the correct font has been identified. After that, the OCR process is relatively quick because most documents use the same font throughout. Subsequent images can therefore be converted to text by comparison with the most recently identified font library.

[00169] The shapes of characters in most commonly used fonts are related. For example, in most fonts, the letter "c" and the letter "e" are visually related – as are "t" and "f," etc. The OCR process is enhanced by use of this relationship to construct templates for letters that have not been scanned yet. For example, where a reader scans a short string of text from a paper document in a previously unencountered font such that the system does not have a set of image templates with which to compare the scanned images the system can leverage the probable relationship between certain characters to construct the font template library even though it has not yet encountered all of the letters in the alphabet. The system can then use the constructed font template library to recognize subsequent scanned text and to further refine the constructed font library.

9.8. Send anything unrecognized (including graphics) to server

[00170] When images cannot be machine-transcribed into a form suitable for use in a search process, the images themselves can be saved for later use by the user, for possible manual transcription, or for processing at a later date when different resources may be available to the system.

10. P-Commerce

[00171] Many of the actions made possible by the system result in some commercial transaction taking place. The phrase *p-commerce* is used herein to describe commercial activities initiated from paper via the system.

10.1. Sales of documents from their physical printed copies.

[00172] When a user captures text from a document, the user may be offered that document for purchase either in paper or electronic form. The user may also be offered related documents, such as those quoted or otherwise referred to in the paper document, or those on a similar subject, or those by the same author.

10.2. Sales of anything else initiated or aided by paper

[00173] The capture of text may be linked to other commercial activities in a variety of ways. The captured text may be in a catalog that is explicitly designed to sell items, in which case the text will be associated fairly directly with the purchase of an item (Section 18.2). The text may also be part of an advertisement, in which case a sale of the item being advertised may ensue.

[00174] In other cases, the user captures other text from which their potential interest in a commercial transaction may be deduced. A reader of a novel set in a particular country, for example, might be interested in a holiday there. Someone reading a review of a new car might be considering purchasing it. The user may capture a particular fragment of text knowing that some commercial opportunity will be presented to them as a result, or it may be a side-effect of their capture activities.

10.3. Capture of labels, icons, serial numbers, barcodes on an item resulting in a sale

[00175] Sometimes text or symbols are actually printed on an item or its packaging. An example is the serial number or product id often found on a label on the back or underside of a piece of electronic equipment. The system can offer the user a convenient way to purchase one or more of the same items by capturing that text. They may also be offered manuals, support or repair services.

10.4. Contextual Advertisements

[00176] In addition to the direct capture of text from an advertisement, the system allows for a new kind of advertising which is not necessarily explicitly in the rendered document, but is nonetheless based on what people are reading.

10.4.1. Advertising based on scan context and history

[00177] In a traditional paper publication, advertisements generally consume a large amount of space relative to the text of a newspaper article, and a limited number of them can be placed around a particular article. In the described system, advertising can be associated with individual words or phrases, and can be selected according to the particular interest the user has shown by capturing that text and possibly taking into account their history of past scans.

[00178] With the described system, it is possible for a purchase to be tied to a particular printed document and for an advertiser to get significantly more feedback about the effectiveness of their advertising in particular print publications.

10.4.2. Advertising based on user context and history

[00179] The system may gather a large amount of information about other aspects of a user's context for its own use (Section 13); estimates of the geographical location of the user are a good example. Such data can also be used to tailor the advertising presented to a user of the system.

10.5. Models of compensation

[00180] The system enables some new models of compensation for advertisers and marketers. The publisher of a printed document containing advertisements may receive some income from a purchase that originated from their document. This may be true whether or not the advertisement existed in the original printed form; it may have been added electronically either by the publisher, the advertiser or some third party, and the sources of such advertising may have been subscribed to by the user.

10.5.1. Popularity-based compensation

[00181] Analysis of the statistics generated by the system can reveal the popularity of certain parts of a publication (Section 14.2). In a newspaper, for example, it might reveal the amount of time readers spend looking at a particular page or article, or the popularity of a particular columnist. In some circumstances, it may be appropriate for an author or publisher to receive compensation based on the activities of the readers rather than on more traditional metrics such as words written or number of copies distributed. An author whose work becomes a frequently read authority on a subject might be considered differently in future contracts from one whose books have sold the same number of copies but are rarely opened. (See also Section 7.6)

10.5.2. Popularity-based advertising

[00182] Decisions about advertising in a document may also be based on statistics about the readership. The advertising space around the most popular columnists may be sold at a premium rate. Advertisers might even be charged or compensated some time after the document is published based on knowledge about how it was received.

10.6. Marketing based on Life Library

[00183] The "Life Library" or scan history described in Sections 6.1 and 16.1 can be an extremely valuable source of information about the interests and habits of a user. Subject to the appropriate consent and privacy issues, such data can inform offers of goods or services to the user. Even in an anonymous form, the statistics gathered can be exceedingly useful.

10.7. Sale/information at later date (when available)

[00184] Advertising and other opportunities for commercial transactions may not be presented to the user immediately at the time of text capture. For example, the opportunity to purchase a sequel to a novel may not be available at the time the user is reading the novel, but the system may present them with that opportunity when the sequel is published.

[00185] A user may capture data that relates to a purchase or other commercial transaction, but may choose not to initiate and/or complete the

transaction at the time the capture is made. In some embodiments, data related to captures is stored in a user's Life Library, and these Life Library entries can remain "active" (i.e., capable of subsequent interactions similar to those available at the time the capture was made). Thus a user may review a capture at some later time, and optionally complete a transaction based on that capture. Because the system can keep track of when and where the original capture occurred, all parties involved in the transaction can be properly compensated. For example, the author who wrote the story - and the publisher who published the story - that appeared next to the advertisement from which the user captured data can be compensated when, six months later, the user visits their Life Library, selects that particular capture from the history, and chooses "Purchase this item at Amazon" from the pop-up menu (which can be similar or identical to the menu optionally presented at the time of the capture).

11. Operating System and Application Integration

[00186] Modern Operating Systems (OSs) and other software packages have many characteristics that can be advantageously exploited for use with the described system, and may also be modified in various ways to provide an even better platform for its use.

11.1. Incorporation of scan and print-related information in metadata and indexing

[00187] New and upcoming file systems and their associated databases often have the ability to store a variety of metadata associated with each file. Traditionally, this metadata has included such things as the ID of the user who created the file, the dates of creation, last modification, and last use. Newer file systems allow such extra information as keywords, image characteristics, document sources and user comments to be stored, and in some systems this metadata can be arbitrarily extended. File systems can therefore be used to store information that would be useful in implementing the current system. For example, the date when a given document was last printed can be stored by the file system, as can details about which text from it has been captured from paper using the described system, and when and by whom.

[00188] Operating systems are also starting to incorporate search engine facilities that allow users to find local files more easily. These facilities can be advantageously used by the system. It means that many of the search-related concepts discussed in Sections 3 and 4 apply not just to today's Internet-based and similar search engines, but also to every personal computer.

[00189] In some cases specific software applications will also include support for the system above and beyond the facilities provided by the OS.

11.2. OS Support for capture devices

[00190] As the use of capture devices such as pen scanners becomes increasingly common, it will become desirable to build support for them into the operating system, in much the same way as support is provided for mice and printers, since the applicability of capture devices extends beyond a single software application. The same will be true for other aspects of the system's operation. Some examples are discussed below. In some embodiments, the entire described system, or the core of it, is provided by the OS. In some embodiments, support for the system is provided by Application Programming Interfaces (APIs) that can be used by other software packages, including those directly implementing aspects of the system.

11.2.1. Support for OCR and other recognition technologies

[00191] Most of the methods of capturing text from a rendered document require some recognition software to interpret the source data, typically a scanned image or some spoken words, as text suitable for use in the system. Some OSs include support for speech or handwriting recognition, though it is less common for OSs to include support for OCR, since in the past the use of OCR has typically been limited to a small range of applications.

[00192] As recognition components become part of the OS, they can take better advantage of other facilities provided by the OS. Many systems include spelling dictionaries, grammar analysis tools, internationalization and localization facilities, for example, all of which can be advantageously employed by the described system for its recognition process, especially

since they may have been customized for the particular user to include words and phrases that he/she would commonly encounter.

[00193] If the operating system includes full-text indexing facilities, then these can also be used to inform the recognition process, as described in Section 9.3.

11.2.2. Action to be taken on scans

[00194] If an optical scan or other capture occurs and is presented to the OS, it may have a default action to be taken under those circumstances in the event that no other subsystem claims ownership of the capture. An example of a default action is presenting the user with a choice of alternatives, or submitting the captured text to the OS's built-in search facilities.

11.2.3. OS has default action for particular documents or document types

[00195] If the digital source of the rendered document is found, the OS may have a standard action that it will take when that particular document, or a document of that class, is scanned. Applications and other subsystems may register with the OS as potential handlers of particular types of capture, in a similar manner to the announcement by applications of their ability to handle certain file types.

[00196] Markup data associated with a rendered document, or with a capture from a document, can include instructions to the operating system to launch specific applications, pass applications arguments, parameters, or data, etc.

11.2.4. Interpretation of gestures and mapping into standard actions

[00197] In Section 12.1.3 the use of "gestures" is discussed, particularly in the case of optical scanning, where particular movements made with a handheld scanner might represent standard actions such as marking the start and end of a region of text.

[00198] This is analogous to actions such as pressing the shift key on a keyboard while using the cursor keys to select a region of text, or using the

wheel on a mouse to scroll a document. Such actions by the user are sufficiently standard that they are interpreted in a system-wide way by the OS, thus ensuring consistent behavior. The same is desirable for scanner gestures and other scanner-related actions.

11.2.5. Set response to standard (and non-standard) iconic/text printed menu items

[00199] In a similar way, certain items of text or other symbols may, when scanned, cause standard actions to occur, and the OS may provide a selection of these. An example might be that scanning the text "[print]" in any document would cause the OS to retrieve and print a copy of that document. The OS may also provide a way to register such actions and associate them with particular scans.

11.3. Support in system GUI components for typical scan-initiated activities

[00200] Most software applications are based substantially on standard Graphical User Interface components provided by the OS.

[00201] Use of these components by developers helps to ensure consistent behavior across multiple packages, for example that pressing the left-cursor key in any text-editing context should move the cursor to the left, without every programmer having to implement the same functionality independently.

[00202] A similar consistency in these components is desirable when the activities are initiated by text-capture or other aspects of the described system. Some examples are given below.

11.3.1. Interface to find particular text content

[00203] A typical use of the system may be for the user to scan an area of a paper document, and for the system to open the electronic counterpart in a software package that is able to display or edit it, and cause that package to scroll to and highlight the scanned text (Section 12.2.1). The first part of this process, finding and opening the electronic document, is typically provided by the OS and is standard across software packages. The second part, however – locating a particular piece of text within a document and causing the

package to scroll to it and highlight it – is not yet standardized and is often implemented differently by each package. The availability of a standard API for this functionality could greatly enhance the operation of this aspect of the system.

11.3.2. Text interactions

[00204] Once a piece of text has been located within a document, the system may wish to perform a variety of operations upon that text. As an example, the system may request the surrounding text, so that the user's capture of a few words could result in the system accessing the entire sentence or paragraph containing them. Again, this functionality can be usefully provided by the OS rather than being implemented in every piece of software that handles text.

11.3.3. Contextual (popup) menus

[00205] Some of the operations that are enabled by the system will require user feedback, and this may be optimally requested within the context of the application handling the data. In some embodiments, the system uses the application pop-up menus traditionally associated with clicking the right mouse button on some text. The system inserts extra options into such menus, and causes them to be displayed as a result of activities such as scanning a paper document.

11.4. Web/network interfaces

[00206] In today's increasingly networked world, much of the functionality available on individual machines can also be accessed over a network, and the functionality associated with the described system is no exception. As an example, in an office environment, many paper documents received by a user may have been printed by other users' machines on the same corporate network. The system on one computer, in response to a capture, may be able to query those other machines for documents which may correspond to that capture, subject to the appropriate permission controls.

11.5. Printing of document causes saving

[00207] An important factor in the integration of paper and digital documents is maintaining as much information as possible about the transitions between the two. In some embodiments, the OS keeps a simple record of when any document was printed and by whom. In some embodiments, the OS takes one or more further actions that would make it better suited for use with the system. Examples include:

[00208] • Saving the digital rendered version of every document printed along with information about the source from which it was printed

[00209] • Saving a subset of useful information about the printed version – for example, the fonts used and where the line breaks occur – which might aid future scan interpretation

[00210] • Saving the version of the source document associated with any printed copy

[00211] • Indexing the document automatically at the time of printing and storing the results for future searching

11.6. My (Printed/Scanned) Documents

[00212] An OS often maintains certain categories of folders or files that have particular significance. A user's documents may, by convention or design, be found in a "My Documents" folder, for example. Standard file-opening dialogs may automatically include a list of recently opened documents.

[00213] On an OS optimized for use with the described system, such categories may be enhanced or augmented in ways that take into account a user's interaction with paper versions of the stored files. Categories such as "My Printed Documents" or "My Recently-Read Documents" might usefully be identified and incorporated in its operations.

11.7. OS-level markup hierarchies

[00214] Since important aspects of the system are typically provided using the "markup" concepts discussed in Section 5, it would clearly be advantageous to have support for such markup provided by the OS in a way

that was accessible to multiple applications as well as to the OS itself. In addition, layers of markup may be provided by the OS, based on its own knowledge of documents under its control and the facilities it is able to provide.

11.8. Use of OS DRM facilities

[00215] An increasing number of operating systems support some form of "Digital Rights Management": the ability to control the use of particular data according to the rights granted to a particular user, software entity or machine. It may inhibit unauthorized copying or distribution of a particular document, for example.

12. User Interface

[00216] The user interface of the system may be entirely on a PC, if the capture device is relatively dumb and is connected to it by a cable, or entirely on the device, if it is sophisticated and with significant processing power of its own. In some cases, some functionality resides in each component. Part, or indeed all, of the system's functionality may also be implemented on other devices such as mobile phones or PDAs.

[00217] The descriptions in the following sections are therefore indications of what may be desirable in certain implementations, but they are not necessarily appropriate for all and may be modified in several ways.

12.1. On the capture device

[00218] With all capture devices, but particularly in the case of an optical scanner, the user's attention will generally be on the device and the paper at the time of scanning. It is very desirable, then, that any input and feedback needed as part of the process of scanning do not require the user's attention to be elsewhere, for example on the screen of a computer, more than is necessary.

12.1.1. Feedback on scanner

[00219] A handheld scanner may have a variety of ways of providing feedback to the user about particular conditions. The most obvious types are direct visual, where the scanner incorporates indicator lights or even a full

display, and auditory, where the scanner can make beeps, clicks or other sounds. Important alternatives include tactile feedback, where the scanner can vibrate, buzz, or otherwise stimulate the user's sense of touch, and projected feedback, where it indicates a status by projecting onto the paper anything from a colored spot of light to a sophisticated display.

[00220] Important immediate feedback that may be provided on the device includes:

[00221] • *feedback on the scanning process* – user scanning too fast, at too great an angle, or drifting too high or low on a particular line

[00222] • *sufficient content* – enough has been scanned to be pretty certain of finding a match if one exists – important for disconnected operation

[00223] • *context known* – a source of the text has been located

[00224] • *unique context known* – one unique source of the text has been located

[00225] • *availability of content* – indication of whether the content is freely available to the user, or at a cost

[00226] Many of the user interactions normally associated with the later stages of the system may also take place on the capture device if it has sufficient abilities, for example, to display part or all of a document.

12.1.2. Controls on scanner

[00227] The device may provide a variety of ways for the user to provide input in addition to basic text capture. Even when the device is in close association with a host machine that has input options such as keyboards and mice, it can be disruptive for the user to switch back and forth between manipulating the scanner and using a mouse, for example.

[00228] The handheld scanner may have buttons, scroll/jog-wheels, touch-sensitive surfaces, and/or accelerometers for detecting the movement of the device. Some of these allow a richer set of interactions while still holding the scanner.

[00229] For example, in response to scanning some text, the system presents the user with a set of several possible matching documents. The user uses a scroll-wheel on the side of the scanner to select one from the list, and clicks a button to confirm the selection.

12.1.3. Gestures

[00230] The primary reason for moving a scanner across the paper is to capture text, but some movements may be detected by the device and used to indicate other user intentions. Such movements are referred to herein as "gestures."

[00231] As an example, the user can indicate a large region of text by scanning the first few words in conventional left-to-right order, and the last few in reverse order, i.e. right to left. The user can also indicate the vertical extent of the text of interest by moving the scanner down the page over several lines. A backwards scan might indicate cancellation of the previous scan operation.

12.1.4. Online / Offline behavior

[00232] Many aspects of the system may depend on network connectivity, either between components of the system such as a scanner and a host laptop, or with the outside world in the form of a connection to corporate databases and Internet search. This connectivity may not be present all the time, however, and so there will be occasions when part or all of the system may be considered to be "offline." It is desirable to allow the system to continue to function usefully in those circumstances.

[00233] The device may be used to capture text when it is out of contact with other parts of the system. A very simple device may simply be able to store the image or audio data associated with the capture, ideally with a timestamp indicating when it was captured. The various captures may be uploaded to the rest of the system when the device is next in contact with it, and handled then. The device may also upload other data associated with the captures, for example voice annotations associated with optical scans, or location information.

[00234] More sophisticated devices may be able to perform some or all of the system operations themselves despite being disconnected. Various techniques for improving their ability to do so are discussed in Section 15.3. Often it will be the case that some, but not all, of the desired actions can be performed while offline. For example, the text may be recognized, but identification of the source may depend on a connection to an Internet-based search engine. In some embodiments, the device therefore stores sufficient information about how far each operation has progressed for the rest of the system to proceed efficiently when connectivity is restored.

[00235] The operation of the system will, in general, benefit from immediately available connectivity, but there are some situations in which performing several captures and then processing them as a batch can have advantages. For example, as discussed in Section 13 below, the identification of the source of a particular capture may be greatly enhanced by examining other captures made by the user at approximately the same time. In a fully connected system where live feedback is being provided to the user, the system is only able to use past captures when processing the current one. If the capture is one of a batch stored by the device when offline, however, the system will be able to take into account any data available from later captures as well as earlier ones when doing its analysis.

12.2. On a host device

[00236] A scanner will often communicate with some other device, such as a PC, PDA, phone or digital camera to perform many of the functions of the system, including more detailed interactions with the user.

12.2.1. Activities performed in response to a capture

[00237] When the host device receives a capture, it may initiate a variety of activities. An incomplete list of possible activities performed by the system after locating and electronic counterpart document associated with the capture and a location within that document follows.

[00238] • The details of the capture may be stored in the user's history. (Section 6.1)

[00239] • The document may be retrieved from local storage or a remote location. (Section 8)

[00240] • The operating system's metadata and other records associated with the document may be updated. (Section 11.1)

[00241] • Markup associated with the document may be examined to determine the next relevant operations. (Section 5)

[00242] • A software application may be started to edit, view or otherwise operate on the document. The choice of application may depend on the source document, or on the contents of the scan, or on some other aspect of the capture. (Section 11.2.2, 11.2.3)

[00243] • The application may scroll to, highlight, move the insertion point to, or otherwise indicate the location of the capture. (Section 11.3)

[00244] • The precise bounds of the captured text may be modified, for example to select whole words, sentences or paragraphs around the captured text. (Section 11.3.2)

[00245] • The user may be given the option to copy the capture text to the clipboard or perform other standard operating system or application-specific operations upon it.

[00246] • Annotations may be associated with the document or the captured text. These may come from immediate user input, or may have been captured earlier, for example in the case of voice annotations associated with an optical scan. (Section 19.4)

[00247] • Markup may be examined to determine a set of further possible operations for the user to select.

12.2.2. Contextual popup menus

[00248] Sometimes the appropriate action to be taken by the system will be obvious, but sometimes it will require a choice to be made by the user. One good way to do this is through the use of "popup menus" or, in cases where the content is also being displayed on a screen, with so-called "contextual menus" that appear close to the content. (See Section 11.3.3). In

some embodiments, the scanner device projects a popup menu onto the paper document. A user may select from such menus using traditional methods such as a keyboard and mouse, or by using controls on the capture device (Section 12.1.2), gestures (Section 12.1.3), or by interacting with the computer display using the scanner (Section 12.2.4). In some embodiments, the popup menus which can appear as a result of a capture include default items representing actions which occur if the user does not respond – for example, if the user ignores the menu and makes another capture.

12.2.3. Feedback on disambiguation

[00249] When a user starts capturing text, there will initially be several documents or other text locations that it could match. As more text is captured, and other factors are taken into account (Section 13), the number of candidate locations will decrease until the actual location is identified, or further disambiguation is not possible without user input. In some embodiments, the system provides a real-time display of the documents or the locations found, for example in list, thumbnail-image or text-segment form, and for the number of elements in that display to reduce in number as capture continues. In some embodiments, the system displays thumbnails of all candidate documents, where the size or position of the thumbnail is dependent on the probability of it being the correct match.

[00250] When a capture is unambiguously identified, this fact may be emphasized to the user, for example using audio feedback.

[00251] Sometimes the text captured will occur in many documents and will be recognized to be a quotation. The system may indicate this on the screen, for example by grouping documents containing a quoted reference around the original source document.

12.2.4. Scanning from screen

[00252] Some optical scanners may be able to capture text displayed on a screen as well as on paper. Accordingly, the term *rendered document* is used herein to indicate that printing onto paper is not the only form of rendering, and that the capture of text or symbols for use by the system may be equally valuable when that text is displayed on an electronic display.

[00253] The user of the described system may be required to interact with a computer screen for a variety of other reasons, such as to select from a list of options. It can be inconvenient for the user to put down the scanner and start using the mouse or keyboard. Other sections have described physical controls on the scanner (Section 12.1.2) or gestures (Section 12.1.3) as methods of input which do not require this change of tool, but using the scanner on the screen itself to scan some text or symbol is an important alternative provided by the system.

[00254] In some embodiments, the optics of the scanner allow it to be used in a similar manner to a light-pen, directly sensing its position on the screen without the need for actual scanning of text, possibly with the aid of special hardware or software on the computer.

13. Context Interpretation

[00255] An important aspect of the described system is the use of other factors, beyond the simple capture of a string of text, to help identify the document in use. A capture of a modest amount of text may often identify the document uniquely, but in many situations it will identify a few candidate documents. One solution is to prompt the user to confirm the document being scanned, but a preferable alternative is to make use of other factors to narrow down the possibilities automatically. Such supplemental information can dramatically reduce the amount of text that needs to be captured and/or increase the reliability and speed with which the location in the electronic counterpart can be identified. This extra material is referred to as "context," and it was discussed briefly in Section 4.2.2. We now consider it in more depth.

13.1. System and capture context

[00256] Perhaps the most important example of such information is the user's capture history.

[00257] It is highly probable that any given capture comes from the same document as the previous one, or from an associated document, especially if the previous capture took place in the last few minutes (Section 6.1.2).

Conversely, if the system detects that the font has changed between two scans, it is more likely that they are from different documents.

[00258] Also useful are the user's longer-term capture history and reading habits. These can also be used to develop a model of the user's interests and associations.

13.2. User's real-world context

[00259] Another example of useful context is the user's geographical location. A user in Paris is much more likely to be reading Le Monde than the Seattle Times, for example. The timing, size and geographical distribution of printed versions of the documents can therefore be important, and can to some degree be deduced from the operation of the system.

[00260] The time of day may also be relevant, for example in the case of a user who always reads one type of publication on the way to work, and a different one at lunchtime or on the train going home.

13.3. Related digital context

[00261] The user's recent use of electronic documents, including those searched for or retrieved by more conventional means, can also be a helpful indicator.

[00262] In some cases, such as on a corporate network, other factors may be usefully considered:

[00263] • Which documents have been printed recently?

[00264] • Which documents have been modified recently on the corporate file server?

[00265] • Which documents have been emailed recently?

[00266] All of these examples might suggest that a user was more likely to be reading a paper version of those documents. In contrast, if the repository in which a document resides can affirm that the document has never been printed or sent anywhere where it might have been printed, then it can be safely eliminated in any searches originating from paper.

13.4. Other statistics – the global context

[00267] Section 14 covers the analysis of the data stream resulting from paper-based searches, but it should be noted here that statistics about the popularity of documents with other readers, about the timing of that popularity, and about the parts of documents most frequently scanned are all examples of further factors which can be beneficial in the search process. The system brings the possibility of Google-type page-ranking to the world of paper.

[00268] See also Section 4.2.2 for some other implications of the use of context for search engines.

14. Data-stream Analysis

[00269] The use of the system generates an exceedingly valuable data-stream as a side effect. This stream is a record of what users are reading and when, and is in many cases a record of what they find particularly valuable in the things they read. Such data has never really been available before for paper documents.

[00270] Some ways in which this data can be useful for the system, and for the user of the system, are described in Section 6.1. This section concentrates on its use for others. There are, of course, substantial privacy issues to be considered with any distribution of data about what people are reading, but such issues as preserving the anonymity of data are well known to those of skill in the art.

14.1. Document Tracking

[00271] When the system knows which documents any given user is reading, it can also deduce who is reading any given document. This allows the tracking of a document through an organization, to allow analysis, for example, of who is reading it and when, how widely it was distributed, how long that distribution took, and who has seen current versions while others are still working from out-of-date copies.

[00272] For published documents that have a wider distribution, the tracking of individual copies is more difficult, but the analysis of the distribution of readership is still possible.

14.2. Read Ranking – popularity of documents and sub-regions

[00273] In situations where users are capturing text or other data that is of particular interest to them, the system can deduce the popularity of certain documents and of particular sub-regions of those documents. This forms a valuable input to the system itself (Section 4.2.2) and an important source of information for authors, publishers and advertisers (Section 7.6, Section 10.5). This data is also useful when integrated in search engines and search indices – for example, to assist in ranking search results for queries coming from rendered documents, and/or to assist in ranking conventional queries typed into a web browser.

14.3. Analysis of users – building profiles

[00274] Knowledge of what a user is reading enables the system to create a quite detailed model of the user's interests and activities. This can be useful on an abstract statistical basis – "35% of users who buy this newspaper also read the latest book by that author" – but it can also allow other interactions with the individual user, as discussed below.

14.3.1. Social Networking

[00275] One example is connecting one user with others who have related interests. These may be people already known to the user. The system may ask a university professor, "Did you know that your colleague at XYZ University has also just read this paper?" The system may ask a user, "Do you want to be linked up with other people in your neighborhood who are also how reading Jane Eyre?" Such links may be the basis for the automatic formation of book clubs and similar social structures, either in the physical world or online.

14.3.2. Marketing

[00276] Section 10.6 has already mentioned the idea of offering products and services to an individual user based on their interactions with the system. Current online booksellers, for example, often make recommendations to a user based on their previous interactions with the bookseller. Such recommendations become much more useful when they are based on interactions with the actual books.

14.4. Marketing based on other aspects of the data-stream

[00277] We have discussed some of the ways in which the system may influence those publishing documents, those advertising through them, and other sales initiated from paper (Section 10). Some commercial activities may have no direct interaction with the paper documents at all and yet may be influenced by them. For example, the knowledge that people in one community spend more time reading the sports section of the newspaper than they do the financial section might be of interest to somebody setting up a health club.

14.5. Types of data that may be captured

[00278] In addition to the statistics discussed, such as who is reading which bits of which documents, and when and where, it can be of interest to examine the actual contents of the text captured, regardless of whether or not the document has been located.

[00279] In many situations, the user will also not just be capturing some text, but will be causing some action to occur as a result. It might be emailing a reference to the document to an acquaintance, for example. Even in the absence of information about the identity of the user or the recipient of the email, the knowledge that somebody considered the document worth emailing is very useful.

[00280] In addition to the various methods discussed for deducing the value of a particular document or piece of text, in some circumstances the user will explicitly indicate the value by assigning it a rating.

[00281] Lastly, when a particular set of users are known to form a group, for example when they are known to be employees of a particular company, the aggregated statistics of that group can be used to deduce the importance of a particular document to that group.

15. Device Features and Functions

[00282] A capture device for use with the system needs little more than a way of capturing text from a rendered version of the document. As described earlier (Section 1.2), this capture may be achieved through a variety of methods including taking a photograph of part of the document or typing some

words into a mobile phone keypad. This capture may be achieved using a small hand-held optical scanner capable of recording a line or two of text at a time, or an audio capture device such as a voice-recorder into which the user is reading text from the document. The device used may be a combination of these – an optical scanner which could also record voice annotations, for example – and the capturing functionality may be built into some other device such as a mobile phone, PDA, digital camera or portable music player.

15.1. Input and output

[00283] Many of the possibly beneficial additional input and output facilities for such a device have been described in Section 12.1. They include buttons, scroll-wheels and touch-pads for input, and displays, indicator lights, audio and tactile transducers for output. Sometimes the device will incorporate many of these, sometimes very few. Sometimes the capture device will be able to communicate with another device that already has them (Section 15.6), for example using a wireless link, and sometimes the capture functionality will be incorporated into such other device (Section 15.7).

15.2. Connectivity

[00284] In some embodiments, the device implements the majority of the system itself. In some embodiments, however, it often communicates with a PC or other computing device and with the wider world using communications facilities.

[00285] Often these communications facilities are in the form of a general-purpose data network such as Ethernet, 802.11 or UWB or a standard peripheral-connecting network such as USB, IEEE-1394 (Firewire), Bluetooth™ or infra-red. When a wired connection such as Firewire or USB is used, the device may receive electrical power through the same connection. In some circumstances, the capture device may appear to a connected machine to be a conventional peripheral such as a USB storage device.

[00286] Lastly, the device may in some circumstances "dock" with another device, either to be used in conjunction with that device or for convenient storage.

15.3. Caching and other online/offline functionality

[00287] Sections 3.5 and 12.1.4 have raised the topic of disconnected operation. When a capture device has a limited subset of the total system's functionality, and is not in communication with the other parts of the system, the device can still be useful, though the functionality available will sometimes be reduced. At the simplest level, the device can record the raw image or audio data being captured and this can be processed later. For the user's benefit, however, it can be important to give feedback where possible about whether the data captured is likely to be sufficient for the task in hand, whether it can be recognized or is likely to be recognizable, and whether the source of the data can be identified or is likely to be identifiable later. The user will then know whether their capturing activity is worthwhile. Even when all of the above are unknown, the raw data can still be stored so that, at the very least, the user can refer to them later. The user may be presented with the image of a scan, for example, when the scan cannot be recognized by the OCR process.

[00288] To illustrate some of the range of options available, both a rather minimal optical scanning device and then a much more full-featured one are described below. Many devices occupy a middle ground between the two.

15.3.1. The SimpleScanner – a low-end offline example

[00289] The SimpleScanner has a scanning head able to read pixels from the page as it is moved along the length of a line of text. It can detect its movement along the page and record the pixels with some information about the movement. It also has a clock, which allows each scan to be time-stamped. The clock is synchronized with a host device when the SimpleScanner has connectivity. The clock may not represent the actual time of day, but relative times may be determined from it so that the host can deduce the actual time of a scan, or at worst the elapsed time between scans.

[00290] The SimpleScanner does not have sufficient processing power to perform any OCR itself, but it does have some basic knowledge about typical word-lengths, word-spacings, and their relationship to font size. It has some basic indicator lights which tell the user whether the scan is likely to be

readable, whether the head is being moved too fast, too slowly or too inaccurately across the paper, and when it determines that sufficient words of a given size are likely to have been scanned for the document to be identified.

[00291] The SimpleScanner has a USB connector and can be plugged into the USB port on a computer, where it will be recharged. To the computer it appears to be a USB storage device on which time-stamped data files have been recorded, and the rest of the system software takes over from this point.

15.3.2. The SuperScanner – a high-end offline example

[00292] The SuperScanner also depends on connectivity for its full operation, but it has a significant amount of on-board storage and processing which can help it make better judgments about the data captured while offline.

[00293] As it moves along the line of text, the captured pixels are stitched together and passed to an OCR engine that attempts to recognize the text. A number of fonts, including those from the user's most-read publications, have been downloaded to it to help perform this task, as has a dictionary that is synchronized with the user's spelling-checker dictionary on their PC and so contains many of the words they frequently encounter. Also stored on the scanner is a list of words and phrases with the typical frequency of their use – this may be combined with the dictionary. The scanner can use the frequency statistics both to help with the recognition process and also to inform its judgment about when a sufficient quantity of text has been captured; more frequently used phrases are less likely to be useful as the basis for a search query.

[00294] In addition, the full index for the articles in the recent issues of the newspapers and periodicals most commonly read by the user are stored on the device, as are the indices for the books the user has recently purchased from an online bookseller, or from which the user has scanned anything within the last few months. Lastly, the titles of several thousand of the most popular publications which have data available for the system are stored so that, in the absence of other information the user can scan the title and have a good idea as to whether or not captures from a particular work are likely to be retrievable in electronic form later.

[00295] During the scanning process, the system informs user that the captured data has been of sufficient quality and of a sufficient nature to make it probable that the electronic copy can be retrieved when connectivity is restored. Often the system indicates to the user that the scan is known to have been successful and that the context has been recognized in one of the on-board indices, or that the publication concerned is known to be making its data available to the system, so the later retrieval ought to be successful.

[00296] The SuperScanner docks in a cradle connected to a PC's Firewire or USB port, at which point, in addition to the upload of captured data, its various onboard indices and other databases are updated based on recent user activity and new publications. It also has the facility to connect to wireless public networks or to communicate via Bluetooth to a mobile phone and thence with the public network when such facilities are available.

15.4. Features for optical scanning

[00297] We now consider some of the features that may be particularly desirable in an optical scanner device.

15.4.1. Flexible positioning and convenient optics

[00298] One of the reasons for the continuing popularity of paper is the ease of its use in a wide variety of situations where a computer, for example, would be impractical or inconvenient. A device intended to capture a substantial part of a user's interaction with paper should therefore be similarly convenient in use. This has not been the case for scanners in the past; even the smallest hand-held devices have been somewhat unwieldy. Those designed to be in contact with the page have to be held at a precise angle to the paper and moved very carefully along the length of the text to be scanned. This is acceptable when scanning a business report on an office desk, but may be impractical when scanning a phrase from a novel while waiting for a train. Scanners based on camera-type optics that operate at a distance from the paper may similarly be useful in some circumstances.

[00299] Some embodiments of the system use a scanner that scans in contact with the paper, and which, instead of lenses, uses an image conduit a bundle of optical fibers to transmit the image from the page to the optical

sensor device. Such a device can be shaped to allow it to be held in a natural position; for example, in some embodiments, the part in contact with the page is wedge-shaped, allowing the user's hand to move more naturally over the page in a movement similar to the use of a highlighter pen. The conduit is either in direct contact with the paper or in close proximity to it, and may have a replaceable transparent tip that can protect the image conduit from possible damage. As has been mentioned in Section 12.2.4, the scanner may be used to scan from a screen as well as from paper, and the material of the tip can be chosen to reduce the likelihood of damage to such displays.

[00300] Lastly, some embodiments of the device will provide feedback to the user during the scanning process which will indicate through the use of light, sound or tactile feedback when the user is scanning too fast, too slow, too unevenly or is drifting too high or low on the scanned line.

15.5. Security, identity, authentication, personalization and billing

[00301] As described in Section 6, the capture device may form an important part of identification and authorization for secure transactions, purchases, and a variety of other operations. It may therefore incorporate, in addition to the circuitry and software required for such a role, various hardware features that can make it more secure, such as a smartcard reader, RFID, or a keypad on which to type a PIN.

[00302] It may also include various biometric sensors to help identify the user. In the case of an optical scanner, for example, the scanning head may also be able to read a fingerprint. For a voice recorder, the voice pattern of the user may be used.

15.6. Device associations

[00303] In some embodiments, the device is able to form an association with other nearby devices to increase either its own or their functionality. In some embodiments, for example, it uses the display of a nearby PC or phone to give more detailed feedback about its operation, or uses their network connectivity. The device may, on the other hand, operate in its role as a security and identification device to authenticate operations performed by the

other device. Or it may simply form an association in order to function as a peripheral to that device.

[00304] An interesting aspect of such associations is that they may be initiated and authenticated using the capture facilities of the device. For example, a user wishing to identify themselves securely to a public computer terminal may use the scanning facilities of the device to scan a code or symbol displayed on a particular area of the terminal's screen and so effect a key transfer. An analogous process may be performed using audio signals picked up by a voice-recording device.

15.7. Integration with other devices

[00305] In some embodiments, the functionality of the capture device is integrated into some other device that is already in use. The integrated devices may be able to share a power supply, data capture and storage capabilities, and network interfaces. Such integration may be done simply for convenience, to reduce cost, or to enable functionality that would not otherwise be available.

[00306] Some examples of devices into which the capture functionality can be integrated include:

[00307] • *an existing peripheral* such as a mouse, a stylus, a USB "webcam" camera, a Bluetooth™ headset or a remote control

[00308] • *another processing/storage device*, such as a PDA, an MP3 player, a voice recorder, a digital camera or a mobile phone

[00309] • *other often-carried items*, just for convenience – a watch, a piece of jewelry, a pen, a car key fob

15.7.1. Mobile phone Integration

[00310] As an example of the benefits of integration, we consider the use of a modified mobile phone as the capture device.

[00311] In some embodiments, the phone hardware is not modified to support the system, such as where the text capture can be adequately done through voice recognition, where they can either be processed by the phone

itself, or handled by a system at the other end of a telephone call, or stored in the phone's memory for future processing. Many modern phones have the ability to download software that could implement some parts of the system. Such voice capture is likely to be suboptimal in many situations, however, for example when there is substantial background noise, and accurate voice recognition is a difficult task at the best of times. The audio facilities may best be used to capture voice annotations.

[00312] In some embodiments, the camera built into many mobile phones is used to capture an image of the text. The phone display, which would normally act as a viewfinder for the camera, may overlay on the live camera image information about the quality of the image and its suitability for OCR, which segments of text are being captured, and even a transcription of the text if the OCR can be performed on the phone.

[00313] In some embodiments, the phone is modified to add dedicated capture facilities, or to provide such functionality in a clip-on adaptor or a separate Bluetooth-connected peripheral in communication with the phone. Whatever the nature of the capture mechanism, the integration with a modern cellphone has many other advantages. The phone has connectivity with the wider world, which means that queries can be submitted to remote search engines or other parts of the system, and copies of documents may be retrieved for immediate storage or viewing. A phone typically has sufficient processing power for many of the functions of the system to be performed locally, and sufficient storage to capture a reasonable amount of data. The amount of storage can also often be expanded by the user. Phones have reasonably good displays and audio facilities to provide user feedback, and often a vibrate function for tactile feedback. They also have good power supplies.

[00314] Most significantly of all, they are a device that most users are already carrying.

PART III – EXAMPLE APPLICATIONS OF THE SYSTEM

[00315] This section lists example uses of the system and applications that may be built on it. This list is intended to be purely illustrative and in no sense exhaustive.

16. Personal Applications

16.1. Life Library

[00316] The Life Library (see also Section 6.1.1) is a digital archive of any important documents that the subscriber wishes to save and is a set of embodiments of services of this system. Important books, magazine articles, newspaper clippings, etc., can all be saved in digital form in the Life Library. Additionally, the subscriber's annotations, comments, and notes can be saved with the documents. The Life Library can be accessed via the Internet and World Wide Web.

[00317] The system creates and manages the Life Library document archive for subscribers. The subscriber indicates which documents the subscriber wishes to have saved in his life library by scanning information from the document or by otherwise indicating to the system that the particular document is to be added to the subscriber's Life Library. The scanned information is typically text from the document but can also be a barcode or other code identifying the document. The system accepts the code and uses it to identify the source document. After the document is identified the system can store either a copy of the document in the user's Life Library or a link to a source where the document may be obtained.

[00318] One embodiment of the Life Library system can check whether the subscriber is authorized to obtain the electronic copy. For example, if a reader scans text or an identifier from a copy of an article in the New York Times (NYT) so that the article will be added to the reader's Life Library, the Life Library system will verify with the NYT whether the reader is subscribed to the online version of the NYT; if so, the reader gets a copy of the article stored in his Life Library account; if not, information identifying the document and how to order it is stored in his Life Library account.

[00319] In some embodiments, the system maintains a subscriber profile for each subscriber that includes access privilege information. Document access information can be compiled in several ways, two of which are: 1) the subscriber supplies the document access information to the Life Library system, along with his account names and passwords, etc., or 2) the Life Library service provider queries the publisher with the subscriber's information and the publisher responds by providing access to an electronic copy if the Life Library subscriber is authorized to access the material. If the Life Library subscriber is not authorized to have an electronic copy of the document, the publisher provides a price to the Life Library service provider, which then provides the customer with the option to purchase the electronic document. If so, the Life Library service provider either pays the publisher directly and bills the Life Library customer later or the Life Library service provider immediately bills the customer's credit card for the purchase. The Life Library service provider would get a percentage of the purchase price or a small fixed fee for facilitating the transaction.

[00320] The system can archive the document in the subscriber's personal library and/or any other library to which the subscriber has archival privileges. For example, as a user scans text from a printed document, the Life Library system can identify the rendered document and its electronic counterpart. After the source document is identified, the Life Library system might record information about the source document in the user's personal library and in a group library to which the subscriber has archival privileges. Group libraries are collaborative archives such as a document repository for: a group working together on a project, a group of academic researchers, a group web log, etc.

[00321] The life library can be organized in many ways: chronologically, by topic, by level of the subscriber's interest, by type of publication (newspaper, book, magazine, technical paper, etc.), where read, when read, by ISBN or by Dewey decimal, etc. In one alternative, the system can learn classifications based on how other subscribers have classified the same document. The system can suggest classifications to the user or automatically classify the document for the user.

[00322] In various embodiments, annotations may be inserted directly into the document or may be maintained in a separate file. For example, when a subscriber scans text from a newspaper article, the article is archived in his Life Library with the scanned text highlighted. Alternatively, the article is archived in his Life Library along with an associated annotation file (thus leaving the archived document unmodified). Embodiments of the system can keep a copy of the source document in each subscriber's library, a copy in a master library that many subscribers can access, or link to a copy held by the publisher.

[00323] In some embodiments, the Life Library stores only the user's modifications to the document (e.g., highlights, etc.) and a link to an online version of the document (stored elsewhere). The system or the subscriber merges the changes with the document when the subscriber subsequently retrieves the document.

[00324] If the annotations are kept in a separate file, the source document and the annotation file are provided to the subscriber and the subscriber combines them to create a modified document. Alternatively, the system combines the two files prior to presenting them to the subscriber. In another alternative, the annotation file is an overlay to the document file and can be overlaid on the document by software in the subscriber's computer.

[00325] Subscribers to the Life Library service pay a monthly fee to have the system maintain the subscriber's archive. Alternatively, the subscriber pays a small amount (e.g., a micro-payment) for each document stored in the archive. Alternatively, the subscriber pays to access the subscriber's archive on a per-access fee. Alternatively, subscribers can compile libraries and allow others to access the materials/annotations on a revenue share model with the Life Library service provider and copyright holders. Alternatively, the Life Library service provider receives a payment from the publisher when the Life Library subscriber orders a document (a revenue share model with the publisher, where the Life Library service provider gets a share of the publisher's revenue).

[00326] In some embodiments, the Life Library service provider acts as an intermediary between the subscriber and the copyright holder (or copyright holder's agent, such as the Copyright Clearance Center, a.k.a. CCC) to facilitate billing and payment for copyrighted materials. The Life Library service provider uses the subscriber's billing information and other user account information to provide this intermediation service. Essentially, the Life Library service provider leverages the pre-existing relationship with the subscriber to enable purchase of copyrighted materials on behalf of the subscriber.

[00327] In some embodiments, the Life Library system can store excerpts from documents. For example, when a subscriber scans text from a paper document, the regions around the scanned text are excerpted and placed in the Life Library, rather than the entire document being archived in the life library. This is especially advantageous when the document is long because preserving the circumstances of the original scan prevents the subscriber from re-reading the document to find the interesting portions. Of course, a hyperlink to the entire electronic counterpart of the paper document can be included with the excerpt materials.

[00328] In some embodiments, the system also stores information about the document in the Life Library, such as author, publication title, publication date, publisher, copyright holder (or copyright holder's licensing agent), ISBN, links to public annotations of the document, readrank, etc. Some of this additional information about the document is a form of paper document metadata. Third parties may create public annotation files for access by persons other than themselves, such the general public. Linking to a third party's commentary on a document is advantageous because reading annotation files of other users enhances the subscriber's understanding of the document.

[00329] In some embodiments, the system archives materials by class. This feature allows a Life Library subscriber to quickly store electronic counterparts to an entire class of paper documents without access to each paper document. For example, when the subscriber scans some text from a

copy of National Geographic magazine, the system provides the subscriber with the option to archive all back issues of the National Geographic. If the subscriber elects to archive all back issues, the Life Library service provider would then verify with the National Geographic Society whether the subscriber is authorized to do so. If not, the Life Library service provider can mediate the purchase of the right to archive the National Geographic magazine collection.

16.2. Life Saver

[00330] A variation on, or enhancement of, the Life Library concept is the "Life Saver," where the system uses the text captured by a user to deduce more about their other activities. The scanning of a menu from a particular restaurant, a program from a particular theater performance, a timetable at a particular railway station, or an article from a local newspaper allows the system to make deductions about the user's location and social activities, and could construct an automatic diary for them, for example as a website. The user would be able to edit and modify the diary, add additional materials such as photographs and, of course, look again at the items scanned.

17. Academic Applications

[00331] Portable scanners supported by the described system have many compelling uses in the academic setting. They can enhance student/teacher interaction and augment the learning experience. Among other uses, students can annotate study materials to suit their unique needs; teachers can monitor classroom performance; and teachers can automatically verify source materials cited in student assignments.

17.1. Children's books

[00332] A child's interaction with a paper document, such as a book, is monitored by a literacy acquisition system that employs a specific set of embodiments of this system. The child uses a portable scanner that communicates with other elements of the literacy acquisition system. In addition to the portable scanner, the literacy acquisition system includes a computer having a display and speakers, and a database accessible by the computer. The scanner is coupled with the computer (hardwired, short range RF, etc.). When the child sees an unknown word in the book, the child scans

it with the scanner. In one embodiment, the literacy acquisition system compares the scanned text with the resources in its database to identify the word. The database includes a dictionary, thesaurus, and/or multimedia files (e.g., sound, graphics, etc.). After the word has been identified, the system uses the computer speakers to pronounce the word and its definition to the child. In another embodiment, the word and its definition are displayed by the literacy acquisition system on the computer's monitor. Multimedia files about the scanned word can also be played through the computer's monitor and speakers. For example, if a child reading "Goldilocks and the Three Bears" scanned the word "bear," the system might pronounce the word "bear" and play a short video about bears on the computer's monitor. In this way, the child learns to pronounce the written word and is visually taught what the word means via the multimedia presentation.

[00333] The literacy acquisition system provides immediate auditory and/or visual information to enhance the learning process. The child uses this supplementary information to quickly acquire a deeper understanding of the written material. The system can be used to teach beginning readers to read, to help children acquire a larger vocabulary, etc. This system provides the child with information about words with which the child is unfamiliar or about which the child wants more information.

17.2. Literacy acquisition

[00334] In some embodiments, the system compiles personal dictionaries. If the reader sees a word that is new, interesting, or particularly useful or troublesome, the reader saves it (along with its definition) to a computer file. This computer file becomes the reader's personalized dictionary. This dictionary is generally smaller in size than a general dictionary so can be downloaded to a mobile station or associated device and thus be available even when the system isn't immediately accessible. In some embodiments, the personal dictionary entries include audio files to assist with proper word pronunciation and information identifying the paper document from which the word was scanned.

[00335] In some embodiments, the system creates customized spelling and vocabulary tests for students. For example, as a student reads an assignment, the student may scan unfamiliar words with the portable scanner. The system stores a list of all the words that the student has scanned. Later, the system administers a customized spelling/vocabulary test to the student on an associated monitor (or prints such a test on an associated printer).

17.3. Music teaching

[00336] The arrangement of notes on a musical staff is similar to the arrangement of letters in a line of text. The same scanning device discussed for capturing text in this system can be used to capture music notation, and an analogous process of constructing a search against databases of known musical pieces would allow the piece from which the capture occurred to be identified which can then be retrieved, played, or be the basis for some further action.

17.4. Detecting plagiarism

[00337] Teachers can use the system to detect plagiarism or to verify sources by scanning text from student papers and submitting the scanned text to the system. For example, a teacher who wishes to verify that a quote in a student paper came from the source that the student cited can scan a portion of the quote and compare the title of the document identified by the system with the title of the document cited by the student. Likewise, the system can use scans of text from assignments submitted as the student's original work to reveal if the text was instead copied.

17.5. Enhanced textbook

[00338] In some embodiments, capturing text from an academic textbook links students or staff to more detailed explanations, further exercises, student and staff discussions about the material, related example past exam questions, further reading on the subject, recordings of the lectures on the subject, and so forth. (See also Section 7.1.)

17.6. Language learning

[00339] In some embodiments, the system is used to teach foreign languages. Scanning a Spanish word, for example, might cause the word to be read aloud in Spanish along with its definition in English.

[00340] The system provides immediate auditory and/or visual information to enhance the new language acquisition process. The reader uses this supplementary information to acquire quickly a deeper understanding of the material. The system can be used to teach beginning students to read foreign languages, to help students acquire a larger vocabulary, etc. The system provides information about foreign words with which the reader is unfamiliar or for which the reader wants more information.

[00341] Reader interaction with a paper document, such as a newspaper or book, is monitored by a language skills system. The reader has a portable scanner that communicates with the language skills system. In some embodiments, the language skills system includes a computer having a display and speakers, and a database accessible by the computer. The scanner communicates with the computer (hardwired, short range RF, etc.). When the reader sees an unknown word in an article, the reader scans it with the scanner. The database includes a foreign language dictionary, thesaurus, and/or multimedia files (sound, graphics, etc.). In one embodiment, the system compares the scanned text with the resources in its database to identify the scanned word. After the word has been identified, the system uses the computer speakers to pronounce the word and its definition to the reader. In some embodiments, the word and its definition are both displayed on the computer's monitor. Multimedia files about grammar tips related to the scanned word can also be played through the computer's monitor and speakers. For example, if the words "to speak" are scanned, the system might pronounce the word "hablar," play a short audio clip that demonstrates the proper Spanish pronunciation, and display a complete list of the various conjugations of "hablar." In this way, the student learns to pronounce the written word, is visually taught the spelling of the word via the multimedia presentation, and learns how to conjugate the verb. The system

can also present grammar tips about the proper usage of "hablar" along with common phrases.

[00342] In some embodiments, the user scans a word or short phrase from a rendered document in a language other than the user's native language (or some other language that the user knows reasonably well). In some embodiments, the system maintains a prioritized list of the user's "preferred" languages. The system identifies the electronic counterpart of the rendered document, and determines the location of the scan within the document. The system also identifies a second electronic counterpart of the document that has been translated into one of the user's preferred languages, and determines the location in the translated document corresponding to the location of the scan in the original document. When the corresponding location is not known precisely, the system identifies a small region (e.g., a paragraph) that includes the corresponding location of the scanned location. The corresponding translated location is then presented to the user. This provides the user with a precise translation of the particular usage at the scanned location, including any slang or other idiomatic usage that is often difficult to accurately translate on a word-by-word basis.

17.7. Gathering research materials

[00343] A user researching a particular topic may encounter all sorts of material, both in print and on screen, which they might wish to record as relevant to the topic in some personal archive. The system would enable this process to be automatic as a result of scanning a short phrase in any piece of material, and could also create a bibliography suitable for insertion into a publication on the subject.

18. Commercial Applications

[00344] Obviously, commercial activities could be made out of almost any process discussed in this document, but here we concentrate on a few obvious revenue streams.

18.1. Fee-based searching and indexing

[00345] Conventional Internet search engines typically provide free search of electronic documents, and also make no charge to the content providers for

including their content in the index. In some embodiments, the system provides for charges to users and/or payments to search engines and/or content providers in connection with the operation and use of the system.

[00346] In some embodiments, subscribers to the system's services pay a fee for searches originating from scans of paper documents. For example, a stockbroker may be reading a Wall Street Journal article about a new product offered by Company X. By scanning the Company X name from the paper document and agreeing to pay the necessary fees, the stockbroker uses the system to search special or proprietary databases to obtain premium information about the company, such as analyst's reports. The system can also make arrangements to have priority indexing of the documents most likely to be read in paper form, for example by making sure all of the newspapers published on a particular day are indexed and available by the time they hit the streets.

[00347] Content providers may pay a fee to be associated with certain terms in search queries submitted from paper documents. For example, in one embodiment, the system chooses a most preferred content provider based on additional context about the provider (the context being, in this case, that the content provider has paid a fee to be moved up the results list). In essence, the search provider is adjusting paper document search results based on pre-existing financial arrangements with a content provider. See also the description of keywords and key phrases in Section 5.2.

[00348] Where access to particular content is to be restricted to certain groups of people (such as clients or employees), such content may be protected by a firewall and thus not generally indexable by third parties. The content provider may nonetheless wish to provide an index to the protected content. In such a case, the content provider can pay a service provider to provide the content provider's index to system subscribers. For example, a law firm may index all of a client's documents. The documents are stored behind the law firm's firewall. However, the law firm wants its employees and the client to have access to the documents through the portable scanner so it provides the index (or a pointer to the index) to the service provider, which in

turn searches the law firm's index when employees or clients of the law firm submit paper-scanned search terms via their portable scanners. The law firm can provide a list of employees and/or clients to the service provider's system to enable this function or the system can verify access rights by querying the law firm prior to searching the law firm's index. Note that in the preceding example, the index provided by the law firm is only of that client's documents, not an index of all documents at the law firm. Thus, the service provider can only grant the law firm's clients access to the documents that the law firm indexed for the client.

[00349] There are at least two separate revenue streams that can result from searches originating from paper documents: one revenue stream from the search function, and another from the content delivery function. The search function revenue can be generated from paid subscriptions from the scanner users, but can also be generated on a per-search charge. The content delivery revenue can be shared with the content provider or copyright holder (the service provider can take a percentage of the sale or a fixed fee, such as a micropayment, for each delivery), but also can be generated by a "referral" model in which the system gets a fee or percentage for every item that the subscriber orders from the online catalog and that the system has delivered or contributed to, regardless of whether the service provider intermediates the transaction. In some embodiments, the system service provider receives revenue for all purchases that the subscriber made from the content provider, either for some predetermined period of time or at any subsequent time when a purchase of an identified product is made.

18.2. Catalogs

[00350] Consumers may use the portable scanner to make purchases from paper catalogs. The subscriber scans information from the catalog that identifies the catalog. This information is text from the catalog, a bar code, or another identifier of the catalog. The subscriber scans information identifying the products that s/he wishes to purchase. The catalog mailing label may contain a customer identification number that identifies the customer to the catalog vendor. If so, the subscriber can also scan this customer identification

number. The system acts as an intermediary between the subscriber and the vendor to facilitate the catalog purchase by providing the customer's selection and customer identification number to the vendor.

18.3. Coupons

[00351] A consumer scans paper coupons and saves an electronic copy of the coupon in the scanner, or in a remote device such as a computer, for later retrieval and use. An advantage of electronic storage is that the consumer is freed from the burden of carrying paper coupons. A further advantage is that the electronic coupons may be retrieved from any location. In some embodiments, the system can track coupon expiration dates, alert the consumer about coupons that will expire soon, and/or delete expired coupons from storage. An advantage for the issuer of the coupons is the possibility of receiving more feedback about who is using the coupons and when and where they are captured and used.

19. General Applications

19.1. Forms

[00352] The system may be used to auto-populate an electronic document that corresponds to a paper form. A user scans in some text or a barcode that uniquely identifies the paper form. The scanner communicates the identity of the form and information identifying the user to a nearby computer. The nearby computer has an Internet connection. The nearby computer can access a first database of forms and a second database having information about the user of the scanner (such as a service provider's subscriber information database). The nearby computer accesses an electronic version of the paper form from the first database and auto-populates the fields of the form from the user's information obtained from the second database. The nearby computer then emails the completed form to the intended recipient. Alternatively, the computer could print the completed form on a nearby printer.

[00353] Rather than access an external database, in some embodiments, the system has a portable scanner that contains the user's information, such as in an identity module, SIM, or security card. The scanner provides information identifying the form to the nearby PC. The nearby PC accesses

the electronic form and queries the scanner for any necessary information to fill out the form.

19.2. Business Cards

[00354] The system can be used to automatically populate electronic address books or other contact lists from paper documents. For example, upon receiving a new acquaintance's business card, a user can capture an image of the card with his/her cellular phone. The system will locate an electronic copy of the card, which can be used to update the cellular phone's onboard address book with the new acquaintance's contact information. The electronic copy may contain more information about the new acquaintance than can be squeezed onto a business card. Further, the onboard address book may also store a link to the electronic copy such that any changes to the electronic copy will be automatically updated in the cell phone's address book. In this example, the business card optionally includes a symbol or text that indicates the existence of an electronic copy. If no electronic copy exists, the cellular phone can use OCR and knowledge of standard business card formats to fill out an entry in the address book for the new acquaintance. Symbols may also aid in the process of extracting information directly from the image. For example, a phone icon next to the phone number on the business card can be recognized to determine the location of the phone number.

19.3. Proofreading/Editing

[00355] The system can enhance the proofreading and editing process. One way the system can enhance the editing process is by linking the editor's interactions with a paper document to its electronic counterpart. As an editor reads a paper document and scans various parts of the document, the system will make the appropriate annotations or edits to an electronic counterpart of the paper document. For example, if the editor scans a portion of text and makes the "new paragraph" control gesture with the scanner, a computer in communication with the scanner would insert a "new paragraph" break at the location of the scanned text in the electronic copy of the document.

19.4. Voice Annotation

[00356] A user can make voice annotations to a document by scanning a portion of text from the document and then making a voice recording that is associated with the scanned text. In some embodiments, the scanner has a microphone to record the user's verbal annotations. After the verbal annotations are recorded, the system identifies the document from which the text was scanned, locates the scanned text within the document, and attaches the voice annotation at that point. In some embodiments, the system converts the speech to text and attaches the annotation as a textual comment.

[00357] In some embodiments, the system keeps annotations separate from the document, with only a reference to the annotation kept with the document. The annotations then become an annotation markup layer to the document for a specific subscriber or group of users.

[00358] In some embodiments, for each capture and associated annotation, the system identifies the document, opens it using a software package, scrolls to the location of the scan and plays the voice annotation. The user can then interact with a document while referring to voice annotations, suggested changes or other comments recorded either by themselves or by somebody else.

19.5. Help In Text

[00359] The described system can be used to enhance paper documents with electronic help menus. In some embodiments, a markup layer associated with a paper document contains help menu information for the document. For example, when a user scans text from a certain portion of the document, the system checks the markup associated with the document and presents a help menu to the user. The help menu is presented on a display on the scanner or on an associated nearby display.

19.6. Use with displays

[00360] In some situations, it is advantageous to be able to scan information from a television, computer monitor, or other similar display. In some embodiments, the portable scanner is used to scan information from computer monitors and televisions. In some embodiments, the portable

optical scanner has an illumination sensor that is optimized to work with traditional cathode ray tube (CRT) display techniques such as rasterizing, screen blanking, etc.

[00361] A voice capture device which operates by capturing audio of the user reading text from a document will typically work regardless of whether that document is on paper, on a display, or on some other medium.

19.6.1. Public Kiosks and Dynamic Session IDs

[00362] One use of the direct scanning of displays is the association of devices as described in Section 15.6. For example, in some embodiments, a public kiosk displays a dynamic session ID on its monitor. The kiosk is connected to a communication network such as the Internet or a corporate intranet. The session ID changes periodically but at least every time that the kiosk is used so that a new session ID is displayed to every user. To use the kiosk, the subscriber scans in the session ID displayed on the kiosk; by scanning the session ID, the user tells the system that he wishes to temporarily associate the kiosk with his scanner for the delivery of content resulting from scans of printed documents or from the kiosk screen itself. The scanner may communicate the Session ID and other information authenticating the scanner (such as a serial number, account number, or other identifying information) directly to the system. For example, the scanner can communicate directly (where "directly" means without passing the message through the kiosk) with the system by sending the session initiation message through the user's cell phone (which is paired with the user's scanner via Bluetooth™). Alternatively, the scanner can establish a wireless link with the kiosk and use the kiosk's communication link by transferring the session initiation information to the kiosk (perhaps via short range RF such as Bluetooth™, etc.); in response, the kiosk sends the session initiation information to the system via its Internet connection.

[00363] The system can prevent others from using a device that is already associated with a scanner during the period (or session) in which the device is associated with the scanner. This feature is useful to prevent others from using a public kiosk before another person's session has ended. As an

example of this concept related to use of a computer at an Internet café, the user scans a barcode on a monitor of a PC which s/he desires to use; in response, the system sends a session ID to the monitor that it displays; the user initiates the session by scanning the session ID from the monitor (or entering it via a keypad or touch screen or microphone on the portable scanner); and the system associates in its databases the session ID with the serial number (or other identifier that uniquely identifies the user's scanner) of his/her scanner so another scanner cannot scan the session ID and use the monitor during his/her session. The scanner is in communication (through wireless link such as Bluetooth™, a hardwired link such as a docking station, etc.) with a PC associated with the monitor or is in direct (i.e., w/o going through the PC) communication with the system via another means such as a cellular phone, etc.

PART IV – SYSTEM DETAILS

[00364] In some embodiments, access to associated electronic materials and actions is achieved through a "markup" layer or process (discussed in further detail in section 5). The markup layer determines how the system behaves and what the user experiences when he captures specific content and/or specific locations within a document. One result of applying markup to a capture (or group of captures) can be to present the user with a "context menu" of actions or choices associated with the capture(s). This menu is somewhat analogous to the right-click menu that is conventionally invoked by right-clicking with a computer mouse at a particular location in an electronic (computer-based) document. In some embodiments, the context menus depend on the individual document, location within the document, specific text captured, nature of scanning (e.g., gestures), user history, and/or permissions.

[00365] In some embodiments, the system displays these menus on a display optionally integrated in the capture device itself. In some embodiments, the system displays these menus on an associated device in the possession of, or near to, the user - for example, on the display of the user's mobile phone (which may have the scanning components integrated),

on the display of the user's laptop or desktop computer, on the display of the user's PDA (Personal Digital Assistant). In some embodiments, the system projects the menu directly onto the rendered document.

[00366] Figure 4 illustrates the interaction between a user, paper document, and the system in one embodiment. The document 401 contains a person's name 402. The user 420 has a connected personal computer 410. The computer has connected to it a scanning device 411 and a mouse 413. By scanning the person's name 402 in the document 401, the user 420 causes the personal computer 410 to display a context menu 412. The context menu 412 contains entries that are relevant to the location scanned, such as "email George" and "phone George." By selecting these entries in the context menu 412 using the mouse 413, the user 420 can contact the person whose name was captured using the scanning device 411.

[00367] In some embodiments, context menus have various default actions. These are actions which can be explicitly selected or de-selected by the user for a particular scan or group of scans, but these actions can also be set to occur if the user does not respond when presented with the context menu by either going on to make additional scans without selecting an item from the menu or allowing the menu to time-out by a period of inactivity. One example of a default scan action is to place a record of the scan in the user's chronological history, and optionally to record that this menu has not yet received any response from the user.

Supplemental markings

[00368] In some embodiments, the described capture device is able to capture and process text data, machine-readable data (such as barcodes, machine-recognizable shapes, and digital codes embedded in human-recognizable icons and shapes), and audio data (e.g., voice). Although not every embodiment will have all of these abilities (in some cases a capture device will only employ one), remarkable interactions are possible when these capabilities are combined.

[00369] In this document, the phrases 'scanner', 'scan' etc are often used but, as in earlier sections of the document, the descriptions should be taken to

apply to optical capture, audio capture, and other methods of capturing data from the rendered document, including combinations of such methods.

[00370] In some embodiments, the described scanner device recognizes data rendered as text and supplemental markings that qualify or enhance that text. In some embodiments, the scanner device has the ability to recognize certain icons and graphics (for example, shapes, underlining, special fonts and font styles, icons, pictures, and symbols), or embedded or associated digital data or codes (1-dimensional and 2-dimensional barcodes, data embedded in images as watermarks, certain text keywords, and numbers).

Markings Acted On By Users

[00371] In some embodiments, it is useful if supplemental markings can be recognized both by the scanner device and by the human user. One example is text in a special color indicating that supplemental electronic content or material is available for, or associated with, a region of colored text. For example, words or phrases in a school textbook for which there are definitions available in the book's glossary, at the publisher's website, or on a companion DVD might be printed in red. When the scanner device captures these words in red the system may perform special actions - for example, put the scanned word in the student's vocabulary list, or take the student's browser to the web page containing the definition of the scanned word or phrase.

[00372] As another example text printed in blue ink could indicate that there are related materials available on the Internet. In some embodiments, the text itself need not be a URL or web address (such as "www.google.com"). Text, whether associated with supplemental markings or otherwise, doesn't need to carry any additional information in order for the described system to make use of it. Rather, the system can maintain or establish an association between a fragment of text and related material based on the content of the text itself - in essence, each fragment of text can be its own link. In the above example, the system doesn't need to recognize the special color of the text when it is captured. In some embodiments the captured text content itself is sufficient to identify a document and/or a

location within a document, and any associations or actions appropriate to the capture can thus be determined and acted upon. Thus, while a capture device may have the ability to recognize supplemental markings, in some instances the supplemental markings are only present to inform the system user that special content or actions are associated with a particular piece of text, and are not necessary for the system to function.

[00373] In some embodiments, when the user scans text that is linked to web content with their capture device, their computer-based browser is automatically launched and directed to the appropriate material on the network. The user's browser might further be instructed how to render this particular piece of content - for example to open a new window if the user's browser is already running. This is one example of how material in paper documents can cause applications and actions to occur on a user's computer or other devices associated with the user. In some embodiments, the details as to which application to launch, how to launch it, and what content to direct it to are described in the markup layers associated with a rendered document.

[00374] Figure 5 is a flow diagram showing steps typically performed by the system to perform an action in response to a text capture. In step 501, the user captures a text sequence from a rendered document. In step 502, the system refers to the loaded markup for the current document to determine what action to take. In step 503, if the action is opening a browser window on a connected personal computer, the system continues to step 505 and launches the browser on the user's personal computer, else the system continues to step 504 and performs the other specified action. Following either step 504 or 505 the action is complete and the system waits for the next scan by the user.

[00375] In some embodiments, a rich interactive experience can be created for a paper document with no special actions or efforts on the part of the document's author, publisher, or printer. The markup, which in part creates the interactive experience, is of a generic nature and can be applied to any document, entire regions of documents, or all documents of a certain type. For example, the system might employ a rule like, "for captures of any

text containing a person's name, check whether this name is in the user's address book." This rule can then be applied to ordinary documents not necessarily created or intended for use with the system.

[00376] Figure 6 is a flow diagram illustrating the steps taken by the system to find appropriate markup for a document in one embodiment. In step 601, the user scans some text in a document. In step 602, the system determines if enough context is known about the current document to consult the markup layer. If there is not enough context information, the system goes to step 603 and prompts the user to scan a portion of the text that will provide the proper context, and then returns to step 602 to repeat the context determination. Once enough context is known, the system proceeds to step 604 where a lookup is performed to determine if the publisher of the document has made markup available for the document. In step 605 a test is performed to determine if publisher markup was found. If publisher markup was found, the search is complete and the markup is used, otherwise the system continues to step 606 where the user's preferences are consulted for markup relevant to unknown documents. In step 607, a test is performed to determine if markup was found in the user's preferences, and if so the search is complete. If markup was not found in the user's preferences, the system continues to step 608 and loads a default set of markup. The search for markup then completes.

[00377] In some embodiments, the markup layer can have a bi-directional effect. In the process of rendering or printing a document, the markup layer can be consulted to determine what elements or regions of the document being rendered are the subject of markup data and thus should receive special treatment or create special actions when a user scans those elements or regions in the rendered document. These elements and regions can then be rendered specially, for example by using special colors, fonts, or iconic symbols in the margin to indicate to the user or other parts of the system the presence of interactive elements in the document and actions associated with these elements or regions. Thus, regions of documents can be specially rendered to reflect the markup processes that will eventually be applied when a user scans the rendered version.

[00378] In some embodiments, supplemental markings in a document are meaningful to the user, to the scanner device, or to both. For example, text in red can indicate a glossary entry to the user, while this relationship or association can be determined by the scanner device or the system in the process of capturing and/or interpreting the scan (for example, a sensor in the scanner device could distinguish that the text being scanned was red in color). However, if there is a markup process employed, the association between captured text and entries in a glossary can optionally be determined through or during a markup process or layer. In some embodiments, this association is made when the system parses captured text and consults a database or list of associated meanings or actions. In the above example of glossary entries, the words found in the book's glossary could comprise such a list, or alternatively could be elements in a database. Thus, while the user potentially saw and responded to special markings (the special text color), the system determined by other means that additional material was associated with the scan.

Iconic Markings

[00379] In some embodiments, supplemental markings used with the described system are iconic - e.g., they are recognizable shapes, symbols, or graphics which have meaning to users. For example, a small head and shoulders silhouette printed in the margin beside a column of text may represent that it is possible to contact the author of this text.

[00380] Further information can be conveyed to the user and/or the system itself in several ways. In some embodiments, the above-mentioned silhouette icon would itself encode contact data (e.g., name, address, phone number, etc.), for example as a 2-dimensional barcode inside the icon, or by otherwise encoding data in the printing of the icon. Thus when the icon is optically scanned the required contact information is determined directly from the scanned data itself. In some embodiments, a checksum or other error-correcting means is employed as well.

[00381] In some embodiments, supplemental markings are associated with a potential purchase. For example, a printed article about a new book

might include a supplemental margin mark in the form of a dollar-sign - which tells the user that the item (in this case a book) about which they are reading is available for immediate purchase. In some embodiments, the user can then scan the dollar-sign icon to effect a purchase of the book being described. The icon (here a dollar-sign) may include embedded or associated digital data to assist in completing the transaction. This data may include, among other things, the item price, the item description, delivery options, a checksum or other security data, an identifier unique to this item, to this offer of this item, to the specific printing in this publication of this offer of this item, or even unique to this user's specific copy of this publication of this offer of this item.

[00382] In some embodiments, an icon representing the availability of supplemental web-based content appears in the margin of a book (for example, an icon of a computer monitor might be used). The icon tells the user that supplemental information is available. The system can be configured (for example in the markup layer) to send the user's browser to this associated web location if they scan any text in near proximity to the icon (for example, anywhere in the paragraph adjacent to the icon) or the icon itself.

[00383] In some embodiments, either or both the technique of scanning an icon directly, and the technique of scanning document text associated with an icon, can be used to initiate an action. The icon indicates to the user that supplemental material or actions are available, and these materials or actions can be accessed either by scanning the icon itself or by scanning rendered material near the icon.

[00384] In some embodiments, icons have names, tags, labels, or descriptions printed with them. The dollar-sign icon might, for example, have the word "purchase" printed under it. A scanner device may optionally scan and interpret this label as well as, or, optionally, in lieu of, scanning and/or processing the icon itself. In some embodiments the scanner device processes both the icon and the associated label, using the multiple types of data to qualify and/or confirm the scan.

[00385] In some embodiments where scanning can be done by voice, labels also serve the purpose of indicating what a user should speak to effect the action. In the above example of a dollar-sign icon having the label "purchase," they could speak the word "purchase." A limited number of "canonical" symbols are established and standard labels are associated with these. In the above example, users learn that the word "purchase" will initiate a purchase process if spoken while scanning near an item that can be purchased - whether or not an icon is present.

[00386] Figure 7 is an example document illustrating several forms of supplemental markings used in some embodiments. The document 701 contains a title 707. This title may be recognized by the system, and in some embodiments scanning the title 707 is sufficient to identify the document to the system. The document also contains a check-mark icon 702 indicating that the publisher has certified that this document is compatible with the system. The check-mark icon 702 also contains a bar code and a document reference number, any of which can be scanned to identify the document to the system and set the current context. The document also contains a silhouette icon 703 next to a person's name, which indicates that the named person can be contacted by scanning the icon, or text near the icon. The icon also contains a label "Contact" which can be spoken into the capture device to contact the named person. The document also contains a computer monitor icon 704 that indicates that more information about the adjacent text is available online by scanning the icon, text near the icon, or speaking the label "More Info" under the icon. The document also contains an advertisement with a dollar sign icon 706, indicating that the advertised item can be purchased using the system. Finally, the document contains a page label 705, that is in a format that the system recognizes to set the current context that that page.

[00387] In some embodiments, the system introduces indirection in supplemental markings. Thus encoded supplemental information (for example a 2-dimensional barcode associated with a purchase icon) may be partly or wholly comprised of a pointer or other reference data that identifies where additional information, or associated data, may be found. In the case

of the silhouette contact data example, the encoded data might be a unique numeric identifier or a URL that directs the system to necessary resources (for example, the contact data of the author of the printed article). In the case of the dollar-sign purchase example, the icon might include or be associated with an identifier or URL pointing to additional purchase data such as the current price, inventory status, or shipping options. In some embodiments, the indirection is comprised of a world wide web URL so that the system can direct the user's browser to a web page where information is available or where the transaction can be completed. Using indirection with supplemental data in documents helps to overcome the static nature of printed documents. Thus the information related to purchasing an item can be changed or updated, even though the purchase is initiated from a static document.

[00388] In some embodiments, scanning an icon indicates to the system that the user wants to obtain author contact information or to contact the author of an article. If the icon itself doesn't encode the specific information, the system checks that it knows the user's context - for example, by noting that the user scanned 2 minutes ago in an article that has associated author contact data - and the system offers the appropriate contact opportunity to connect the user to the author associated with that specific content or context. Alternatively, if the context is not known, the system indicates to the user that it requires additional context data - and when this data is supplied, for example by the user scanning additional material in the article, the requested contact opportunity is realized (e.g., by an offer to send an email, storing contact data in the user's address book, or dialing the phone).

Markings Acted On By The System

[00389] In some embodiments, supplemental markings in a rendered document are recognized and processed by the system, but are ignored or unseen by the user. For example, all of the terms that appear in a textbook glossary might be printed both as ordinary text, and with an invisible, machine readable code that generally isn't noticed by the user but is picked up when a document is scanned. Such markings can be printed in invisible ink, encoded in visible magnetic ink, or encoded in subtle details of ordinary printing (for example by introducing slight variations of pixel locations, color, or intensity).

In some embodiments a word or phrase appears underlined to the user, but the underlining actually includes embedded data that is machine readable - for example, as a miniature barcode. Many of these and other data-embedding techniques can employ technologies that are well-known in the art of printing (for example, they can make use of the various technologies used for "watermarking" documents which are to be printed or displayed).

[00390] Figure 9 illustrates supplemental markings using text styles and other system-readable information. The figure shows a document 901, which contains a first paragraph of text 902 that is printed in red to indicate to the user that selecting words in that paragraph will cause the words to be looked up in a glossary that accompanies the document. The document 901 contains a second paragraph 903 that is printed in blue text indicating that the user may select words in that paragraph to view supplemental information about the paragraph available on the Internet or a companion DVD to the document. The document 901 contains a third paragraph 904 that is underlined using an underlining method that allows the system to encode information in the underline that is not visible to a human reader. In the depicted example, the underlining encodes a web link 905 that directs the system to a specified web site if the user scans the underlined text.

[00391] In some embodiments, supplemental markings can be employed to assist the system in disambiguating captured data. For example, a specific instance of a frequently used word can be made unique and distinguishable by adding supplemental markings. Thus the system can determine precisely which of the many occurrences of this word is being scanned.

[00392] In some embodiments, one or more occurrences of a frequently used, ambiguous word or phrase leads to a specific and well-defined behavior by virtue of being associated with supplemental markings. For example, certain occurrences of a word that appears in a glossary can be rendered with a supplemental mark so that capturing any of these occurrences of this word or phrase causes the system to display the definition from the glossary.

[00393] In some embodiments, similar behavior is accomplished without supplemental markings by including instructions or data in the markup layers

or processes which cause the definition of a particular word or phrase to be displayed whenever it is scanned - even though the scan is otherwise ambiguous. [See also keywords - 5.2] In some embodiments, this is handled by programming the system to handle ambiguous scans in a specific way. For example, the system can be instructed to always look for a vocabulary or glossary entry if a capture has fewer than three words, or if a capture is generally ambiguous. Thus special actions can be programmed or set for various circumstances encountered in scans. Some example rules include:

[00394] If the scan begins with capital letters, check to see whether it is a proper name.

[00395] If a scan is a proper name, check whether it is a person's name.

[00396] If a scan is a person's name, see whether this person is in the user's address book.

[00397] If a scanned name is in the user's address book, offer to contact this person by sending an email to them, dialing their phone number, taking a voice memo for delivery to them, etc.

[00398] If the scanned name is a proper name and is not in the user's address book, offer to add this person to the user's address book (or do so by default).

[00399] If the scanned name is a proper name and the system has data about one or more people, places, etc having this name, present, or offer to present this data to the user (e.g., would you like to email [proper name]).

[00400] If a scan is very short (for example, 1 or 2 or 3 words), treat it as a special query.

[00401] If a scan is very short check whether a definition is available for this word or phrase - optionally display the definition and optionally add it to the user's vocabulary list.

[00402] If a scan is very short, optionally combine it (or consider combining it) with subsequent scans that follow (or that preceded) within a short period of time, occurred in the same font, etc. This could automatically

detect, for example, when the user is selecting multiple items from a list of items, or scanning items which occur near each other in a document.

[00403] In some embodiments, a region of text is rendered so as to indicate the extent that a user would ideally capture. For example, the extent of the indicated region might correspond to the amount of text that needs to be captured in order for the scan to be unambiguous. In some embodiments, the extent of the marking indicates the amount of text which will be effectively captured by a scan of a smaller portion. For example; a 100-word quote might appear in green ink, indicating that if the user scans some identifying portion of the quote, all 100 words will be captured into their life library (see sections 11.3.2 and 12.2.1).

[00404] In some embodiments, users instruct the system to perform the above "capture a quote" function by setting a configuration option. For example, the behavior of the system can be set, for some or all users, as a default behavior or as a user option, to extend a capture forwards and/or backwards in a text region to terminate at particular points inside or outside the region of the capture itself. These end-points might be any of: the beginning and end of a quotation containing or overlapping the captured data, the region having the same font or typeface or color as the captured data, or the line, sentence, paragraph, or page containing or overlapping the captured data. These points can be set for material which is near to, surrounding, containing, contained by, or overlapping the captured data.

Functionality Without Markings

[00405] In some embodiments of the system the user doesn't need to scan an associated supplemental icon or visual indicator itself to obtain the various actions or results indicated as available by that icon. Even short fragments of written text are unique, especially within limited contexts like a specific article, volume of a periodical, or specific book and may optionally be further identified if the user has recently established their probable position within a sub-region in these documents (established, for example, by a recent scan in or near this sub-region). Thus, scanned text itself is used in some

embodiments to initiate actions, or to present the user with menu options when the user captures data near a supplemental marking.

[00406] In the purchase-the-described-book example above, it wouldn't be necessary for the user to scan the dollar-sign purchase icon itself. Rather, the system recognizes that the user scanned somewhere in proximity to this symbol, infers from this proximity and/or other data that a purchase might be desired, and presents to the user the option to make a purchase (for example on the display of their computer, on a display on the scanning device itself, on the display of the user's mobile phone, etc.). Thus the dollar-sign icon appearing in the margin as mentioned above can indicate to the user that they can purchase the item by scanning the icon itself, or by scanning content near the icon. In some embodiments of the system either or both of these methods are implemented.

Menu Items

[00407] In some embodiments, the interactive paper experience enabled by the system includes interactions with printed (as distinct from dynamically-displayed) menus. Thus publishers can print menus of possible choices or interactions in their documents, and users can select these menu items by scanning them. These menu items can be comprised of pure text items, pure graphical symbols, graphical symbols with embedded or encoded data, or graphical symbols and associated text descriptions. The scanner device is capable of recognizing the printed text, the graphical symbol, the embedded or encoded data, or combinations of the above. In embodiments where speech is an available form of capture, the user may read the printed words associated with a menu item, speak the name commonly associated with a menu item (for example, "purchase," or "information"), or say the number of the menu item to select a particular menu choice.

[00408] In some embodiments frequently used groups of printed menu items are employed. For example, advertisements which are made interactive by the use of the system may be accompanied by standard menu items such as, "purchase this now" (perhaps a dollar-sign icon), "request information about this" (perhaps a letter "i" in a circle), or "remember this for

later" (perhaps a "!" in a circle). These menu items might include a representative graphic, a printed text description, a number indicating menu position, or other means by which the user can indicate a choice.

[00409] Figure 8 illustrates a printed menu recognized by the system in some embodiments. The advertisement 801 contains a printed menu 802. The printed menu contains an entry 803 that contains a dollar-sign icon, a number "1" and a label "Purchase Now." In some embodiments the user may purchase the advertised item by scanning the icon, the number, or the label, or by speaking the label. The printed menu 802 also contains an entry 804 with an icon, number, and label that allow the user to request further information about the advertised product. The printed menu 802 also contains a third entry 805 with an icon, number, and label that allow the user to save the advertisement for later viewing. In some embodiments printed menus may contain only icons, numbers, labels, or any combination of these.

[00410] In some embodiments, certain elements of text in a printed document may have special significance when entered by a scanner device in the described system. For example, scanning a page number (optionally accompanied by the word "page," or surrounded by hyphens - for example "page 23" or "-23-") may indicate to the system that a bookmark should be placed at that page. Similarly, scanning a single word may indicate that the user would like to add that word to a vocabulary list or that the user would like to see a definition of that word. Thus, in some embodiments, the described system includes the ability to process and act upon scanned text, optionally modified by context data - including non-text elements included in a scan or group of scans.

Capabilities While Disconnected

[00411] In some embodiments, the scanner device can perform actions and provide information to the user even when it is disconnected from other elements of the system. In some embodiments, the scanned data can be stored for subsequent processing. In some embodiments, the system stores sufficient data and instructions on the scanner device to determine appropriate interactive behavior when the device is offline.

[00412] For example, consider a scanning device that incorporates a visual indication to the user when the material they are scanning is coming from a known context (e.g., the current document is recognized and known to be indexed), and optionally an additional visual indication when the specific location of the current scan within the current context is also recognized and known. Before the scanner device is disconnected from the network, the system can download to the scanner device information about the 1000 most popular current periodicals – determined by the individual user's history of scanning, and/or by using information from what other system users typically are scanning. In some embodiments, the downloaded information includes indices which tell when a particular scan is known to exist in a particular corpus of documents, when the scan is known to be unique in a corpus of documents (e.g., its location is known), and various markup data describing how the scanner device or system should behave if the user scans at particular locations within the 1000 periodicals for which data is being cached. Thus the user may experience some or all of the "interactive paper" experience, even though they are using their scanner device in a disconnected mode (see sections 3.5, 12.1.4, and 15.3).

[00413] In some embodiments, information downloaded to the scanner device further includes data about fonts, specific layouts and renderings of documents, specified interactive behavior for various sub-regions (e.g., markup data), permissions data, readership statistics, rankings, popularity of documents and sub-regions, annotations associated with documents, and/or data required to complete off-line purchases and other commercial transactions. Downloaded information includes those items of data which are available while the scanner device is connected to the system, and which have a reasonable probability of being useful to the user, the scanner device, or the scanner system while the scanner is disconnected.

[00414] In some embodiments, available data is ranked according to the probability that it will be required or used by the user, optionally including a weighting factor for how much storage space is required to store said information, in order to determine which data to download to the device. The most valuable data can then be given some degree of priority when allocating

download bandwidth, storage space available on the scanner device, and other resources.

[00415] For example, consider a scanner device with appropriate programmed behavior and a small display. The user is using this scanner device to scan data from an in-flight magazine on an airplane - hence the scanner device may be disconnected from any desktop-computer-based or network-based elements of the system. A dollar-sign icon appears in the margin of the in-flight magazine next to an article about a recently published book. The user scans the icon indicating his or her interest in possibly purchasing the book. The scanner device recognizes the dollar-sign icon representing a purchase.

[00416] In some embodiments, there is digital data printed under the icon in the form of a two-dimensional barcode. The scanner device reads this data and presents on the scanner display all the necessary steps and data to make a purchase, including the expected time for delivery, the vendor who will be providing the item, the item cost, tax if appropriate, and so forth. Thus the supplemental data is independently read and interpreted by the scanner device, and the scanner device has programmed behavior which allows a transaction to be initiated (and in some cases completed) while disconnected from any network or host.

[00417] In some embodiments, there is no digital data accompanying the dollar-sign, however, the scanner device uses internal logic and processing to detect that a purchase transaction has been initiated. This action is recorded for subsequent action. Additional context can be saved or maintained, such as the time of the (dollar-sign) scan and scan data related to prior and following scans. In some embodiments, the user then receives indication that a purchase transaction has been recognized (for example, a dollar-sign symbol can appear on the optional local display of their scanner device). If the scanner logic detects that the context of the dollar-sign/purchase scan is uncertain, a "Context Uncertain" indicator may be displayed, prompting the user to scan additional information about the purchase. The context may be

uncertain if, for example, no scan had been received in the 20 seconds immediately preceding the scan of the dollar-sign icon.

[00418] In some embodiments, a scanner device may not have available either the necessary capabilities or necessary data to complete transactions, information requests, or other similar actions when the scanner device is disconnected from any network or host. In these cases the scanner device may only capture (and optionally pre-process) scanned data. When the scanner device is later reconnected to a network or host, any captured data can be re-examined to determine whether further actions are required or appropriate. For example, in the case where the user scanned a purchase icon while disconnected, the system can examine the data associated with the purchase icon (and/or determine the context in which this icon was scanned), and can then initiate an interactive process with the user – for example on the user's personal computer via a browser connected to an e-commerce web site - where this delayed interactive process effects the completion of the intended purchase.

[00419] In some embodiments, the scanning device uses inexpensive data storage to carry substantial amounts of information with it for use while disconnected from the rest of the system. This leads to a blended experience that combines some of the online and offline experiences described above. For example, the system may determine that a particular user is likely to encounter some specific set of all commercial offers or transactions which are available at a particular time - for example, offers or purchase transactions which have appeared within the last 6 months in the 100 periodical publications most likely to be read by this user. While the scanner device has access to a network it can download and store information about this finite set of possible transactions so that, should the user encounter one of these transactions while disconnected, the scanner device has much or all of the necessary information to interact with the user and complete the transaction.

Marked Documents

[00420] One challenge of interactive paper documents is that some of a particular document's interactivity may depend on whether the described

system is "aware of" that particular document. For example, documents that have been indexed by the system, or have markup layers associated with them, are likely to be richer in interactions than documents lacking this relationship. Thus, some embodiments of the system may include a symbol or icon or graphic which indicates that the document, chapter, section, page, advertisement, periodical, book, or paper to which the symbol is attached is known, or is likely to be known, or will be known in the future to the system.

[00421] In some embodiments, publishers give users direct assurance that a particular document is known to, and compatible with, the system, by printing a special human-recognizable mark at one or more locations in a document. These system-compatible symbols may also be machine-readable, for example by the described scanning device - so that scanning such a mark sets a definitive context for the subsequent scanning about to occur, or for previous scanning which has already occurred in a document. In some embodiments, the system-compatible marks will be used with document sub-components. For example, they may appear printed in a particular advertisement, in association with a specific article, or next to an image. In some embodiments, each occurrence of the system-compatible mark is unique - thus the scanner device or system knows its context exactly when such a mark is scanned, and the user is assured that they are setting or establishing a known context when they scan such a mark. In some embodiments extra data, for example a document reference number, is printed close to the mark, and scanning the combination of the mark and the data establishes the context.

[00422] In some embodiments, the scanning device itself is capable of recognizing, recording, and/or responding to system-compatible symbols; such that the context can be recognized even when the scanner device is disconnected from other elements of the system. For example, when a user is traveling and their scanning device is disconnected from the network and/or disconnected from various other components of the system, the user might turn to the back page of the daily newspaper they just purchased, note the "system-compatible" icon printed next to the commercial barcode, and be assured that their activities within this document will be meaningful - for

example when they reconnect to the network, and/or when this document is included in the system (e.g., is indexed, and/or receives a markup layer, and/or has an electronic counterpart available to the user, etc.).

[00423] In some embodiments, a document is published for immediate use by system users before the document is indexed or otherwise integrated into the system and the results of some user interactions with the document are deferred. The publisher assures a system user that actions the user takes in such a document will be meaningful by pre-specifying that the document will become interactive. For example a user may be assured that portions of a document that the user highlights with the scanning device will be viewable later with its surrounding context (because the source document will have become available to the system), or that a user's interest registered in advertised products or in products mentioned within the document will eventually lead to an opportunity to purchase these products electronically (because any necessary additional markup or links to e-commerce web sites will become available, in particular via scan data accumulating in the user's Life Library (see section 16.1)).

Document Registry

[00424] In some embodiments, a registry (or multiple registries) of documents known to the system is maintained. Portions of this registry can be downloaded to the user's scanner device, or to a device near the scanner device, so that the user or scanner device can determine whether a particular document is or will be known to the system. In some embodiments, the registry includes the title of the included documents. This allows the user to scan the title of a document (for example by voice or by optical scan) and have the scanner device indicate (for example via a visible LED) that the document context is known - even if details about that context may not presently be available or known, or the document in question is not yet indexed by the system. In some embodiments, the registry also includes information that allows the scanner device to determine whether a particular document is known by scanning the barcode, or numbers associated with the barcode, printed on the document or on the shipping label used to deliver the document.

[00425] Printed barcodes generally include a set of machine-readable bars, and also a set of human-readable printed digits that are encoded as data in the bars. Companies using U.P.C. (Universal Product Code) bar codes on their products each have a Manufacturer Identification Number that uniquely identifies each company; this company prefix is also encoded in the U.P.C. Also encoded in the U.P.C. is the Item Number that uniquely identifies this particular item among those made by this manufacturer.

[00426] In some embodiments, the system maintains a registry of documents known to the system for which there is an associated UPC code. In some embodiments, the system also maintains a registry or list of UPC codes for documents specifically known to not be included in the system (for example, indexed by the system). This latter set of documents might include those whose publishers asked not to be included, or documents for which the indexing or maintenance fee has not been paid.

[00427] In some embodiments, all, or a portion of, the above mentioned document indices are placed on a device near to (for example, in direct communication with) a scanner device, or onto the scanner device itself. Thus the scanner device can determine whether a particular document is (or eventually will be) known to the system. In some embodiments, the scanner device determines whether a document is known to the system when the user directly scans the UPC code on a document. In the case of a voice scan/capture, the data input may be the spoken digits printed with the UPC code.

[00428] In embodiments where the known-document registry includes document titles, for example "The New York Times" or "Gone with the Wind", an alternative way to determine that a document is known to the system - and, optionally, to set the current context - is to capture its title.

[00429] Figure 10 is a data structure used in some embodiments for storing the known-document registry. The data structure 1001 contains a first column 1002 labeled "UPC" which contains a bar code associated with a known-document entry. The data structure 1001 contains a second column 1003 labeled "Title" which stores a title used to refer to a known-document

entry. The data structure 1001 contains a third column 1004 labeled "markup" which contains markup relevant to a known-document entry. The data structure 1001 contains a sample entry 1005 which shows a particular known document, its UPC code, its title, and the markup used to interact with the document.

[00430] In some embodiments, the extended environment of a document can be developed for that document without requiring any additional effort or action by the author or publisher. Since the elements of printed text and markings in ordinary printed documents form a rich set of document identifiers and location-within-document identifiers, interactive behavior can be superimposed on new or pre-existing documents by establishing markup rules and/or markup descriptions which describe how the system should behave when the user scans particular content and/or at a particular location.

[00431] In some embodiments, the interactions available with a document are determined by classifying the document in a hierarchical representation. In some cases, a document will not have any markup processes or data specific to that document, but will have behaviors provided by the system for a larger class of documents to which this individual document belongs. Some documents may only belong to the class "all documents," i.e., their interactions and behaviors are shared by every document or by every document not belonging to a more specific class. Thus, documents may participate in markup and behavior description in a model based on classes and inheritance. In some embodiments, documents may participate in default markup and/or default behavior descriptions.

An example implementation

[00432] The example below illustrates how such a system might provide a user with access to associated materials from a rendered document.

[00433] A user in his office is reading a book, which has icons printed in it to indicate to the reader that further electronic material is available. The user is particularly interested in one section and wishes to find out more. At the top of each page is printed the name of the publisher and the title of the book.

The user picks up his portable scanning device and scans this top line, followed by the heading of the section in which he is interested.

[00434] The scanner device sends the scanned images to the user's computer using a wireless link such as Bluetooth. Software on the computer performs image analysis and OCR of the scanned text and sends the recognized text for both scans to an internet-based service.

[00435] The internet-based service analyses the text using standard fuzzy matching techniques to identify the publisher of the document from a list of publishers in its database. Also listed in the database is the URL of the publisher's 'associated materials' web service. The internet-based service forwards the text to the publisher's service, where similar matching techniques are used to extract the title of the book and the heading of the section by comparing them to those known in the publisher's database.

[00436] The publisher's web service locates the portion of its web site containing associated materials, discussion forums, errata, further reading, author's comments etc and it sends the URL for that site back to the user. The aforementioned software running on the user's computer causes that web page to be retrieved and displayed.

Conclusion

[00437] It will be appreciated by those skilled in the art that the above-described system may be straightforwardly adapted or extended in various ways. While the foregoing description makes reference to particular embodiments, the scope of the invention is defined solely by the claims that following and the elements recited therein.

CLAIMS

We claim:

[c1] 1. A method in a computing system for providing supplemental electronic materials relating to a rendered document, comprising:

receiving an indication of a text capture operation performed by a user from the rendered document, the indication specifying text captured as part of the text capture operation;

retrieving supplemental electronic materials relating to the rendered document;

in response receiving the indication:

using the captured text specified by the indication to identify a position in the rendered document at which the indicated text capture operation was performed;

causing a portion of the retrieved supplemental electronic materials that relates to the identified position to be displayed to the user.

[c2] 2. The method of claim 1 wherein the rendered document is an academic textbook.

[c3] 3. The method of claim 1, further comprising identifying the rendered document based upon the captured text specified by the indication.

[c4] 4. The method of claim 1 wherein the retrieved supplemental electronic materials were developed after the rendered document was rendered.

[c5] 5. The method of claim 1 wherein the retrieved supplemental electronic materials were developed by a developer unrelated to the publisher of the rendered document.

[c6] 6. The method of claim 1, further comprising:
receiving an indication of a second text capture operation performed by a user from the rendered document, the indication specifying text captured as part of the second text capture operation;
in response receiving the indication of the second text capture operation:
using the captured text specified by the indication of the second text capture operation to identify a second position in the rendered document at which the indicated second text capture operation was performed;
causing a portion of the retrieved supplemental electronic materials that relates to the identified second position to be displayed to the user.

[c7] 7. A text capture device, comprising:
a visual input subsystem for capturing images of text sequences from rendered documents; and
a processing subsystem for extracting from images captured by the visual input subsystem both human-readable content and non-human-readable content.

[c8] 8. The text capture device of claim 7, further containing an audio input subsystem for capturing audio clips of users reading aloud from rendered documents.

[c9] 9. A method in a computing system for processing a visual capture operation, comprising:
receiving an indication of a visual capture operation performed from rendered document, the indication specifying both a text sequence captured as part of the capture operation and a supplemental marking captured as part of the capture operation; and

determining an action to perform in response to receiving the indication, based both upon the text sequence specified by the indication and the supplemental marking specified by the indication.

[c10] 10. The method of claim 9 wherein the supplemental marking captured as part of the capture operation is a non-textual symbol.

[c11] 11. The method of claim 9 wherein the indicated visual capture operation is performed by capture device, the method further comprising recognizing the supplemental marking in the capture device.

[c12] 12. The method of claim 9 wherein the indicated visual capture operation is performed by capture device, the method further comprising recognizing the supplemental marking in a computer system external to the capture device.

[c13] 13. The method of claim 9 wherein indicated visual capture operation is performed by a distinguished user, and wherein the determined action is displaying selected information on computer display visible to the distinguished user.

[c14] 14. A method in a computing system for processing a textual capture operation, comprising:

receiving an indication of a textual capture operation performed from rendered document, the indication specifying both a text sequence captured as part of the capture operation and a visual characteristic of the captured text sequence; and

determining an action to perform in response to receiving the indication, based both upon the text sequence specified by the indication and the visual characteristic specified by the indication.

[c15] 15. The method of claim 14 wherein the visual characteristic of the captured text sequence specified by the received indication is a font.

[c16] 16. The method of claim 14 wherein the visual characteristic of the captured text sequence specified by the received indication is a font size.

[c17] 17. The method of claim 14 wherein the visual characteristic of the captured text sequence specified by the received indication is a color.

[c18] 18. The method of claim 14 wherein the visual characteristic of the captured text sequence specified by the received indication is a formatting option.

[c19] 19. The method of claim 14 wherein the visual characteristic of the captured text sequence specified by the received indication is a border.

1/10

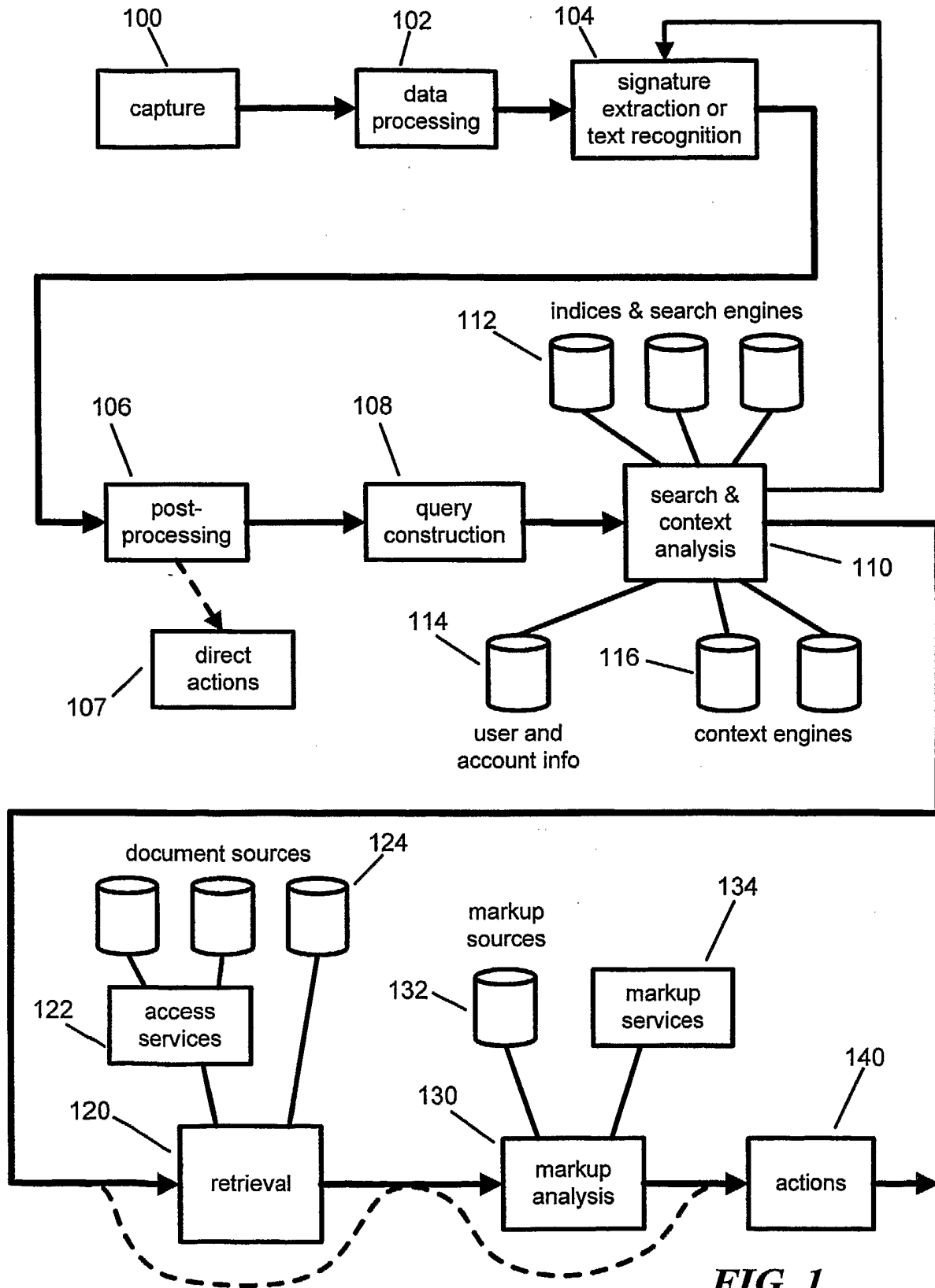


FIG. 1

2/10

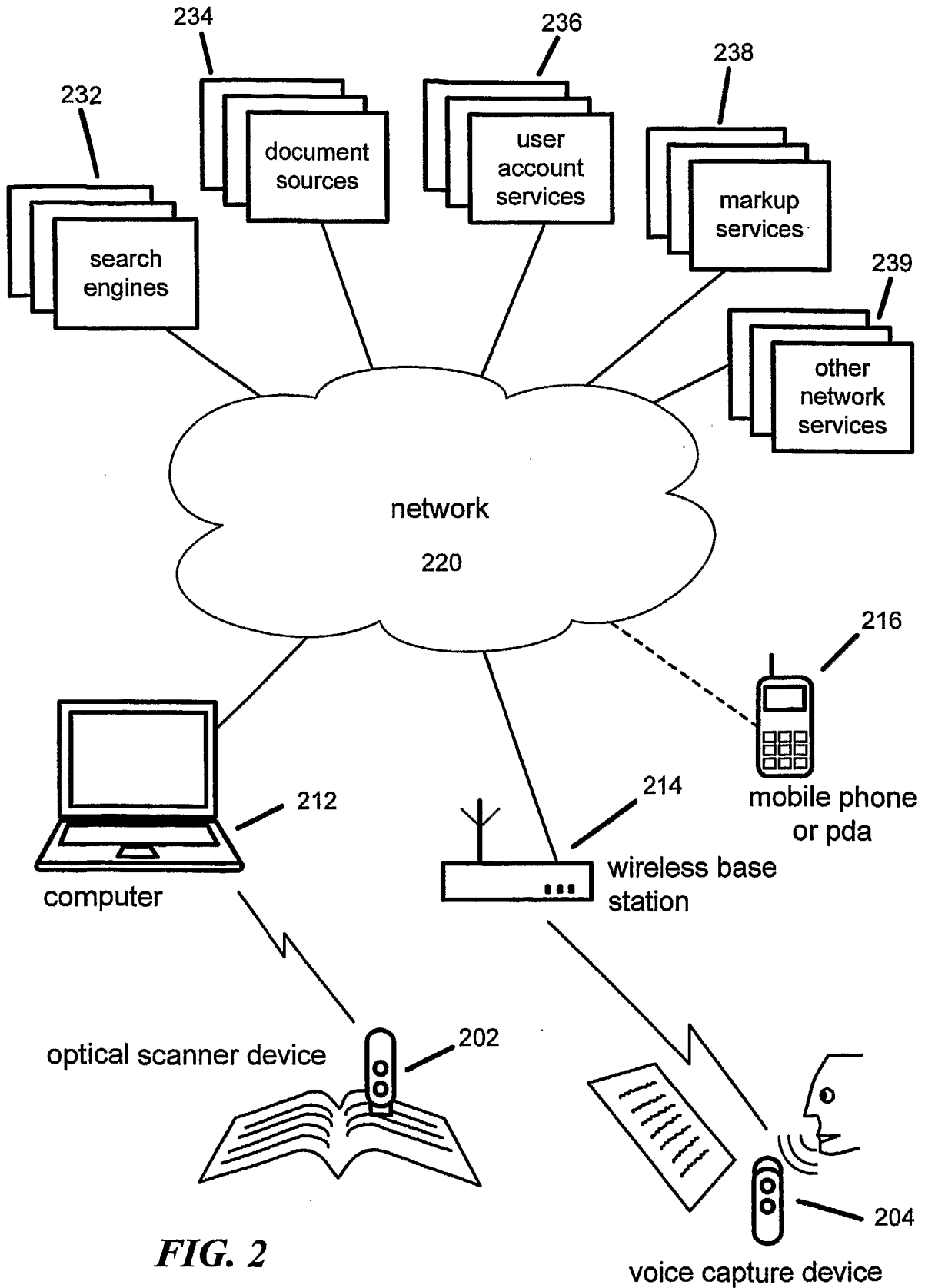


FIG. 2

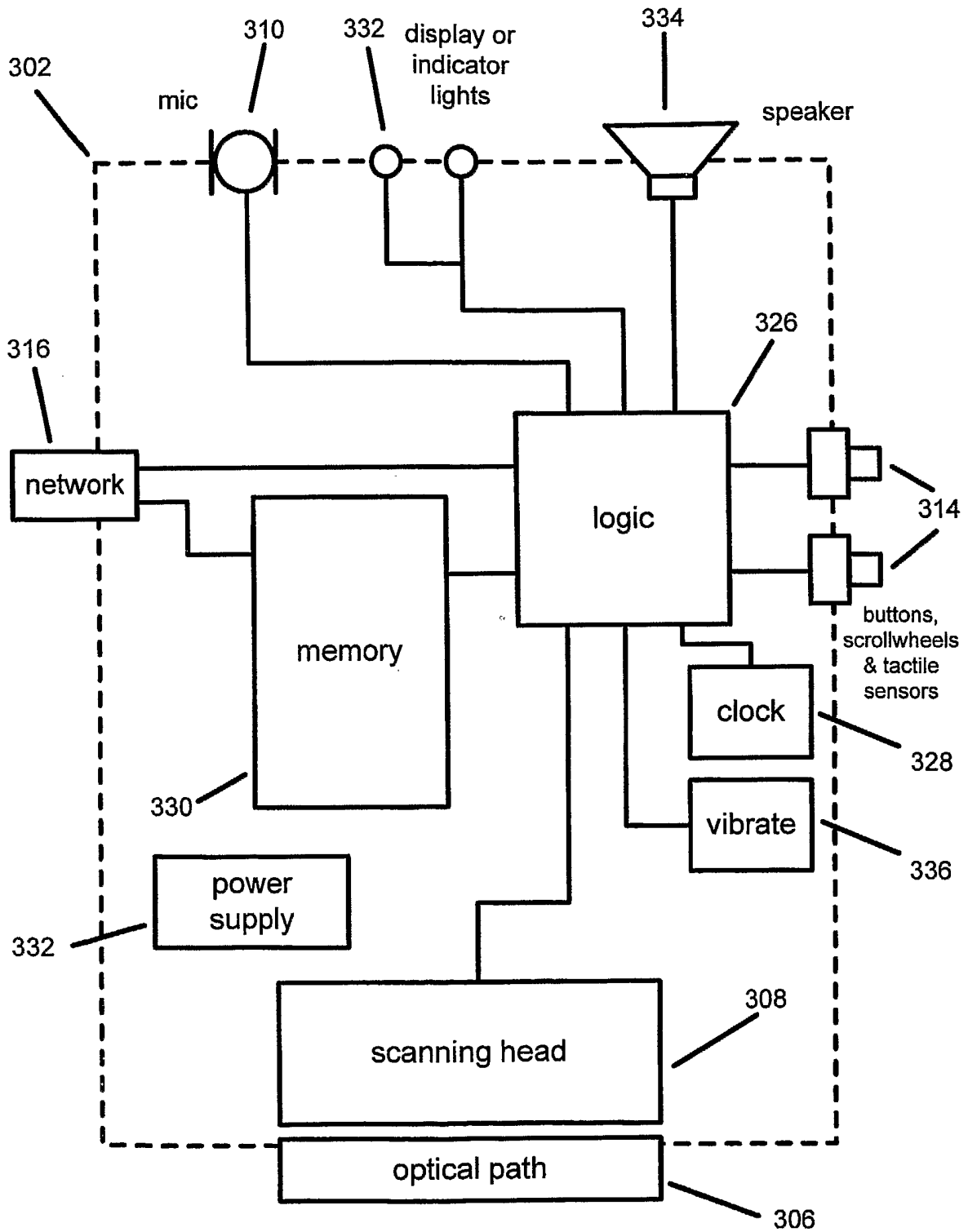


FIG. 3

4/10

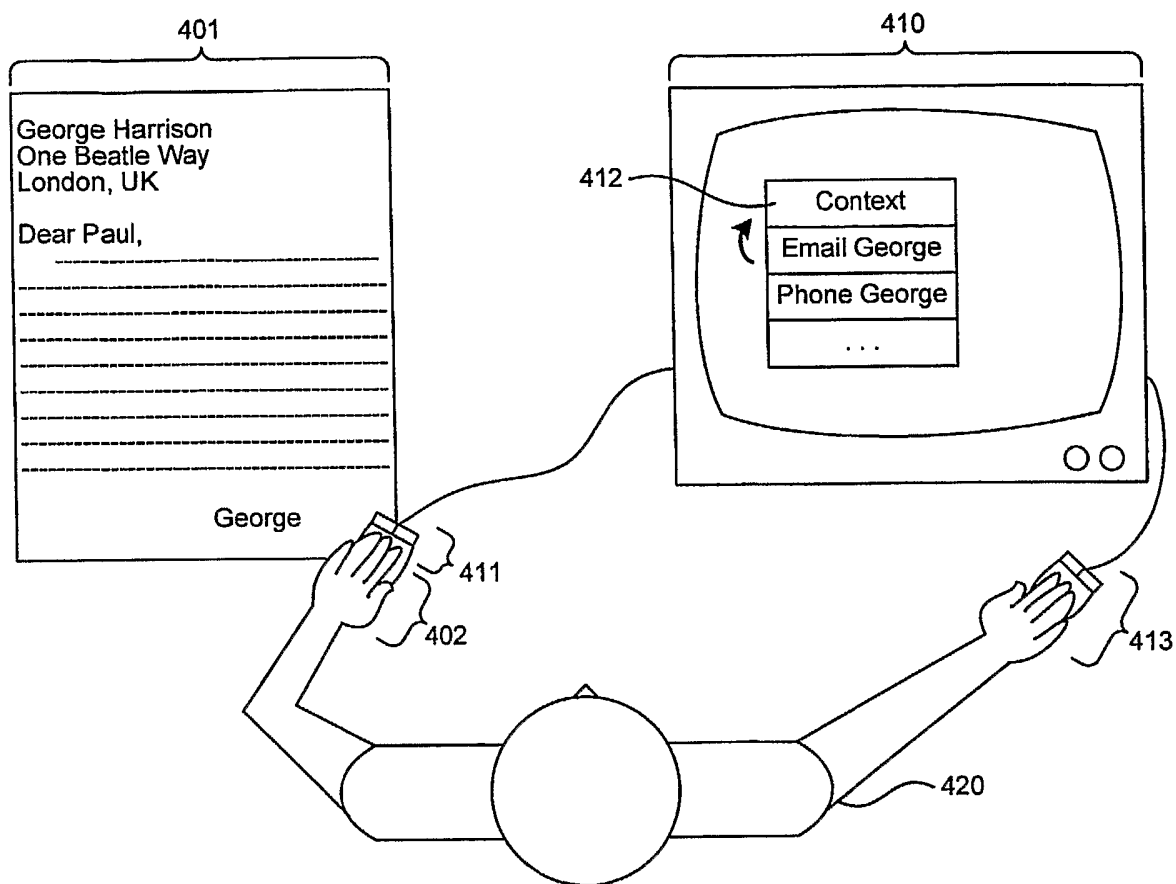


FIG. 4

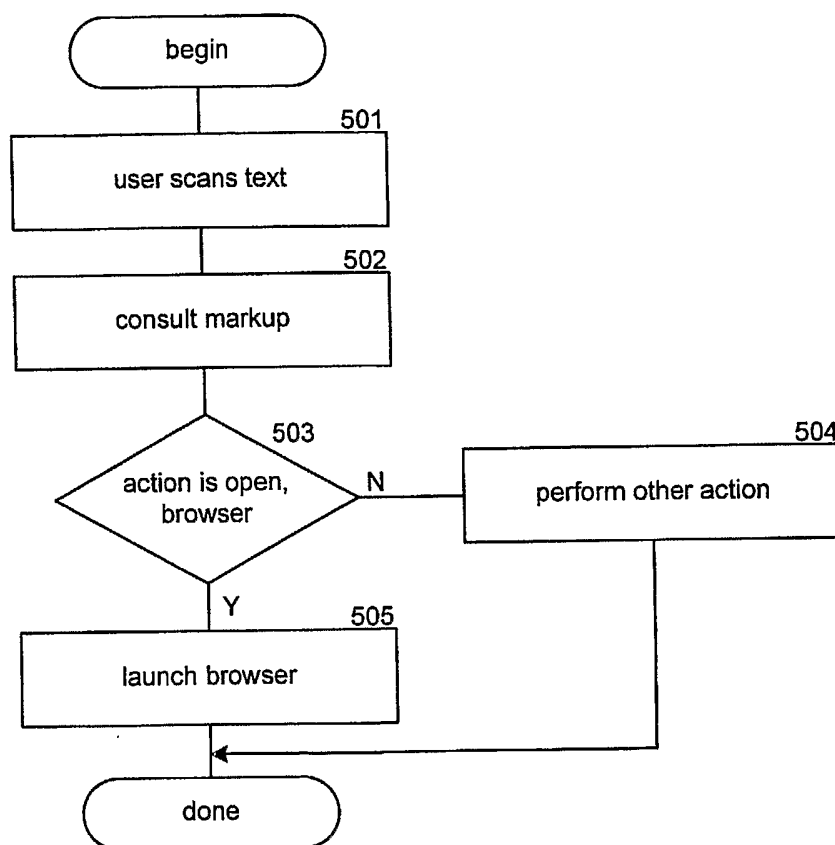


FIG. 5

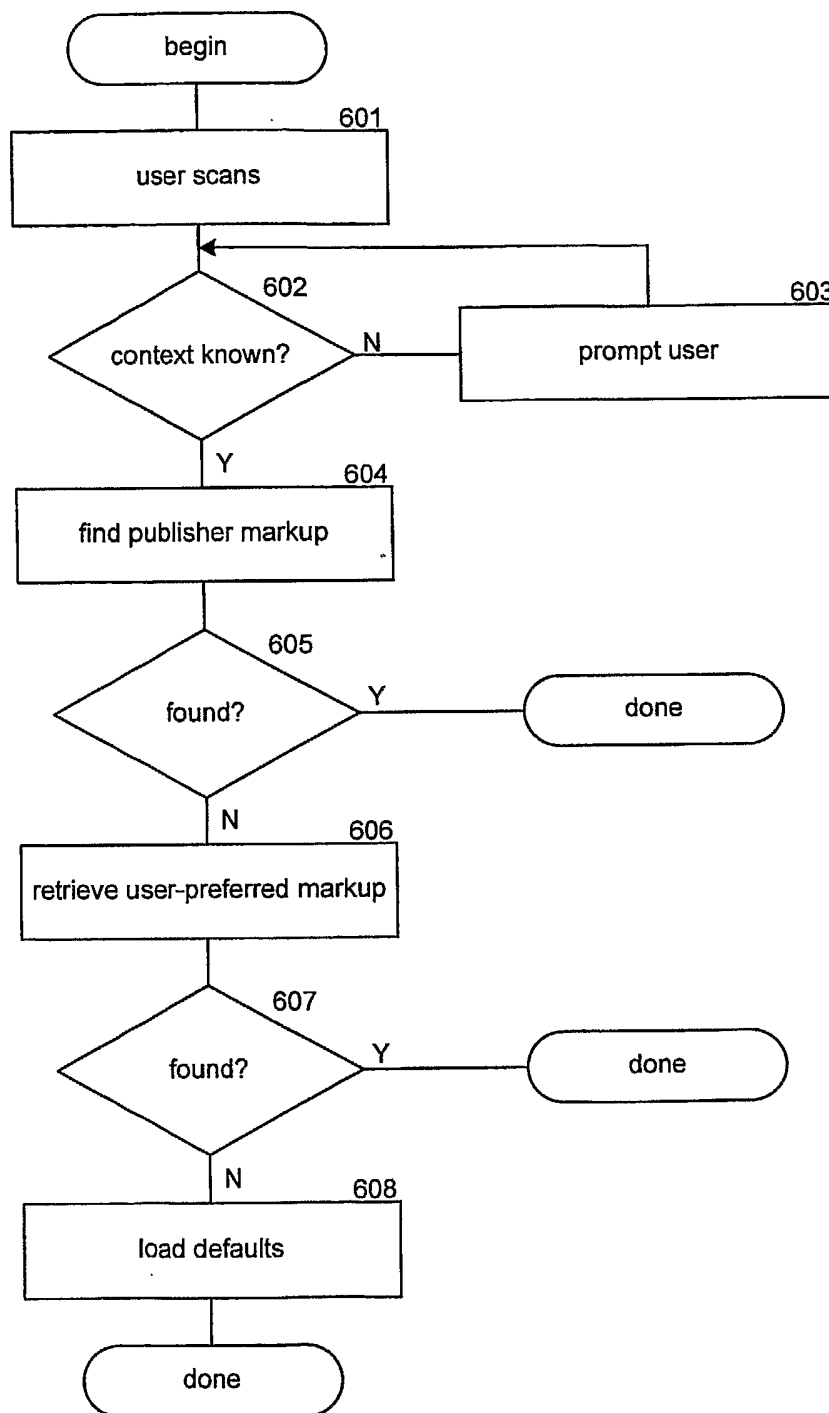


FIG. 6

7/10

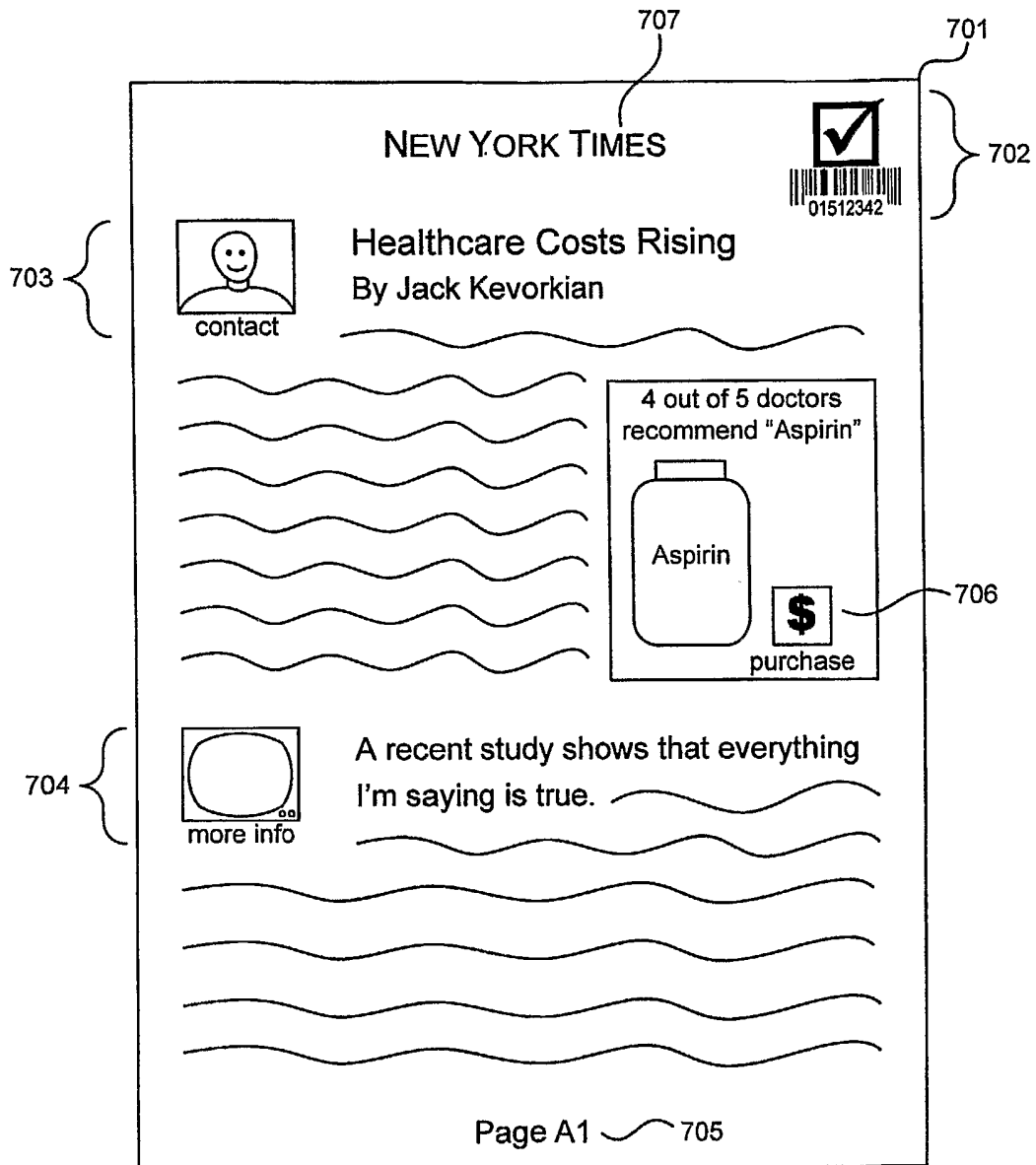


FIG. 7

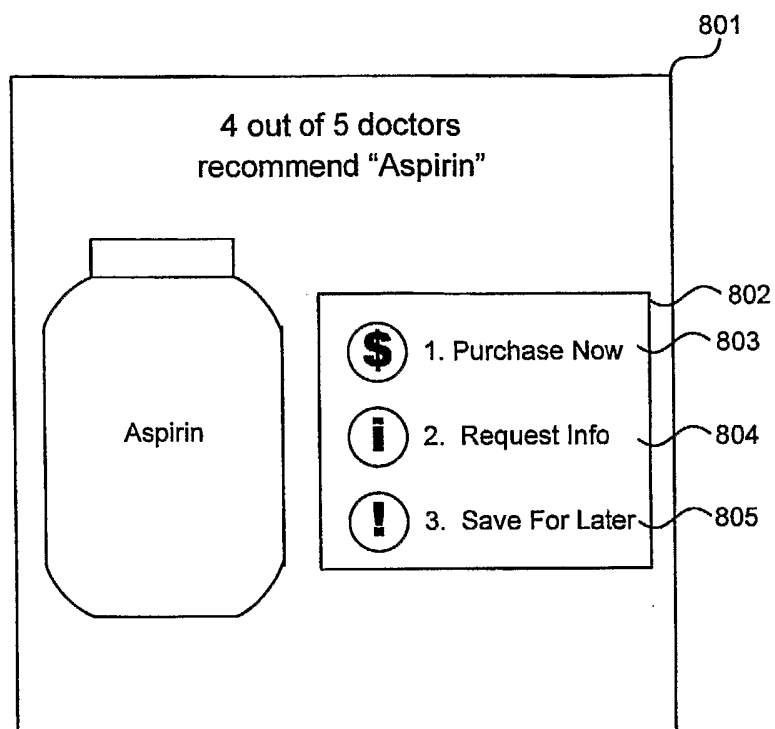


FIG. 8

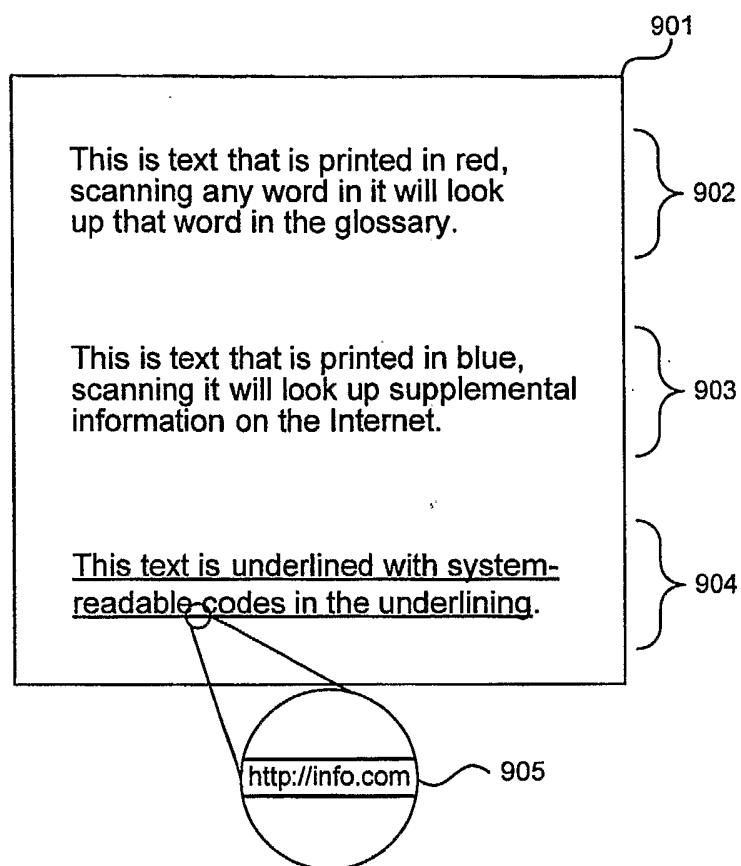


FIG. 9

10/10

A table with three columns and three rows. The columns are labeled 'UPC', 'Title', and 'Markup'. The first row contains the headers. The second row contains '1234567', 'Gone With The Wind', and a redacted area. The third row contains '...', an empty cell, and another empty cell. Reference numerals 1001-1005 point to various parts of the table.


UPC	Title	Markup
1234567	Gone With The Wind	
...		

FIG. 10