**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G06F 17/30 | A1

(11) International Publication Number: WO 98/04979

(43) International Publication Date: 5 February 1998 (05.02.98)

(21) International Application Number: PCT/GB97/01949

(22) International Filing Date: 18 July 1997 (18.07.97)

(30) Priority Data:
96305879.7 31 July 1996 (31.07.96) EP
*(34) Countries for which the regional or international application was filed:* GB et al.

(71) Applicant *(for all designated States except US)*: BRITISH TELECOMMUNICATIONS PUBLIC LIMITED COMPANY [GB/GB]; 81 Newgate Street, London EC1A 7AJ (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: LEGH-SMITH, Jon [GB/GB]; 32 Finchley Road, Ipswich, Suffolk IP4 2HU (GB). APPLEBY, Steve [GB/GB]; 23 Richard Avenue, Wivenhoe, Colchester, Essex CO7 9JQ (GB). FAIRMAN, Ian [GB/GB]; 28 Camden Road, Ipswich, Suffolk IP3 8JW (GB). EYLES, Sarah [GB/GB]; 37 Baronsdale Close, Ipswich, Suffolk IP1 4JY (GB).

(74) Agent: EVERSHED, Michael; BT Group Legal Services, Intellectual Property Dept., 8th floor, 120 Holborn, London EC1N 2TE (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

**Published**
*With international search report.*

(54) Title: DATA ACCESS SYSTEM

(57) Abstract

A method of automatically creating a database (112) on the basis of a set of category headings is described, where a set of keywords is provided for each category heading. The keywords are used by a processing platform (100) to define searches to be carried out on a plurality of search engines (130a, 130b, 130c) connected to the processing platform (100) via the Internet (120). The search results are processed by the processing platform (100) to identify the URLs embedded in the search results. The URLs are then used to retrieve the pages to which they refer from remote data sources in the Internet (120). The processing platform (100) then filters and scores the pages to determine which pages are the most relevant to the original categories. Internet location information for the most relevant pages is stored in the database (112).

