



- (51) **International Patent Classification:**
G06F 17/00 (2006.01) *G06F 3/14* (2006.01)
- (21) **International Application Number:**
PCT/US2013/070524
- (22) **International Filing Date:**
18 November 2013 (18.11.2013)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
61/728,100 19 November 2012 (19.11.2012) US
13/831,121 14 March 2013 (14.03.2013) US
- (71) **Applicant:** GLOBYS, INC. [—/US]; 705 5th Avenue South, Suite 700, Seattle, WA 98104 (US).
- (72) **Inventors:** CAZZANTI, Luca; 705 5th Avenue South, Suite 700, Seattle, WA 98104 (US). MEHANIAN, Cour-osh; 705 5th Avenue South, Suite 700, Seattle, WA 98104 (US). PENZOTTI, Julie; 705 5th Avenue South, Suite 700, Seattle, WA 98104 (US). DOWNS, Oliver; 705 5th Avenue South, Suite 700, Seattle, WA 98104 (US).

SCOTT, Doug; 705 5th Avenue South, Suite 700, Seattle, WA 98104 (US).

(74) **Agent:** WIEGAND, Jamie, L.; Frommer Lawrence & Haug LLP, 745 5th Ave., New York, NY 10151 (US).

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

(84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,

[Continued on next page]

(54) **Title:** UNSUPERVISED PRIORITIZATION AND VISUALIZATION OF CLUSTERS

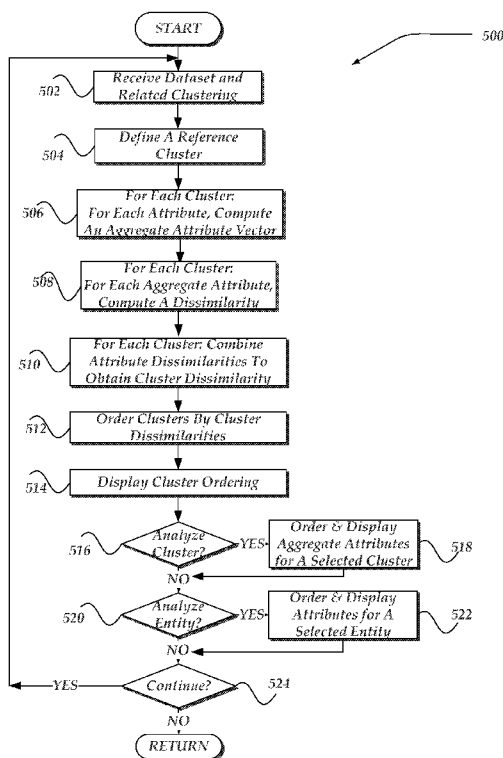


FIG. 5

(57) **Abstract:** Techniques are disclosed that automatically identify and order the most differentiated clusters from a given collection of clusters within a dataset. A measure of dissimilarity is computed for each cluster from a defined reference cluster, and the clusters are ordered according to the chosen dissimilarity. At least N clusters are selected as the most differentiated clusters relative to the defined reference. Within each cluster, the top-M most distinguishing cluster attributes can be automatically identified by an analogous process that computes the dissimilarity of each cluster attribute to its corresponding attribute in the reference cluster, and orders the attributes by dissimilarity. This then allows for automatic surfacing of what it is about a cluster that differentiates its members relative to the population as a whole, and to provide insight on what action or treatment might be made to address that specific segment of the underlying population.



EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, **Published:**

LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,

SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,

GW, KM, ML, MR, NE, SN, TD, TG).

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