



(51) International Patent Classification:

G06F 17/21 (2006.01) *G06F 3/048* (2006.01)
G06F 17/24 (2006.01) *G06F 3/14* (2006.01)

(21) International Application Number:

PCT/US2012/022454

(22) International Filing Date:

24 January 2012 (24.01.2012)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

13/014,520 26 January 2011 (26.01.2011) US

(71) Applicant (for all designated States except US): MICROSOFT CORPORATION [US/US]; One Microsoft Way, Redmond, Washington 98052-6399 (US).

(72) Inventors: ROTHSCILLER, Chad; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). BATTAGIN, Daniel; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). BENEDICT, Christopher; c/o Microsoft Corporation, LCA - International Patents, One Mi-

crosoft Way, Redmond, Washington 98052-6399 (US). MOREIRA-SILVEIRA, Rodrigo; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). COHEN, Eric; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). GULWANI, Sumit; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). ROUHANA, Dany; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). SINGH, Rishabh; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). ZORN, Benjamin Goth; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). VENKATESAN, Ramarathnam; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). DANILOV, Dmitri O.; c/o Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

[Continued on next page]

(54) Title: FORMATTING DATA BY EXAMPLE

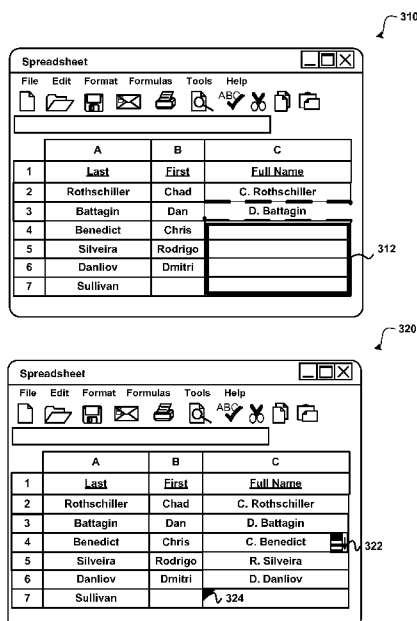


FIG.3

(57) Abstract: Data formatting rules to convert data from one form to another form are automatically determined based on a user's edits. A machine learning heuristic is applied to a user's edits to determine a data formatting rule that may be applied to data. For example, a user may make edits that add/remove characters from data, concatenate data, extract data, rename data, and the like. The machine learning heuristic may be automatically triggered in response to an event (e.g. after a predetermined number of edits are made to a same type of data) or manually triggered (e.g. selecting a user interface option). The data formatting rule may be applied to other data and the results of the formatting reviewable by the user. Based on further edits/reviews, the data formatting rule may be updated. The data formatting rules may be stored for later use.



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

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

DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

- *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*