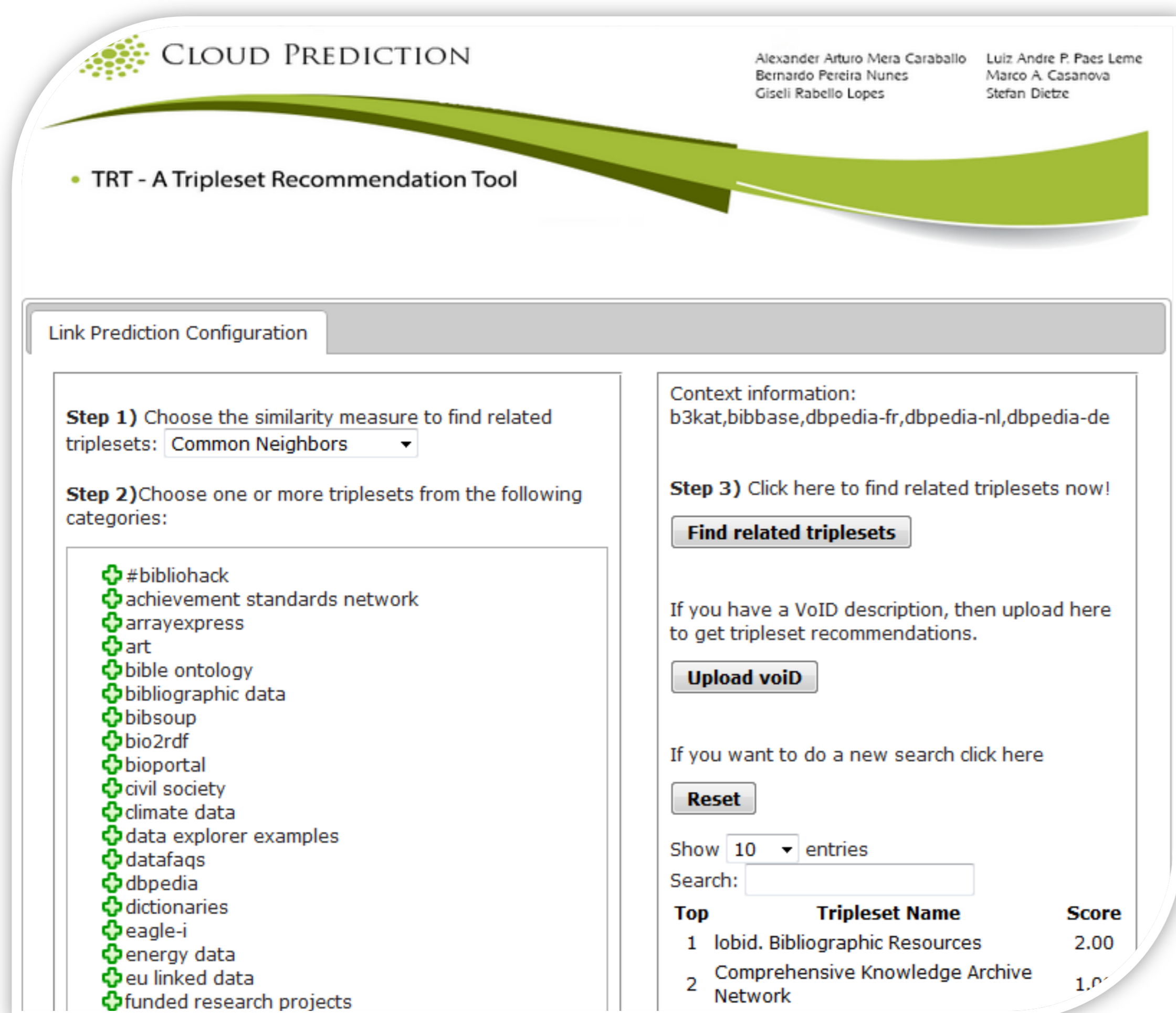# TRT - A Tripleset Recommendation Tool

Alexander Arturo Mera Caraballo[1], Bernardo Pereira Nunes[1], Giseli Rabello Lopes[1],
Luiz André P. Paes Leme[2], Marco A. Casanova[1], Stefan Dietze[3]

[1]Pontifical Catholic University of Rio de Janeiro (Brazil), [2]Fluminense Federal University (Brazil), [3]Leibniz University Hannover (Germany)

TRT in action – http://web.ccead.puc-rio.br:8080/Uncover

Tripleset recommendation problem

A considerable number of triplesets, following the Linked Data principles, have already been published in a large number of areas.

This growth makes it difficult to choose which triplesets should be interlinked with a given tripleset. Therefore, we developed a tool for helping this process on recommending related triplesets to a data source. The tripleset recommendation problem can be defined as:

Given a tripleset t and a set of triplesets S, rank the triplesets in S based on the probability of interlinking t with them.

TRT - Tripleset recommendation tool

The tool initially builds the Linked Data Network $G = (S,C)$ defined by the metadata stored in the DataHub repository. Then, If we have a user that is working on tripleset t and wants to discover one or more triplesets u, the user proceeds to define the input data the tool requires:

1 Select a prediction index

2 Define a target context Ct for t in one of the two ways:
-By providing a VoID descriptor vt
-By manually selecting triplesets from the categories the tool displays

3 The tool outputs a ranked list of triplesets

Evaluation

We evaluate the following prediction indices: Common Neighbours (CN), Salton, Jaccard, Sørensen, Hub Promoted index (HPI), Hub Depressed index (HDI), Leicht-Holme-Newman index (LHN1), Preferencial Attachment (PA), Adamic-Adar (AA) and Resource Allocation (RA).

To evaluate the prediction indices, we used three standard metrics: Area Under the receiver operating characteristic Curve (AUC), Mean Average Precision (MAP) and Recall. Table 1 summarizes the results for different target context sizes (shown in the first column of the table). The entries corresponding to the highest results among the 12 indices are emphasized in light blue.

Table 1: AUC, MAP and Recall of the local and quasi-local indices

AUC	CN	Salton	Jaccard	Sørensen	HPI	HDI	LHN	PA	AA	RA	LP	LRW
1	70.52	47.79	69.84	69.28	48.94	69.31	48.00	83.74	71.31	70.53	70.74	69.67
5	87.10	55.73	81.20	80.93	58.78	80.17	52.24	90.76	88.45	88.02	92.70	83.21
10	92.42	57.14	85.06	84.85	60.84	83.79	52.87	92.81	92.37	92.40	92.25	86.69
20	92.77	58.47	88.34	88.03	59.45	86.54	51.39	94.33	92.53	92.64	92.76	88.22
50	92.84	59.10	92.96	92.99	56.27	92.09	52.30	95.90	92.17	92.72	91.91	90.26
MAP	CN	Salton	Jaccard	Sørensen	HPI	HDI	LHN	PA	AA	RA	LP	LRW
1	18.17	14.49	16.30	14.73	17.08	15.00	14.8	37.83	18.06	17.8	18.46	15.57
5	49.48	25.07	21.80	20.36	35.14	19.02	18.38	48.26	52.20	51.48	58.23	26.05
10	63.49	30.99	30.40	28.71	41.81	24.41	19.44	52.62	63.43	63.71	62.63	31.91
20	71.20	34.22	44.37	43.56	38.14	34.14	17.9	53.97	71.46	72.38	70.59	34.66
50	71.13	27.73	69.49	70.55	20.64	66.14	15.92	47.3	70.99	72.42	67.51	39.03
Recall	CN	Salton	Jaccard	Sørensen	HPI	HDI	LHN	PA	AA	RA	LP	LRW
1	48.72	49.69	49.86	49.76	49.68	49.55	50.02	96.40	50.81	48.74	48.81	49.12
5	81.45	83.8	82.69	83.03	83.68	82.23	82.43	98.45	83.63	82.83	86.9	82.42
10	89.52	88.73	89.42	89.29	89.35	89.85	89.17	98.74	89.49	89.28	88.96	89.21
20	90.03	90.31	89.68	89.18	89.12	89.53	89.19	99.80	89.50	89.58	89.84	90.01
50	90.05	90.16	90.15	90.21	90.04	89.38	88.45	99.56	89.06	89.58	89.02	89.71

Interested on other projects in this room?

We recommend:

-Cite4Me
- Profiles of LD

Vote for us!

ISWC 2013 - Australia