

Automatic SDG Budget Tagging: Building Public Financial
Management Capacity through Natural Language Processing

Appendix

Daniele Guariso¹, Omar A. Guerrero¹, and Gonzalo Castañeda²

¹The Alan Turing Institute, London, UK

²Centro de Investigación y Docencia Económica (CIDE), Mexico City, Mexico

A List of budget programs features

Table A.1: Budget Programs Features by Country

Country	<i>N</i> of Features	Features
Mexico	21	Tranche Description, Description Responsible Unit, Functional Group Description, Function Description, Subfunction Description, Description Institutional Activity, Description Modality, Description Budget Program, Description National Development Plan, National Development Plan Objective, National Development Plan Program Description, National Development Plan Program Objective Description, Strategic Objective, Objective Description, External Conditions, Indicator Definition, Indicator Description, Matrix Description, Theme Description, Concept Description, Funding Source Description.
Colombia	10	Entity Name, Entity, Sector as of 2012, Grouped Sectors, Subsector, Other Classification, Medium Term Expenditure Framework, Accounts, Name, Main Programs.
Uruguay	28	Organism Name, Executing Unit Name, Grouped Executing Unit Name, Program Area Name, Program Name, Project Name, Financing Name, Financing Source Name, Expense Type Name, Group Name, Subgroup Name, Expense Auxiliary Object Name, Scope Program Area, Program Area Indicators, Program Objective, Program Scope, Program Indicators, Agency Mission, Agency Vision, Agency Scope, Agency Duties, Agency Objective, Agency Objective Indicators, Executing Unit Mission, Executing Unit Vision, Executing Unit Duties, Executing Unit Objective, Executing Unit Objective Indicators.

B Cosine similarity and soft cosine similarity across SDGs

Table B.1: Similarity Scores Across SDGs

	Mexico		Colombia		Uruguay	
	<i>CosSim</i>	<i>SoftCosSim</i>	<i>CosSim</i>	<i>SoftCosSim</i>	<i>CosSim</i>	<i>SoftCosSim</i>
<i>SDG1</i>	0.367	0.198	0.209	0.238	0.396	0.409
<i>SDG2</i>	0.338	0.259	0.275	0.319	0.487	0.359
<i>SDG3</i>	0.378	0.289	0.147	0.314	0.327	0.251
<i>SDG4</i>	0.310	0.308	0.188	0.421	0.339	0.354
<i>SDG5</i>	0.299	0.247	0.211	0.276	0.523	0.420
<i>SDG6</i>	0.265	0.186	0.157	0.297	0.421	0.301
<i>SDG7</i>	0.302	0.229	0.152	0.201	0.275	0.332
<i>SDG8</i>	0.467	0.375	0.305	0.330	0.490	0.433
<i>SDG9</i>	0.450	0.327	0.184	0.270	0.472	0.335
<i>SDG10</i>	0.370	0.209	0.288	0.376	0.422	0.442
<i>SDG11</i>	0.397	0.310	0.358	0.404	0.479	0.382
<i>SDG12</i>	-	-	0.358	0.364	0.428	0.319
<i>SDG13</i>	-	-	0.354	0.314	0.382	0.254
<i>SDG14</i>	-	-	0.289	0.334	0.314	0.177
<i>SDG15</i>	0.228	0.213	0.221	0.245	0.402	0.301
<i>SDG16</i>	0.447	0.333	0.281	0.251	0.283	0.256
<i>SDG17</i>	0.377	0.253	0.306	0.289	0.422	0.324

Note: For Mexico, we exclude SDG 12 and 13, as there is only one BP for each of them. SDG 14 instead, is not represented in the BPs.