

Lezione

Distributional Semantics

Luigi Di Caro

Distributional Semantics

- Panoramica
 - Semantica distribuzionale: rivisitazione in chiave linguistica delle tecniche di Text Mining
 - Similarità: concetti e problemi
 - Esempi ed esercizi

DS: Serie di citazioni

- *“Words that occur in the same contexts tend to have similar meanings”*. [Harris ‘54]
- *“A word is characterized by the company it keeps”*. [Firth ‘57]
- *“The conjunctive use of words primarily serves to more narrowly specify the object of discourse”*. [Furnas ‘83]
- *“There is some underlying latent semantic structure in the data that is partially obscured by the randomness of word choice with respect to retrieval”*. [Deerwester ‘90]
- *“Topic document will probabilistically influence the author’s choice of words when writing the document”*. [Blei ‘03]
- *“Pairs of words that co-occur in similar patterns tend to have similar semantic relations”*. [Turney ‘03]

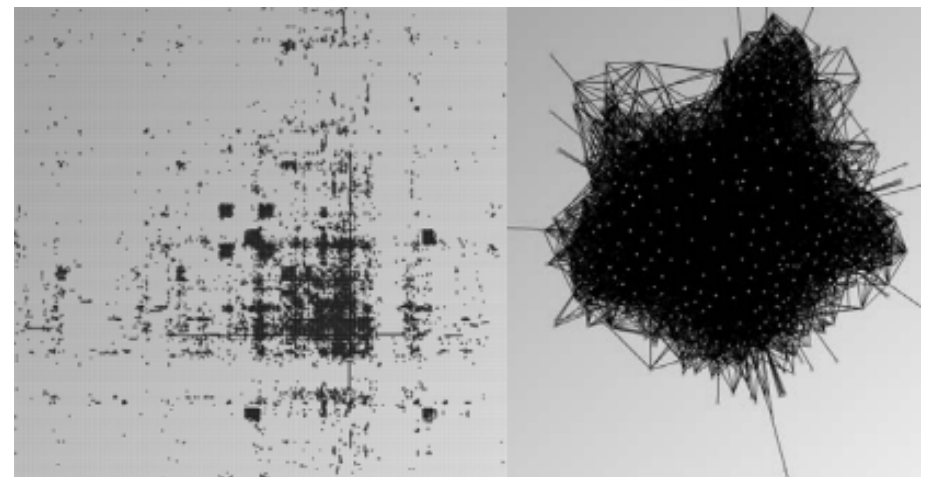
DS: aree diverse, nomi diversi

- Distributional Hypothesis / Distributional Analysis [Linguistics]
- Distributional Semantics [Natural Language Proc.]
- Vector Space Model [Information Retrieval]
- Conceptual spaces [Cognitive Science]
- Latent Semantics Analysis [Information Retrieval]
- Graded Categorization [Cognitive Science]
- Hyperspace Analogue to Language [Psychometrics]
- Adjacency matrix [Graph Theory]

Text Mining: breve ripasso

- Sistema semplice per rappresentare dati multidimensionali
 - Testi, ma anche...
 - Networks
 - E qualsiasi x-to-y data
 - Person - Question matrix
 - Object - Feature matrix
 - ...

$$\begin{bmatrix} & \text{Term}_1 & \text{Term}_2 & \dots & \dots & \dots & \text{Term}_N \\ \text{Doc}_1 & w_{11} & w_{12} & \dots & \dots & \dots & w_{1N} \\ \text{Doc}_2 & w_{21} & w_{22} & \dots & \dots & \dots & w_{2N} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \text{Doc}_K & w_{K1} & w_{K2} & \dots & \dots & \dots & w_{KN} \end{bmatrix}$$



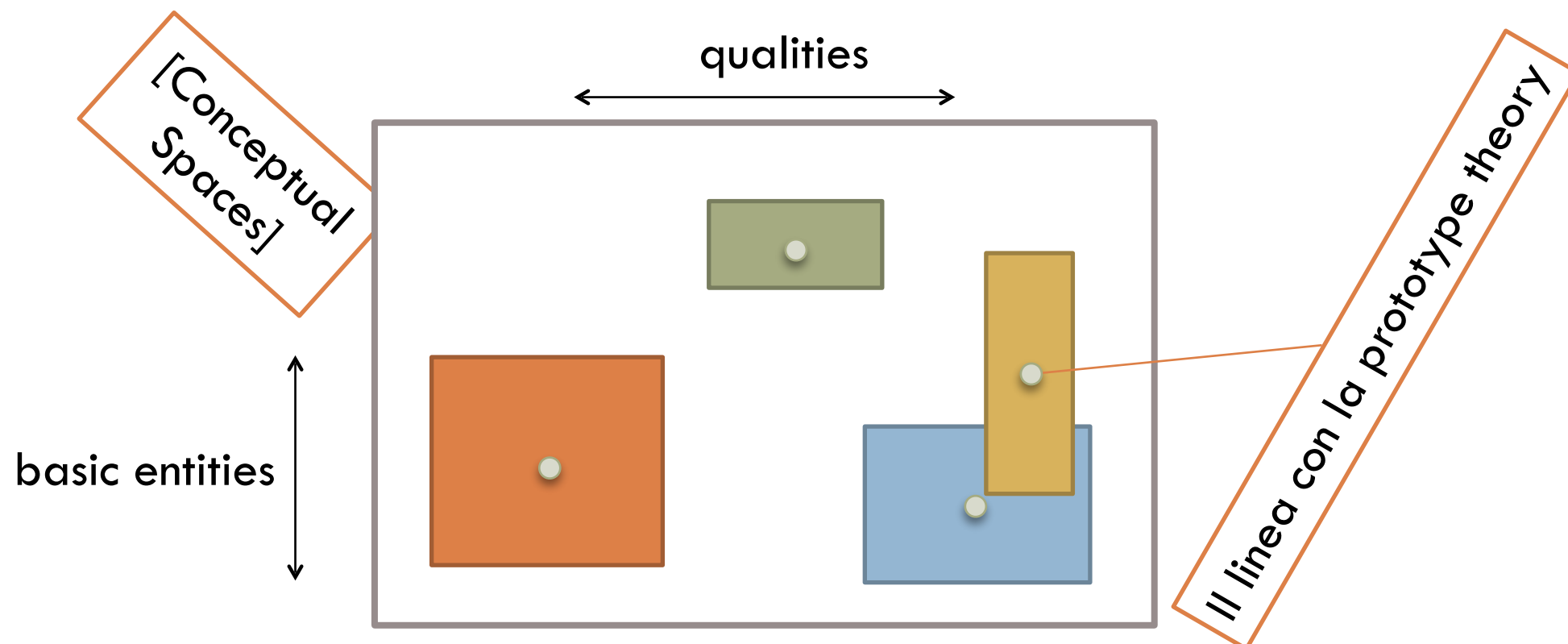
- Diventa molto semplice calcolare la similarità tra elementi
 - *Similarity*: concetto fondamentale per l'NLP

Perché usare Matrici?

- Approssimazione
 - La comprensione umana del linguaggio è di fatto un concetto basato sull'approssimazione
 - *Significato e linguaggio* sono due concetti sfumati
 - Metodi “esatti” non catturerebbero le sfumature dell'uso del linguaggio
 - DS “funziona” in molti casi: è anche questo un motivo che ne giustifica l'utilizzo!

Perché usare Matrici?

- Tra symbolic representations e associationism/connectionism
- Il significato diventa una *regione geometrica*
- Rappresentazione che facilita la *condivisione della conoscenza*



Parametri, tecniche, misure, operazioni...

- Righe / Colonne: similarità (cosine sim., jaccard, ...)
- Strategie di pesatura (freq, tf-idf, pmi, ...)
- Trasformazioni matriciali (SVD, NNMF, ...)
- Clustering (K-means, X-means, EM, ...)

2202	2229	2256	2283	2310	2337	2364	2391	2418	2445	2472	2499	1	28	55	82	109	136	163	190	217	244	271	298	325	352	379	404	431	458	485	512	539	566	593	620	647	674	701	728	755	782	809	836	863	890	917	944	971	998	1025	1052	1079	1106	1133	1160	1187	1214	1241	1268	1295	1322	1349	1376	1403	1430	1457	1484	1511	1538	1565	1592	1619	1646	1673	1700	1727	1754	1781	1808	1835	1862	1889	1916	1943	1970	1997	2024	2051	2078	2105	2132	2159	2186	2213	2240	2267	2294	2321	2348	2375	2402	2429	2456	2483	2510	2537	2564	2591	2618	2645	2672	2699	2726	2753	2780	2807	2834	2861	2888	2915	2942	2969	2996	3023	3050	3077	3104	3131	3158	3185	3212	3239	3266	3293	3320	3347	3374	3401	3428	3455	3482	3509	3536	3563	3590	3617	3644	3671	3698	3725	3752	3779	3806	3833	3860	3887	3914	3941	3968	3995	4022	4049	4076	4103	4130	4157	4184	4211	4238	4265	4292	4319	4346	4373	4400	4427	4454	4481	4508	4535	4562	4589	4616	4643	4670	4697	4724	4751	4778	4805	4832	4859	4886	4913	4940	4967	4994	5021	5048	5075	5102	5129	5156	5183	5210	5237	5264	5291	5318	5345	5372	5399	5426	5453	5480	5507	5534	5561	5588	5615	5642	5669	5696	5723	5750	5777	5804	5831	5858	5885	5912	5939	5966	5993	6020	6047	6074	6101	6128	6155	6182	6209	6236	6263	6290	6317	6344	6371	6398	6425	6452	6479	6506	6533	6560	6587	6614	6641	6668	6695	6722	6749	6776	6803	6830	6857	6884	6911	6938	6965	6992	7019	7046	7073	7100	7127	7154	7181	7208	7235	7262	7289	7316	7343	7370	7397	7424	7451	7478	7505	7532	7559	7586	7613	7640	7667	7694	7721	7748	7775	7802	7829	7856	7883	7910	7937	7964	7991	8018	8045	8072	8099	8126	8153	8180	8207	8234	8261	8288	8315	8342	8369	8396	8423	8450	8477	8504	8531	8558	8585	8612	8639	8666	8693	8720	8747	8774	8801	8828	8855	8882	8909	8936	8963	8990	9017	9044	9071	9098	9125	9152	9179	9206	9233	9260	9287	9314	9341	9368	9395	9422	9449	9476	9503	9530	9557	9584	9611	9638	9665	9692	9719	9746	9773	9800	9827	9854	9881	9908	9935	9962	9989	10016	10043	10070	10097	10124	10151	10178	10205	10232	10259	10286	10313	10340	10367	10394	10421	10448	10475	10502	10529	10556	10583	10610	10637	10664	10691	10718	10745	10772	10799	10826	10853	10880	10907	10934	10961	10988	11015	11042	11069	11096	11123	11150	11177	11204	11231	11258	11285	11312	11339	11366	11393	11420	11447	11474	11501	11528	11555	11582	11609	11636	11663	11690	11717	11744	11771	11798	11825	11852	11879	11906	11933	11960	11987	12014	12041	12068	12095	12122	12149	12176	12203	12230	12257	12284	12311	12338	12365	12392	12419	12446	12473	12500	12527	12554	12581	12608	12635	12662	12689	12716	12743	12770	12797	12824	12851	12878	12905	12932	12959	12986	13013	13040	13067	13094	13121	13148	13175	13202	13229	13256	13283	13310	13337	13364	13391	13418	13445	13472	13499	13526	13553	13580	13607	13634	13661	13688	13715	13742	13769	13796	13823	13850	13877	13904	13931	13958	13985	14012	14039	14066	14093	14120	14147	14174	14201	14228	14255	14282	14309	14336	14363	14390	14417	14444	14471	14498	14525	14552	14579	14606	14633	14660	14687	14714	14741	14768	14795	14822	14849	14876	14903	14930	14957	14984	15011	15038	15065	15092	15119	15146	15173	15200	15227	15254	15281	15308	15335	15362	15389	15416	15443	15470	15497	15524	15551	15578	15605	15632	15659	15686	15713	15740	15767	15794	15821	15848	15875	15902	15929	15956	15983	16010	16037	16064	16091	16118	16145	16172	16199	16226	16253	16280	16307	16334	16361	16388	16415	16442	16469	16496	16523	16550	16577	16604	16631	16658	16685	16712	16739	16766	16793	16820	16847	16874	16901	16928	16955	16982	17009	17036	17063	17090	17117	17144	17171	17198	17225	17252	17279	17306	17333	17360	17387	17414	17441	17468	17495	17522	17549	17576	17603	17630	17657	17684	17711	17738	17765	17792	17819	17846	17873	17900	17927	17954	17981	18008	18035	18062	18089	18116	18143	18170	18197	18224	18251	18278	18305	18332	18359	18386	18413	18440	18467	18494	18521	18548	18575	18602	18629	18656	18683	18710	18737	18764	18791	18818	18845	18872	18899	18926	18953	18980	19007	19034	19061	19088	19115	19142	19169	19196	19223	19250	19277	19304	19331	19358	19385	19412	19439	19466	19493	19520	19547	19574	19601	19628	19655	19682	19709	19736	19763	19790	19817	19844	19871	19898	19925	19952	19979	20006	20033	20060	20087	20114	20141	20168	20195	20222	20249	20276	20303	20330	20357	20384	20411	20438	20465	20492	20519	20546	20573	20600	20627	20654	20681	20708	20735	20762	20789	20816	20843	20870	20897	20924	20951	20978	21005	21032	21059	21086	21113	21140	21167	21194	21221	21248	21275	21302	21329	21356	21383	21410	21437	21464	21491	21518	21545	21572	21599	21626	21653	21680	21707	21734	21761	21788	21815	21842	21869	21896	21923	21950	21977	22004	22031	22058	22085	22112	22139	22166	22193	22220	22247	22274	22301	22328	22355	22382	22409	22436	22463	22490	22517	22544	22571	22598	22625	22652	22679	22706	22733	22760	22787	22814	22841	22868	22895	22922	22949	22976	23003	23030	23057	23084	23111	23138	23165	23192	23219	23246	23273	23300	23327	23354	23381	23408	23435	23462	23489	23516	23543	23570	23597	23624	23651	23678	23705	23732	23759	23786	23813	23840	23867	23894	23921	23948	23975	24002	24029	24056	24083	24110	24137	24164	24191	24218	24245	24272	24299	24326	24353	24380	24407	24434	24461	24488	24515	24542	24569	24596	24623	24650	24677	24704	24731	24758	24785	24812	24839	24866	24893	24920	24947	24974	25001	25028	25055	25082	25109	25136	25163	25190	25217	25244	25271	25298	25325	25352	25379	25406	25433	25460	25487	25514	25541	25568	25595	25622	25649	25676	25703	25730	25757	25784	25811	25838	25865	25892	25919	25946	25973	26000	26027	26054	26081	26108	26135	26162	26189	26216	26243	26270	26297	26324	26351	26378	26405	26432	26459	26486	26513	26540	26567	26594	26621	26648	26675	26702	26729	26756	26783	26810	26837	26864	26891	26918	26945	26972	27000	27027	27054	27081	27108	27135	27162	27189	27216	27243	27270	27297	27324	27351	27378	27405	27432	27459	27486	27513	27540	27567	27594	27621	27648	27675	27702	27729	27756	27783	27810	27837	27864	27891	27918	27945	27972	28000	28027	28054	28081	28108	28135	28162	28189	28216	28243	28270	28297	28324	28351	28378	28405	28432	28459	28486	28513	28540	28567	28594	28621	28648	28675	28702	28729	28756	28783	28810	28837	28864	28891	28918	28945	28972	29000	29027	29054	29081	29108	29135	29162	29189	29216	29243	29270	29297	29324	29351	29378	29405	29432	29459	29486	29513	29540	29567	29594	29621	29648	29675	29702	29729	29756	29783	29810	29837	29864	29891	29918	29945	29972	30000	30027	30054	30081	30108	30135	30162	30189	30216	30243	30270	30297	30324	30351	30378	30405	30432	30459	30486	30513	30540	30567	30594	30621	30648	30675	30702	30729	30756	30783	30810	30837	30864	30891	30918	30945	30972	31000	31027	31054	31081	31108	31135	31162	31189	31216	31243	31270	31297	31324	31351	31378	31405	31432	31459	31486	31513	31540	31567	31594	31621	31648	31675	31702	31729	31756	31783	31810	31837	31864	31891	31918	31945	31972	32000	32027	32054	32081	32108	32135	32162	32189	32216	32243	32270	32297	32324	32351	32378	32405	32432	32459	32486	32513	32540	32567	32594	32621	32648	32675	32702	32729	32756	32783	32810	32837	32864	32891	32918	32945	32972	33000	33027	33054	33081	33108	33135	33162	33189	33216	33243	33270	33297	33324	33351	33378	33405	33432	33459	33486	33513	33540	33567	33594	33621	33648	33675	33702	33729
------	------	------	------	------	------	------	------	------	------	------	------	---	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

DS oltre il Text Mining: pre-processing e annotazioni linguistiche

(morphological, lexical and syntactic)

Normalization

Tokens to Types

Basic English

Stemming

Lemmatization

...



Named Entities

Semantic roles

...

(semantic)

Denormalization

Configurazioni matriciali

- Term-Document Matrix
- Term-Context Matrix
- Pair-pattern Matrix

*Peter Turney. From Frequency to Meaning:
Vector Space Models of Semantics*
<https://www.jair.org/media/2934/live-2934-4846-jair.pdf>

?

?

2202	2229	2256	2283	2310	2337	2364	2391	2418	2445	2472	2499	1	28	55	82	109	136	163	190	217	244	271	298	325	352	379	406	433	460	487	514	541	568	595	622	649	676	703	730	757	784	811	838	865	892	919	946	973	1000	1027	1054	1081	1108	1135	1162	1189	1216	1243	1270	1297	1324	1351	1378	1405	1432	1459	1486	1513	1540	1567	1594	1621	1648	1675	1702	1729	1756	1783	1810	1837	1864	1891	1918	1945	1972	1999	2026	2053	2080	2107	2134	2161	2188	2215	2242	2269	2296	2323	2350	2377	2404	2431	2458	2485	2512	2539	2566	2593	2620	2647	2674	2701	2728	2755	2782	2809	2836	2863	2890	2917	2944	2971	2998	3025	3052	3079	3106	3133	3160	3187	3214	3241	3268	3295	3322	3349	3376	3403	3430	3457	3484	3511	3538	3565	3592	3619	3646	3673	3700	3727	3754	3781	3808	3835	3862	3889	3916	3943	3970	3997	4024	4051	4078	4105	4132	4159	4186	4213	4240	4267	4294	4321	4348	4375	4402	4429	4456	4483	4510	4537	4564	4591	4618	4645	4672	4699	4726	4753	4780	4807	4834	4861	4888	4915	4942	4969	4996	5023	5050	5077	5104	5131	5158	5185	5212	5239	5266	5293	5320	5347	5374	5401	5428	5455	5482	5509	5536	5563	5590	5617	5644	5671	5698	5725	5752	5779	5806	5833	5860	5887	5914	5941	5968	5995	6022	6049	6076	6103	6130	6157	6184	6211	6238	6265	6292	6319	6346	6373	6400	6427	6454	6481	6508	6535	6562	6589	6616	6643	6670	6697	6724	6751	6778	6805	6832	6859	6886	6913	6940	6967	6994	7021	7048	7075	7102	7129	7156	7183	7210	7237	7264	7291	7318	7345	7372	7399	7426	7453	7480	7507	7534	7561	7588	7615	7642	7669	7696	7723	7750	7777	7804	7831	7858	7885	7912	7939	7966	7993	8020	8047	8074	8101	8128	8155	8182	8209	8236	8263	8290	8317	8344	8371	8398	8425	8452	8479	8506	8533	8560	8587	8614	8641	8668	8695	8722	8749	8776	8803	8830	8857	8884	8911	8938	8965	8992	9019	9046	9073	9100	9127	9154	9181	9208	9235	9262	9289	9316	9343	9370	9397	9424	9451	9478	9505	9532	9559	9586	9613	9640	9667	9694	9721	9748	9775	9802	9829	9856	9883	9910	9937	9964	9991	10018	10045	10072	10099	10126	10153	10180	10207	10234	10261	10288	10315	10342	10369	10396	10423	10450	10477	10504	10531	10558	10585	10612	10639	10666	10693	10720	10747	10774	10801	10828	10855	10882	10909	10936	10963	10990	11017	11044	11071	11098	11125	11152	11179	11206	11233	11260	11287	11314	11341	11368	11395	11422	11449	11476	11503	11530	11557	11584	11611	11638	11665	11692	11719	11746	11773	11800	11827	11854	11881	11908	11935	11962	11989	12016	12043	12070	12097	12124	12151	12178	12205	12232	12259	12286	12313	12340	12367	12394	12421	12448	12475	12502	12529	12556	12583	12610	12637	12664	12691	12718	12745	12772	12799	12826	12853	12880	12907	12934	12961	12988	13015	13042	13069	13096	13123	13150	13177	13204	13231	13258	13285	13312	13339	13366	13393	13420	13447	13474	13501	13528	13555	13582	13609	13636	13663	13690	13717	13744	13771	13798	13825	13852	13879	13906	13933	13960	13987	14014	14041	14068	14095	14122	14149	14176	14203	14230	14257	14284	14311	14338	14365	14392	14419	14446	14473	14500	14527	14554	14581	14608	14635	14662	14689	14716	14743	14770	14797	14824	14851	14878	14905	14932	14959	14986	15013	15040	15067	15094	15121	15148	15175	15202	15229	15256	15283	15310	15337	15364	15391	15418	15445	15472	15499	15526	15553	15580	15607	15634	15661	15688	15715	15742	15769	15796	15823	15850	15877	15904	15931	15958	15985	16012	16039	16066	16093	16120	16147	16174	16201	16228	16255	16282	16309	16336	16363	16390	16417	16444	16471	16498	16525	16552	16579	16606	16633	16660	16687	16714	16741	16768	16795	16822	16849	16876	16903	16930	16957	16984	17011	17038	17065	17092	17119	17146	17173	17200	17227	17254	17281	17308	17335	17362	17389	17416	17443	17470	17497	17524	17551	17578	17605	17632	17659	17686	17713	17740	17767	17794	17821	17848	17875	17902	17929	17956	17983	18010	18037	18064	18091	18118	18145	18172	18199	18226	18253	18280	18307	18334	18361	18388	18415	18442	18469	18496	18523	18550	18577	18604	18631	18658	18685	18712	18739	18766	18793	18820	18847	18874	18901	18928	18955	18982	19009	19036	19063	19090	19117	19144	19171	19198	19225	19252	19279	19306	19333	19360	19387	19414	19441	19468	19495	19522	19549	19576	19603	19630	19657	19684	19711	19738	19765	19792	19819	19846	19873	19900	19927	19954	19981	20008	20035	20062	20089	20116	20143	20170	20197	20224	20251	20278	20305	20332	20359	20386	20413	20440	20467	20494	20521	20548	20575	20602	20629	20656	20683	20710	20737	20764	20791	20818	20845	20872	20899	20926	20953	20980	21007	21034	21061	21088	21115	21142	21169	21196	21223	21250	21277	21304	21331	21358	21385	21412	21439	21466	21493	21520	21547	21574	21601	21628	21655	21682	21709	21736	21763	21790	21817	21844	21871	21898	21925	21952	21979	22006	22033	22060	22087	22114	22141	22168	22195	22222	22249	22276	22303	22330	22357	22384	22411	22438	22465	22492	22519	22546	22573	22600	22627	22654	22681	22708	22735	22762	22789	22816	22843	22870	22897	22924	22951	22978	23005	23032	23059	23086	23113	23140	23167	23194	23221	23248	23275	23302	23329	23356	23383	23410	23437	23464	23491	23518	23545	23572	23599	23626	23653	23680	23707	23734	23761	23788	23815	23842	23869	23896	23923	23950	23977	24004	24031	24058	24085	24112	24139	24166	24193	24220	24247	24274	24301	24328	24355	24382	24409	24436	24463	24490	24517	24544	24571	24598	24625	24652	24679	24706	24733	24760	24787	24814	24841	24868	24895	24922	24949	24976	25003	25030	25057	25084	25111	25138	25165	25192	25219	25246	25273	25300	25327	25354	25381	25408	25435	25462	25489	25516	25543	25570	25597	25624	25651	25678	25705	25732	25759	25786	25813	25840	25867	25894	25921	25948	25975	26002	26029	26056	26083	26110	26137	26164	26191	26218	26245	26272	26299	26326	26353	26380	26407	26434	26461	26488	26515	26542	26569	26596	26623	26650	26677	26704	26731	26758	26785	26812	26839	26866	26893	26920	26947	26974	27001	27028	27055	27082	27109	27136	27163	27190	27217	27244	27271	27298	27325	27352	27379	27406	27433	27460	27487	27514	27541	27568	27595	27622	27649	27676	27703	27730	27757	27784	27811	27838	27865	27892	27919	27946	27973	28000	28027	28054	28081	28108	28135	28162	28189	28216	28243	28270	28297	28324	28351	28378	28405	28432	28459	28486	28513	28540	28567	28594	28621	28648	28675	28702	28729	28756	28783	28810	28837	28864	28891	28918	28945	28972	29000	29027	29054	29081	29108	29135	29162	29189	29216	29243	29270	29297	29324	29351	29378	29405	29432	29459	29486	29513	29540	29567	29594	29621	29648	29675	29702	29729	29756	29783	29810	29837	29864	29891	29918	29945	29972	30000	30027	30054	30081	30108	30135	30162	30189	30216	30243	30270	30297	30324	30351	30378	30405	30432	30459	30486	30513	30540	30567	30594	30621	30648	30675	30702	30729	30756	30783	30810	30837	30864	30891	30918	30945	30972	31000	31027	31054	31081	31108	31135	31162	31189	31216	31243	31270	31297	31324	31351	31378	31405	31432	31459	31486	31513	31540	31567	31594	31621	31648	31675	31702	31729	31756	31783	31810	31837	31864	31891	31918	31945	31972	32000	32027	32054	32081	32108	32135	32162	32189	32216	32243	32270	3229
------	------	------	------	------	------	------	------	------	------	------	------	---	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	------

Configurazioni matriciali

- Term-Document Matrix

- Document *similarity*

- Document *clustering*

- Document *classification*

- Document *segmentation*

- (Partial) *question answering*

- Question analysis, doc retrieval, passage retrieval, answer extraction

terms

docs

2229	2250	2263	2309	2307	2364	2391	2438	2445	2472	2489	1	29	35	62	109	136	143	190	217	241	271	286	325	357	389	443	369	393	412	429	438	450	459	465	472	478	480	487	491	493	495	497	499	501	503	505	507	509	511	513	515	517	519	521	523	525	527	529	531	533	535	537	539	541	543	545	547	549	551	553	555	557	559	561	563	565	567	569	571	573	575	577	579	581	583	585	587	589	591	593	595	597	599	601	603	605	607	609	611	613	615	617	619	621	623	625	627	629	631	633	635	637	639	641	643	645	647	649	651	653	655	657	659	661	663	665	667	669	671	673	675	677	679	681	683	685	687	689	691	693	695	697	699	701	703	705	707	709	711	713	715	717	719	721	723	725	727	729	731	733	735	737	739	741	743	745	747	749	751	753	755	757	759	761	763	765	767	769	771	773	775	777	779	781	783	785	787	789	791	793	795	797	799	801	803	805	807	809	811	813	815	817	819	821	823	825	827	829	831	833	835	837	839	841	843	845	847	849	851	853	855	857	859	861	863	865	867	869	871	873	875	877	879	881	883	885	887	889	891	893	895	897	899	901	903	905	907	909	911	913	915	917	919	921	923	925	927	929	931	933	935	937	939	941	943	945	947	949	951	953	955	957	959	961	963	965	967	969	971	973	975	977	979	981	983	985	987	989	991	993	995	997	999	1001	1003	1005	1007	1009	1011	1013	1015	1017	1019	1021	1023	1025	1027	1029	1031	1033	1035	1037	1039	1041	1043	1045	1047	1049	1051	1053	1055	1057	1059	1061	1063	1065	1067	1069	1071	1073	1075	1077	1079	1081	1083	1085	1087	1089	1091	1093	1095	1097	1099	1101	1103	1105	1107	1109	1111	1113	1115	1117	1119	1121	1123	1125	1127	1129	1131	1133	1135	1137	1139	1141	1143	1145	1147	1149	1151	1153	1155	1157	1159	1161	1163	1165	1167	1169	1171	1173	1175	1177	1179	1181	1183	1185	1187	1189	1191	1193	1195	1197	1199	1201	1203	1205	1207	1209	1211	1213	1215	1217	1219	1221	1223	1225	1227	1229	1231	1233	1235	1237	1239	1241	1243	1245	1247	1249	1251	1253	1255	1257	1259	1261	1263	1265	1267	1269	1271	1273	1275	1277	1279	1281	1283	1285	1287	1289	1291	1293	1295	1297	1299	1301	1303	1305	1307	1309	1311	1313	1315	1317	1319	1321	1323	1325	1327	1329	1331	1333	1335	1337	1339	1341	1343	1345	1347	1349	1351	1353	1355	1357	1359	1361	1363	1365	1367	1369	1371	1373	1375	1377	1379	1381	1383	1385	1387	1389	1391	1393	1395	1397	1399	1401	1403	1405	1407	1409	1411	1413	1415	1417	1419	1421	1423	1425	1427	1429	1431	1433	1435	1437	1439	1441	1443	1445	1447	1449	1451	1453	1455	1457	1459	1461	1463	1465	1467	1469	1471	1473	1475	1477	1479	1481	1483	1485	1487	1489	1491	1493	1495	1497	1499	1501	1503	1505	1507	1509	1511	1513	1515	1517	1519	1521	1523	1525	1527	1529	1531	1533	1535	1537	1539	1541	1543	1545	1547	1549	1551	1553	1555	1557	1559	1561	1563	1565	1567	1569	1571	1573	1575	1577	1579	1581	1583	1585	1587	1589	1591	1593	1595	1597	1599	1601	1603	1605	1607	1609	1611	1613	1615	1617	1619	1621	1623	1625	1627	1629	1631	1633	1635	1637	1639	1641	1643	1645	1647	1649	1651	1653	1655	1657	1659	1661	1663	1665	1667	1669	1671	1673	1675	1677	1679	1681	1683	1685	1687	1689	1691	1693	1695	1697	1699	1701	1703	1705	1707	1709	1711	1713	1715	1717	1719	1721	1723	1725	1727	1729	1731	1733	1735	1737	1739	1741	1743	1745	1747	1749	1751	1753	1755	1757	1759	1761	1763	1765	1767	1769	1771	1773	1775	1777	1779	1781	1783	1785	1787	1789	1791	1793	1795	1797	1799	1801	1803	1805	1807	1809	1811	1813	1815	1817	1819	1821	1823	1825	1827	1829	1831	1833	1835	1837	1839	1841	1843	1845	1847	1849	1851	1853	1855	1857	1859	1861	1863	1865	1867	1869	1871	1873	1875	1877	1879	1881	1883	1885	1887	1889	1891	1893	1895	1897	1899	1901	1903	1905	1907	1909	1911	1913	1915	1917	1919	1921	1923	1925	1927	1929	1931	1933	1935	1937	1939	1941	1943	1945	1947	1949	1951	1953	1955	1957	1959	1961	1963	1965	1967	1969	1971	1973	1975	1977	1979	1981	1983	1985	1987	1989	1991	1993	1995	1997	1999	2001	2003	2005	2007	2009	2011	2013	2015	2017	2019	2021	2023	2025	2027	2029	2031	2033	2035	2037	2039	2041	2043	2045	2047	2049	2051	2053	2055	2057	2059	2061	2063	2065	2067	2069	2071	2073	2075	2077	2079	2081	2083	2085	2087	2089	2091	2093	2095	2097	2099	2101	2103	2105	2107	2109	2111	2113	2115	2117	2119	2121	2123	2125	2127	2129	2131	2133	2135	2137	2139	2141	2143	2145	2147	2149	2151	2153	2155	2157	2159	2161	2163	2165	2167	2169	2171	2173	2175	2177	2179	2181	2183	2185	2187	2189	2191	2193	2195	2197	2199	2201	2203	2205	2207	2209	2211	2213	2215	2217	2219	2221	2223	2225	2227	2229	2231	2233	2235	2237	2239	2241	2243	2245	2247	2249	2251	2253	2255	2257	2259	2261	2263	2265	2267	2269	2271	2273	2275	2277	2279	2281	2283	2285	2287	2289	2291	2293	2295	2297	2299	2301	2303	2305	2307	2309	2311	2313	2315	2317	2319	2321	2323	2325	2327	2329	2331	2333	2335	2337	2339	2341	2343	2345	2347	2349	2351	2353	2355	2357	2359	2361	2363	2365	2367	2369	2371	2373	2375	2377	2379	2381	2383	2385	2387	2389	2391	2393	2395	2397	2399	2401	2403	2405	2407	2409	2411	2413	2415	2417	2419	2421	2423	2425	2427	2429	2431	2433	2435	2437	2439	2441	2443	2445	2447	2449	2451	2453	2455	2457	2459	2461	2463	2465	2467	2469	2471	2473	2475	2477	2479	2481	2483	2485	2487	2489	2491	2493	2495	2497	2499	2501	2503	2505	2507	2509	2511	2513	2515	2517	2519	2521	2523	2525	2527	2529	2531	2533	2535	2537	2539	2541	2543	2545	2547	2549	2551	2553	2555	2557	2559	2561	2563	2565	2567	2569	2571	2573	2575	2577	2579	2581	2583	2585	2587	2589	2591	2593	2595	2597	2599	2601	2603	2605	2607	2609	2611	2613	2615	2617	2619	2621	2623	2625	2627	2629	2631	2633	2635	2637	2639	2641	2643	2645	2647	2649	2651	2653	2655	2657	2659	2661	2663	2665	2667	2669	2671	2673	2675	2677	2679	2681	2683	2685	2687	2689	2691	2693	2695	2697	2699	2701	2703	2705	2707	2709	2711	2713	2715	2717	2719	2721	2723	2725	2727	2729	2731	2733	2735	2737	2739	2741	2743	2745	2747	2749	2751	2753	2755	2757	2759	2761	2763	2765	2767	2769	2771	2773	2775	2777	2779	2781	2783	2785	2787	2789	2791	2793	2795	2797	2799	2801	2803	2805	2807	2809	2811	2813	2815	2817	2819	2821	2823	2825	2827	2829	2831	2833	2835	2837	2839	2841	2843	2845	2847	2849	2851	2853	2855	2857	2859	2861	2863	2865	2867	2869	2871	2873	2875	2877	2879	2881	2883	2885	2887	2889	2891	2893	2895	2897	2899	2901	2903	2905	2907	2909	2911	2913	2915	2917	2919	2921	2923	2925	2927	2929	2931	2933	2935	2937	2939	2941	2943	2945	2947	2949	2951	2953	2955	2957	2959	2961	2963	2965	2967	2969	2971	2973	2975	2977	2979	2981	2983	2985	2987	2989	2991	2993	2995	2997	2999	3001	3003	3005	3007	3009	3011	3013	3015	3017	3019	3021	3023	3025	3027	3029	3031	3033	3035	3037	3039	3041	3043	3045	3047	3049	3051	3053	3055	3057	3059	3061	3063	3065	3067	3069	3071	3073	3075	3077	3079	3081	3083	3085	3087	3089	3091	3093	3095	3097	3099	3101	3103	3105	3107	3109	3111	3113	3115	3117	3119	312
------	------	------	------	------	------	------	------	------	------	------	---	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----

Configurazioni matriciali

- Term-Context Matrix
 - Word *similarity*
 - Word *clustering*
 - Word *classification*
 - Automatic *thesaurus generation*
 - Word sense *disambiguation*
 - Spelling *correction*
 - Semantic *role labeling*
 - Query *expansion*
 - Information *extraction*
 - Named *Entity Recognition*

Words / terms

contexts

222	225	228	233	2309	2384	2391	2438	2445	2472	2489	1	29	55	82	109	136	140	150	157	174	271	286	325	357	387	394	440	469	495	712	720	730	736	749	751	753	1278	901	707	730	741	748	815	842	849	896	923	951	
1229	2255	2292	2309	2320	2383	2390	2437	2444	2473	2489	17	54	81	108	135	142	149	158	163	179	287	324	328	360	380	407	407	464	471	478	1270	1276	1282	1303	1308	1320	740	750	757	814	841	868	895	922	949	951			
2254	2291	2306	2325	2362	2390	2416	2443	2470	2487	1499	1901	51	80	107	134	161	168	215	242	260	296	323	350	352	362	369	374	384	403	429	452	1269	1268	1273	1274	1276	1303	1320	742	750	756	813	840	867	894	921	948	977	977
2290	2307	2324	2381	2398	2415	2442	2469	2496	2498	1925	1927	79	106	133	160	167	214	241	268	295	322	349	351	375	395	402	429	456	483	510	1270	1268	1273	1274	1276	1303	1320	743	750	756	813	840	868	895	922	949	976	1003	
2306	2333	2380	2387	2414	2441	2468	2495	1987	1924	1926	1953	105	132	159	186	213	240	267	294	321	348	375	377	404	461	488	515	542	569	1270	1268	1273	1274	1276	1303	1320	744	751	813	840	868	895	922	949	976	1004			
2358	2385	2412	2439	2466	2493	1985	1922	1949	1951	1978	2005	137	164	211	238	265	292	319	346	373	400	402	429	436	470	527	554	581	608	1271	1269	1273	1274	1276	1303	1320	745	752	814	841	869	896	923	950	977	1005			
2384	2411	2438	2465	2492	1984	1921	1948	1953	1977	2004	2031	163	210	237	264	291	318	345	372	399	401	428	435	462	519	546	573	600	1272	1270	1274	1275	1276	1303	1320	746	753	814	841	869	896	923	950	977	1006				
2410	2437	2464	2491	1983	1950	1947	1974	1975	2003	2008	2057	209	236	263	290	317	344	371	398	425	452	479	481	508	565	592	619	646	1273	1271	1275	1276	1277	1304	1321	747	754	815	842	870	897	924	951	978	1007				
2436	2463	2490	1982	1959	1946	1973	2000	2002	2029	2056	2083	235	262	289	316	343	370	397	424	451	478	480	507	534	591	618	645	672	1274	1272	1276	1277	1278	1305	1322	748	755	816	843	871	898	925	952	979	1008				
2462	2489	1981	1958	1945	1972	1999	2001	2028	2055	2082	2109	261	288	315	342	369	396	423	450	477	504	531	533	560	623	650	677	704	1275	1273	1277	1278	1279	1306	1323	749	756	817	844	872	900	927	954	981	1009				
1988	1990	1997	1944	1971	1998	2005	2027	2054	2081	2108	2135	267	294	321	348	375	402	429	456	483	510	537	539	566	625	652	679	706	1276	1274	1278	1279	1280	1307	1324	750	757	818	845	873	901	928	955	982	1009				
1916	1943	1970	1997	2024	2051	2078	2105	2132	2159	2186	2213	269	296	323	350	377	404	431	458	485	512	539	541	568	625	652	679	706	1277	1275	1279	1280	1281	1308	1325	751	758	819	846	874	902	929	956	983	1010				
1915	1942	1969	1996	2023	2050	2077	2104	2131	2158	2185	2212	269	296	323	350	377	404	431	458	485	512	539	541	568	625	652	679	706	1278	1276	1280	1281	1282	1309	1326	752	759	820	847	875	903	930	957	984	1011				
1941	1968	1995	2022	2049	2076	2103	2130	2157	2184	2211	2238	269	296	323	350	377	404	431	458	485	512	539	541	568	625	652	679	706	1279	1277	1281	1282	1283	1310	1327	753	760	821	848	876	904	931	958	985	1012				
1967	1994	2021	2048	2075	2102	2129	2156	2183	2210	2237	2264	269	296	323	350	377	404	431	458	485	512	539	541	568	625	652	679	706	1280	1278	1282	1283	1284	1311	1328	754	761	822	849	877	905	932	959	986	1013				
1993	2020	2047	2074	2101	2128	2155	2182	2209	2236	2263	2290	269	296	323	350	377	404	431	458	485	512	539	541	568	625	652	679	706	1281	1279	1283	1284	1285	1312	1329	755	762	823	850	878	906	933	960	987	1014				
2019	2046	2073	2100	2127	2154	2181	2208	2235	2262	2289	2316	269	296	323	350	377	404	431	458	485	512	539	541	568	625	652	679	706	1282	1280	1284	1285	1286	1313	1330	756	763	824	851	879	907	934	961	988	1015				
2045	2072	2099	2126	2153	2180	2207	2234	2261	2288	2315	2342	269	296	323	350	377	404	431	458	485	512	539	541	568	625	652	679	706	1283	1281	1285	1286	1287	1314	1331	757	764	825	852	880	908	935	962	989	1016				
2071	2098	2125	2152	2179	2206	2233	2260	2287	2314	2341	2368	269	296	323	350	377	404	431	458	485	512	539	541	568	625	652	679	706	1284	1282	1286	1287	1288	1315	1332	758	765	826	853	881	909	936	963	990	1017				
2097	2124	2151	2178	2205	2232	2259	2286	2313	2340	2367	2394	269	296	323	350	377	404	431	458	485	512	539	541	568	625	652	679	706	1285	1283	1287	1288	1289	1316	1333	759	766	827	854	882	910	937	964	991	1018				
2123	2150	2177	2204	2231	2258	2285	2312	2339	2366	2393	2420	269	296	323	350	377	404	431	458	485	512	539	541	568	625	652	679	706	1286	1284	1288	1289	1290	1317	1334	760	767	828	855	883	911	938	965	992	1019				
2149	2176	2203	2230	2257	2284	2311	2338	2365	2392	2419	2446	269	296	323	350	377	404	431	458	485	512	539	541	568	625	652	679	706	1287	1285	1289	1290	1291	1318	1335	761	768	829	856	884	912	939	966	993	1020				
2175	2202	2229	2256	2283	2310	2337	2364	2391	2418	2445	2472	269	296	323	350	377	404	431	458	485	512	539	541	568	625	652	679	706	1288	1286	1290	1291	1292	1319	1336	762	769	830	857	885	913	940	967	994	1021				
2170	2200	2230	2260	2290	2320	2350	2380	2410	2440	2470	2500	269	296	323	350	377	404	431	458	485	512	539	541	568	625	652	679	706	1289	1287	1291	1292	1293	1320	1337	763	770	831	858	886	914	941	968	995	1022				
227	314	361	408	455	482	489	536	543	570	597	624	1670	1980	2000	1957	1984	2011	2038	2065	2092	2119	2146	2173	2200	952	979	1006	1033	1060	1087	1114	1141	1168	1195	1222	1249	1276	1303	1330	1357	1384	1411	1438	1465	1492				
314	360	407	454	481	488	515	542	569	596	623	650	1671	1981	2001	1958	1985	2012	2039	2066	2093	2120	2147	2174	2201	953	980	1007	1034	1061	1088	1115	1142	1169	1196	1223	1250	1277	1304	1331	1358	1385	1412	1439	1466	1493				
408	455	482	489	495	514	541	568	595	622	649	676	1672	1982	2002	1959	1986	2013	2040	2067	2094	2121	2148	2175	2202	954	981	1008	1035	1062	1089	1116	1143	1170	1197	1224	1251	1278	1305	1332	1359	1386	1413	1440	1467	1494				
455	482	489	496	513	540	567	594	621	648	675	702	1673	1983	2003	1960	1987	2014	2041	2068	2095	2122	2149	2176	2203	955	982	1009	1036	1063	1090	1117	1144	1171	1198	1225	1252	1279	1306	1333	1360	1387	1414	1441	1468	1495				
482	489	512	539	566	593	620	647	674	701	728	755	1674	1984	2004	1961	1988	2015	2042	2069	2096	2123	2150	2177	2204	956	983	1010	1037	1064	1091	1118	1145	1172	1200	1227	1254	1281	1308	1335	1362	1389	1416	1443	1470	1497				
487	494	511	538	565	592	619	646	673	700	727	754	1675	1985	2005	1962	1989	2016	2043	2070	2097	2124	2151	2178	2205	957	984	1011	1038	1065	1092	1119	1146	1173	1201	1228	1255	128228												

Configurazioni matriciali

- Pair-Pattern Matrix
 - Relational *similarity*
 - Pattern *similarity*
 - Paraphrase *detection*
 - NL generation, QA, IR
 - Relational *clustering*
 - Relational *classification*
 - Classificazione semantica di coppie di verbi (es. taint:poison classificato come “strength”, oppure assess:review as “abilitazione”)
 - Classificazione semantica di nomi composti (e.g., flu:virus classified as “cause”, home:town as “location”)
 - Relational *search*
 - “lista tutte le X tale che X causa il cancro”

pattern

wordX:wordY

224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------

X is solved by Y

224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	--

X is solved by Y
X solves Y
X is resolved by Y
...

Similarità: ruolo

- Human cognition fully leverages on the concept of similarity.
- As W. V. Quine stated, *“similarity is fundamental for learning and thought”*, and it *“allows us to order things into kinds so that these can function as stimulus meanings”*.
- We learn from past events for taking decisions in case of new similar events.
- We are able to predict functionalities of unknown objects if they appear similar to something we know

Similarità: tante similarità

- **Semantic Similarity**
 - Sinonimi, o “quasi-sinonimi”
 - Spesso malinterpretata ed usata!
- **Semantic Relatedness**
 - Concetti che condividono proprietà. Esempi meronyms, antonyms, ...
 - Significato molto generico
- **Attributional Similarity**
 - Simile a Relatedness. Miglior nome, ma minor utilizzo
- **Taxonomical Similarity**
 - Concetti che condividono iperonimi
- **Relational Similarity**
 - Similarità tra coppie di parole (i.e., dog:bark, cat:meow)
- **Semantic Association**
 - Parole che co-occorrono frequentemente (i.e., cradle, baby)
 - Simile a Relatedness, ma orientata alla corpus-analysis

*Attenzione alla differenza
tra **parole** e **concetti***

*Spesso si usa il termine **parola**
ma si intende **concetto***

Similarità: tante similarità

Dataset per sperimentazioni:

- SimLex-999 <https://www.cl.cam.ac.uk/~fh295/simlex.html>

Problema principale: word order

- Le rappresentazioni matriciali non tengono in considerazione l'ordine delle parole nel testo:
 - Word Choice vs Word Order – *“Word choice rather than word order plays a more important role in assigning meaning (...) about 80% of the information content could be contained in word choice.”* [Landauer '97]
 - Language-dependent – “Warlpiri” è una lingua australiana che è completamente is order-free [Kashket '86]
- Le matrici pair-pattern sono sensibili all'ordine [Turney '10]
- L'ordine delle parole può essere preso in considerazione attraverso combinazioni vettoriali [Lapata '08]
- L'ordine delle parole può essere preso in considerazione attraverso vettori d'ordinamento [Sahlgren]
- E' possibile rappresentare congiunzione, disgiunzione e negazione usando vettori [Salton '83, Widdow 2004, Rijsbergen 2004]

Secondo problema: rappresentazione non composizionale

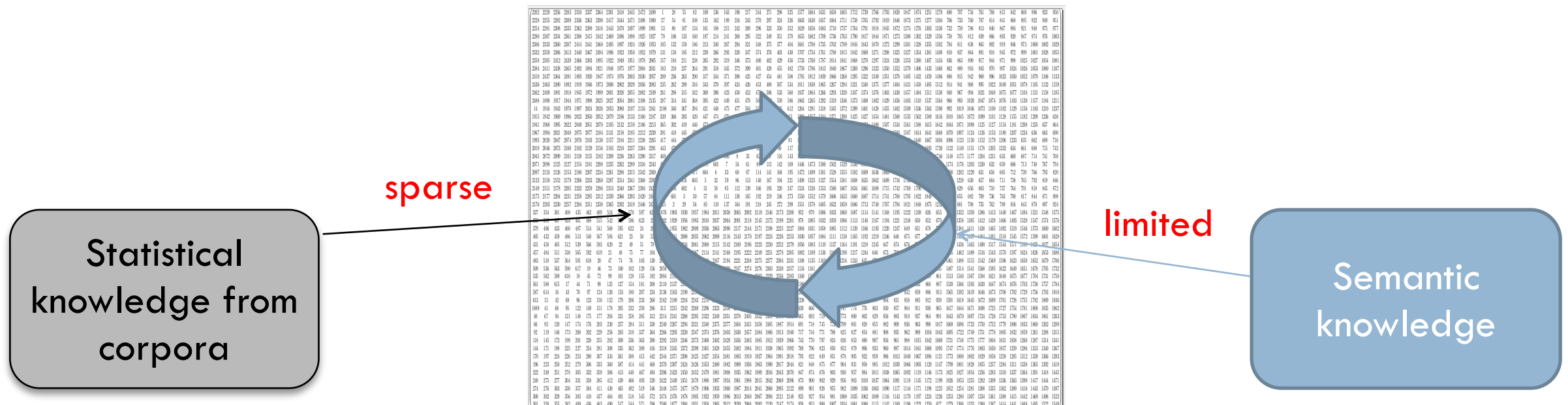
- Le rappresentazioni matriciali sono orientate al significato di singole parole
- Soluzione spesso adottata:
 - Combinazioni vettoriali (addition, multiplication, ...)
- Vi ricordate delle teorie sulla costruzione del significato?

Esempi di altri approcci

- Validità di composizioni nomi-aggettivi [Baroni/Zamparelli 2010]
- Uso di tensori word-link-word [Baroni/Lenci 2010]
- Composizionalità attraverso multiplication / addition for [Lapata 2010]
- Uso di tensori verbi-argumenti [Grefenstette 2011, 2013]

Arricchimento matriciale

- Matrici possono essere costruite da corpus, ed arricchite da risorse semantiche esistenti
- Matrici possono essere costruite in maniera vincolata rispetto a risorse semantiche esistenti
- Entrambe le cose



Significato filamentare

- Filaments of Meaning in Word Space [Jussi Karlgren et al]
 - La semantica delle parole contenuta negli spazi vettoriali è da ritrovarsi in piccole prozioni / filamenti relazionali (local spaces)
 - Il significato di una parola è una rappresentazione ristretta, mentre VSM è una rappresentazione ad alta dimensionalità

Esercizio di laboratorio “*let’splay*”

- **Due** esercizi a scelta tra:
 - **La ghigliottina!** Date 5 parole, trovare la “sesta”
 - **Estrazione automatica di figure retoriche** (sineddoche, iperbole, metafore ecc.) Ad es. usando librerie di Sentiment Analysis per estrazione aggettivi “positivi” legati a sostantivi “negativi” (e viceversa)
 - **Creazione di una nuova lingua!** Rimozione ambiguità attraverso allineamento di più lingue, e creazione del relativo dizionario non ambiguo

Esercizio di laboratorio “*let’splay*”

- **Plagiarism detection (di codice!)** - usare strumenti e conoscenze apprese per creare un sistema che dati due file (di codice python) provi a dare in output delle evidenze di possibile plagio (suggerimenti: usare keyword extraction, metadati, nomi variabili, ...)
- **False friends.** Estrazione di parole simili o identiche (come stringhe) ma diverse come semantica.
- **“Frutta con la M”.** Esercizio in cui si specifica una categoria (ad es. “fruit” oppure “at supermarket” (più difficile) assieme ad una lettera di inizio parola. Creare un gioco in cui il sistema crea combinazioni casuali di input e prova a sfidare sul tempo un utente