

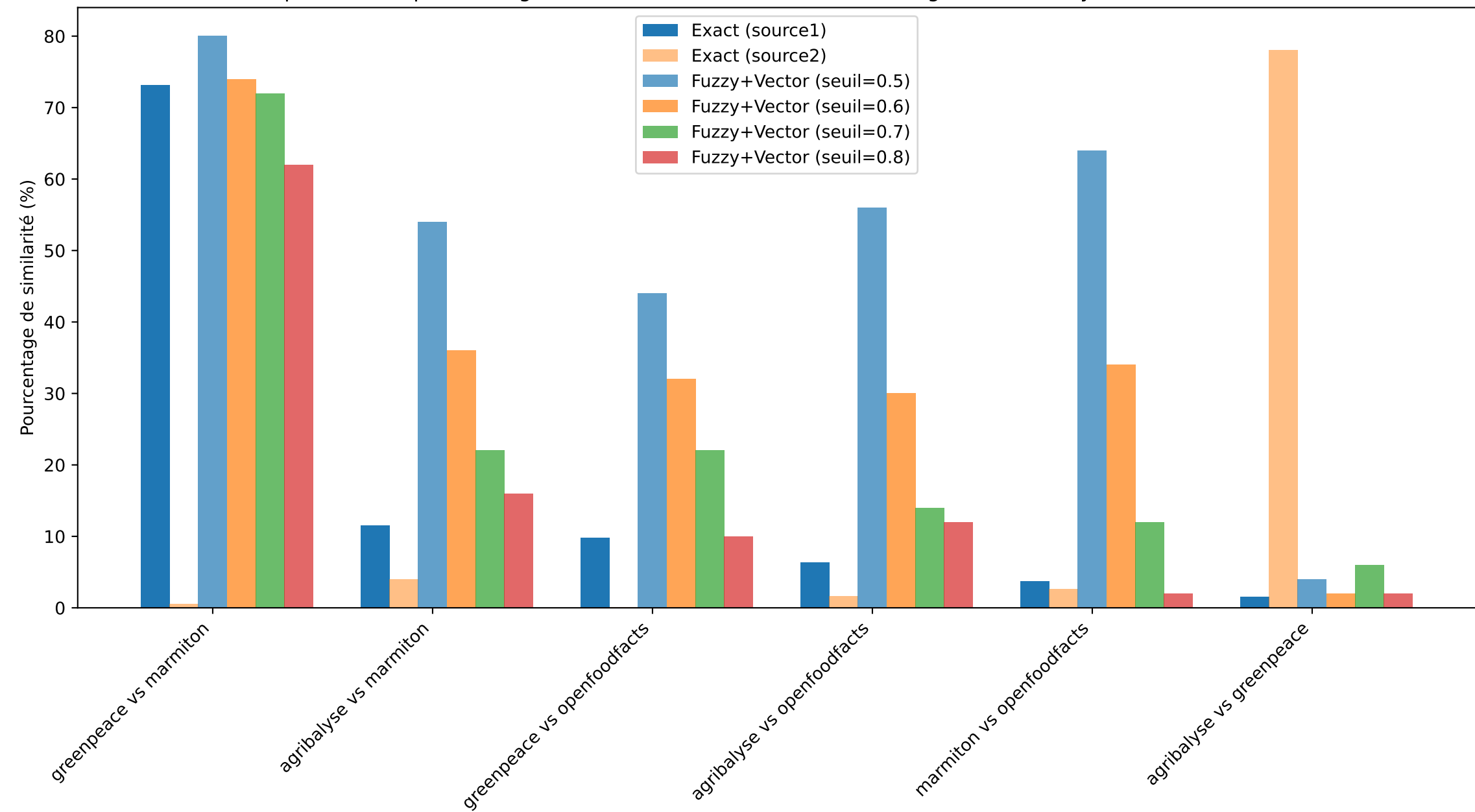
Correspondance exacte entre sources

source1	source2	common_count	total_source1	total_source2	percent_source1	percent_source2
greenpeace	marmiton	30	41	5940	73.17	0.51
agribalyse	marmiton	239	2070	5940	11.55	4.02
greenpeace	openfoodfacts	4	41	8315	9.76	0.05
agribalyse	openfoodfacts	132	2070	8315	6.38	1.59
marmiton	openfoodfacts	218	5940	8315	3.67	2.62
agribalyse	greenpeace	32	2070	41	1.55	78.05

Similarité fuzzy + vector (échantillon, 2 sens, seuil=0.65) entre sources

source1	source2	similar_count_source1	sampled_total_source1	percent_source1	similar_count_source2	sampled_total_source2	percent_source2
agribalyse	greenpeace	19	1000	1.9	36	41	87.8
agribalyse	marmiton	290	1000	29.0	179	1000	17.9
agribalyse	openfoodfacts	220	1000	22.0	65	1000	6.5
greenpeace	marmiton	36	41	87.8	12	1000	1.2
greenpeace	openfoodfacts	12	41	29.27	2	1000	0.2
marmiton	openfoodfacts	160	1000	16.0	85	1000	8.5

Comparaison du pourcentage de similarité entre sources (matching exact et fuzzy+vector, multi-seuils)



Echantillon de paires matchées (global\_score > 0.5)

source1	name1	source2	name2	code_source2	global_score	vector_similarity	fuzzy_similarity	seuil
greenpeace	haricot	marmiton	haricot		1.0	1.0	1.0	0.5
agribalyse	thym	marmiton	thym		1.0	1.0	1.0	0.5
marmiton	sardines mediterrannee	openfoodfacts	sardines	0020100000076	0.589	0.885	0.391	0.5
greenpeace	pomme terre conservation	marmiton	pomme terre		0.573	0.713	0.48	0.5
agribalyse	soupe creme marrons rayon	marmiton	creme marrons		0.611	0.719	0.538	0.5
marmiton	fruits mer	agribalyse	pizza fruits mer	25463	0.718	0.825	0.647	0.5
agribalyse	pate gibier	marmiton	pate		0.513	0.656	0.417	0.5
openfoodfacts	figue	agribalyse	figue	13012	1.0	1.0	1.0	0.5
greenpeace	laitue	marmiton	laitue		1.0	1.0	1.0	0.5
openfoodfacts	new york vanilla cheesecake	agribalyse	cheesecake	19689	0.526	0.726	0.393	0.5
marmiton	pomme terre hachees	greenpeace	pomme terre primeur		0.526	0.649	0.444	0.5
greenpeace	chou	openfoodfacts	chouchou	00361361	0.703	0.686	0.714	0.5
soupe legumes	agribalyse deshydratee reconstituee	openfoodfacts	soupe legumes	0022314451484	0.503	0.78	0.318	0.5
marmiton	restes tomate	openfoodfacts	veloute tomates	0000112302711	0.504	0.616	0.429	0.5
greenpeace	pois	marmiton	pois		1.0	1.0	1.0	0.5

Echantillon de paires matchées (global\_score > 0.6)

source1	name1	source2	name2	code_source2	global_score	vector_similarity	fuzzy_similarity	seuil
openfoodfacts	tabasco hot habanero sauce	marmiton	cafe tabasco sauce		0.612	0.829	0.467	0.6
greenpeace	celeri	openfoodfacts	celery	0033383653228	0.7	0.918	0.556	0.6
marmiton	poisson	agribalyse	poisson pane	26029	0.719	0.798	0.667	0.6
greenpeace	echalote	agribalyse	echalote	20097	1.0	1.0	1.0	0.6
greenpeace	echalote	marmiton	echalote		1.0	1.0	1.0	0.6
agribalyse	confiture lait	openfoodfacts	confiture lait	00000262	1.0	1.0	1.0	0.6
greenpeace	topinambour	marmiton	topinambour		1.0	1.0	1.0	0.6
greenpeace	haricot	agribalyse	haricot	20061_1	1.0	1.0	1.0	0.6
greenpeace	salsifi	marmiton	salsifis		0.795	0.937	0.7	0.6
marmiton	basilic	openfoodfacts	pesto basilic	0001001270350	0.672	0.824	0.571	0.6
agribalyse	hareng	openfoodfacts	hareng fume	0003038680063	0.674	0.809	0.583	0.6
greenpeace	poireau	marmiton	poireau		1.0	1.0	1.0	0.6
agribalyse	estragon	marmiton	estragon		1.0	1.0	1.0	0.6
greenpeace	asperge	agribalyse	asperge	20282	1.0	1.0	1.0	0.6
agribalyse	vermicelle riz	openfoodfacts	vermicelles riz	0001003409007	0.875	0.953	0.824	0.6

Echantillon de paires matchées (global\_score > 0.7)

source1	name1	source2	name2	code_source2	global_score	vector_similarity	fuzzy_similarity	seuil
greenpeace	haricot	marmiton	haricot		1.0	1.0	1.0	0.7
agribalyse	jambon decouenne degraissee	openfoodfacts	jambon decouenne	0002000001797	0.763	0.887	0.68	0.7
openfoodfacts	pistache grille non sale	agribalyse	pistache grillee salee	15009	0.761	0.805	0.731	0.7
greenpeace	poireau	marmiton	poireau		1.0	1.0	1.0	0.7
greenpeace	epinard	openfoodfacts	epinard	0032601952976	1.0	1.0	1.0	0.7
greenpeace	echalote	marmiton	echalote		1.0	1.0	1.0	0.7
greenpeace	oignon	agribalyse	oignon	20034	1.0	1.0	1.0	0.7
greenpeace	aubergine	agribalyse	aubergine	20053	1.0	1.0	1.0	0.7
openfoodfacts	noisette	agribalyse	noisette	15004	1.0	1.0	1.0	0.7
greenpeace	citrouille	agribalyse	citrouille pulpe	20166	0.711	0.807	0.647	0.7
agribalyse	chou chinois	marmiton	chou chinois		1.0	1.0	1.0	0.7
marmiton	truite fumee	agribalyse	truite delevage fumee	27029	0.702	0.869	0.591	0.7
greenpeace	fenouil	agribalyse	fenouil	20028	1.0	1.0	1.0	0.7
greenpeace	concombre	openfoodfacts	thon concombre	00169066	0.704	0.759	0.667	0.7
greenpeace	rutabaga	marmiton	rutabaga		1.0	1.0	1.0	0.7

Echantillon de paires matchées (global\_score > 0.8)

source1	name1	source2	name2	code_source2	global_score	vector_similarity	fuzzy_similarity	seuil
greenpeace	carotte	agribalyse	carotte	20009	1.0	1.0	1.0	0.8
greenpeace	epinard	agribalyse	epinard	20059	1.0	1.0	1.0	0.8
greenpeace	panais	marmiton	panais		1.0	1.0	1.0	0.8
greenpeace	topinambour	marmiton	topinambour		1.0	1.0	1.0	0.8
agribalyse	rutabaga	greenpeace	rutabaga		1.0	1.0	1.0	0.8
greenpeace	panais	agribalyse	panais	20181	1.0	1.0	1.0	0.8
greenpeace	betterave	agribalyse	betterave	20091	1.0	1.0	1.0	0.8
greenpeace	rutabaga	marmiton	rutabaga		1.0	1.0	1.0	0.8
agribalyse	son davoine	openfoodfacts	son davoine	0003175681103	1.0	1.0	1.0	0.8
greenpeace	aubergine	agribalyse	aubergine	20053	1.0	1.0	1.0	0.8
marmiton	coppa	agribalyse	coppa	28850	1.0	1.0	1.0	0.8
greenpeace	echalote	agribalyse	echalote	20097	1.0	1.0	1.0	0.8
marmiton	echalotes	greenpeace	echalote		0.803	0.917	0.727	0.8
greenpeace	haricot	marmiton	haricot		1.0	1.0	1.0	0.8
agribalyse	noisette	openfoodfacts	noisette	00010865	1.0	1.0	1.0	0.8