

► ISCO Revision: Cross-mapping National Occupational Classifications with ISCO

Shutong Ding

ILO Department of STATISTICS

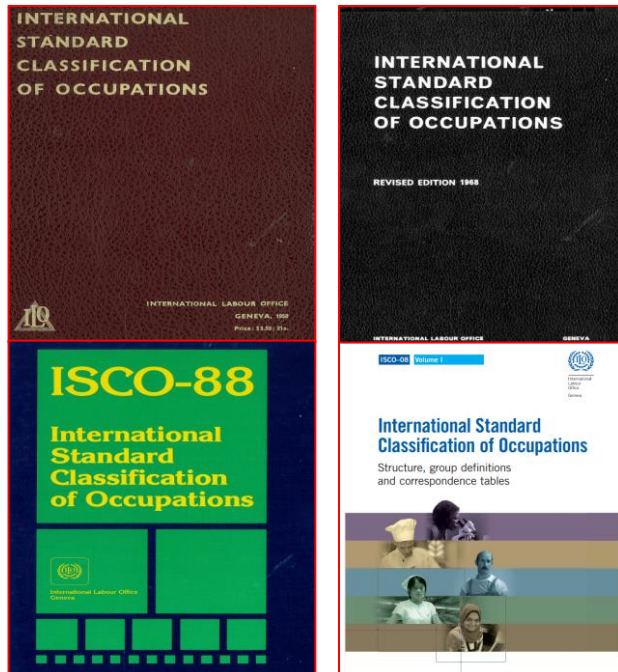
30/11/2024



ISCO revision

Conducted in a 20-year cycle since ISCO-68

- ▶ Requested/ Mandated by the ICLS
- ▶ Facilitated by the office with the assistance of a working group



- ▶ Information and data sources used in revising ISCO:

- ☐ TWG: advice, experience and proposals for improvement,
- ☒ Recently updated National Occupational Classifications (NOCs),
- ☐ Literature review, analysis of available information/data,
- ☐ Ad-hoc consultations with experts and working groups,
- ☐ Etc.

► List of common categories identified across recently updated NOCs/ROCs

Year **2017+**, from across the globe (wide regional representation, various income levels, etc)

- **Americas**: US SOC, Canada, Mexico, Colombia, Chile, etc
- **Europe**: UK, France, Switzerland, Italy, ESCO, etc
- **Africa**: Kenya, etc
- **MENA and Gulf**: KSA, etc
- **Asia / Pacific**: ANZSCO, Singapore, Malaysia, Indonesia, Hong Kong, Viet Nam, Korea, etc

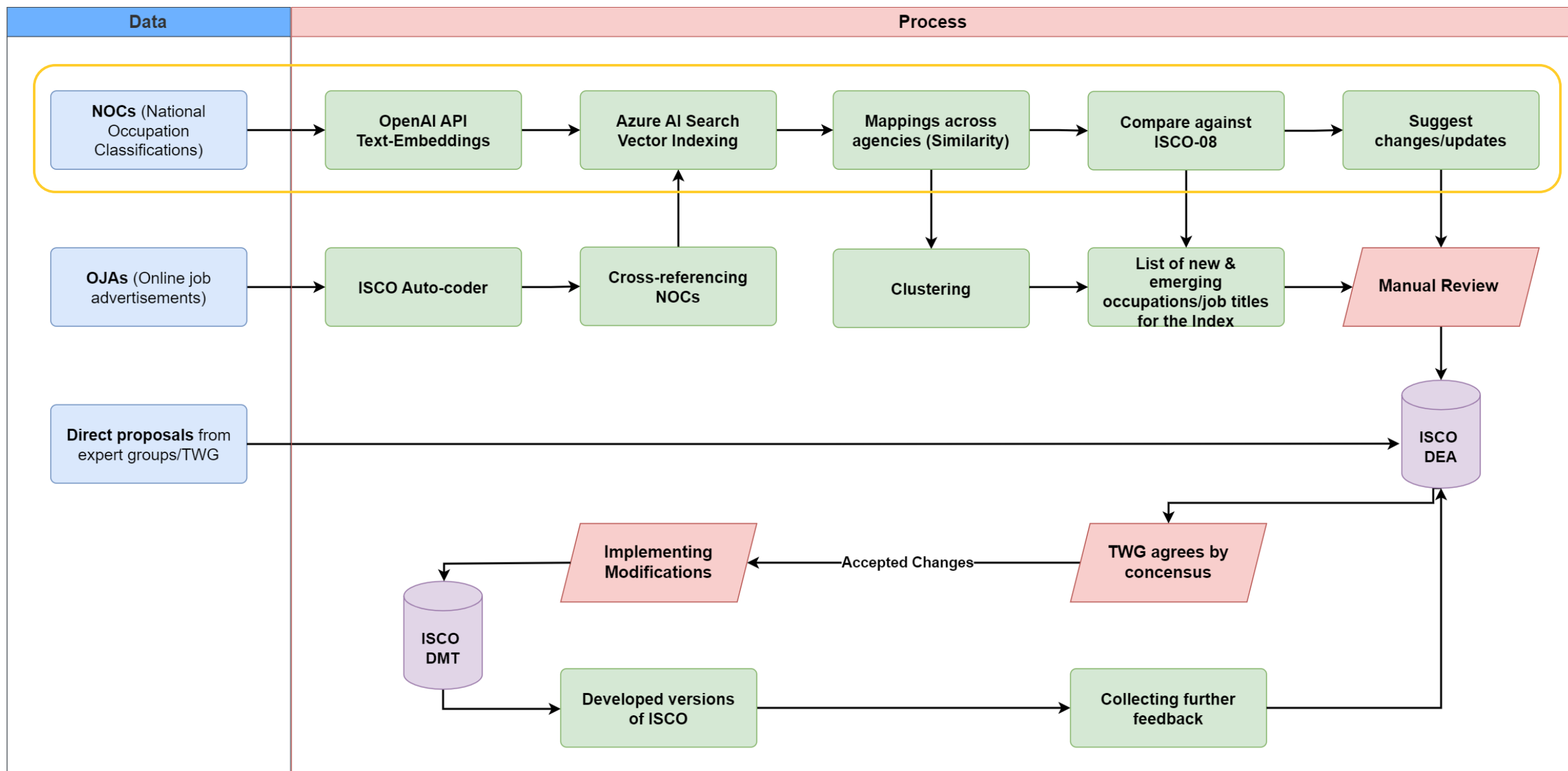
Need to facilitate extensive analysis of NOCs across different languages and volumes!

UK SOC 2020 (extended)		ANZSCO 2022		Singapore SSOC 2024		Canada NOC 2021		US SOC 2018	
2131	IT project managers	26 ICT Professionals		2166	Graphic and Multimedia Designer and Artists	2122 Computer and information systems professionals		15-1200 Computer Occupations	
2131/00	IT project manager	261 Business and Systems Analysts, and Programmers		21661	Graphic designer	21220	Cybersecurity specialists	15-1210 Computer and Information Analysts	
2132	IT managers	2611 ICT Business and Systems Analysts		21662	Multimedia designer/artist/ animator	21221	Business systems specialists	15-1211	Computer Systems Analysts
2132/01	IT information manager	261111	ICT Business Analyst	21669	Graphic/Multimedia designer/artist n.e.c.	21222	Information systems specialists	15-1211	Health Informatics Specialists
2132/02	IT product manager	261112	Systems Analyst			21223	Database analyst and data administrators	15-1212	Information Security Analysts
2132/03	IT service delivery manager	261113	User Experience Designer (ICT)	25	INFORMATION AND COMMUNICATIONS TECHNOLOGY (ICT) PROFESSIONALS	2123 Computer, software and Web designers and developers		15-1220 Computer and Information Research Scientists	
2132/04	IT systems manager	2612 Multimedia Specialists and Web Developers		251	SOFTWARE AND APPLICATIONS DEVELOPERS AND ANALYSTS	21230	Computer systems developers and programmers	15-1221	Computer and Information Research Scientists
2132/05	IT test manager	261211	Multimedia Specialist	2511	Systems Analysts	21231	Software engineers and designers	15-1230 Computer Support Specialists	
2132/06	Network manager	261212	Web Developer	25111	Systems designer/analyst	21232	Software developers and programmers	15-1231	Computer Network Support Specialists
2132/07	Software development manager	2613 Software and Applications Programmers		25112	ICT business process consultant/business analyst	21233	Web designers	15-1232	Computer User Support Specialists
2132/08	Technical support manager	261311	Analyst Programmer	25113	Enterprise/Solution/Software architect	21234	Web developers and programmers	15-1240 Database and Network Administrators and Architects	
2132/99	IT manager n.e.c.	261312	Developer Programmer	25114	ICT product manager			15-1241	Computer Network Architect
2133	IT business analyst, architect and systems designers	261313	Software Engineer	2512	Software, Web, Multimedia and Games Developers and Designers			15-1241	Telecommunications Engineering Specialists
2133/01	Computer analyst and scientist	261314	Software Tester	25121	Software developer			15-1242	Database Administrators
2133/02	Data architect	261315	Cyber Security Engineer	25122	Web/Mobile applications developer			15-1243	Database Architect
2133/03	Data engineer	261316	DevOps Engineer	25123	Multimedia-games developer/designer			15-1243	Data Warehousing Specialists
2133/04	IT systems architect	261317	Penetration Tester	25124	Interaction designer			15-1244	Network and Computer Systems Administrators
2133/05	IT business analyst	261399	Software and Applications Programmers n.e.c.	25125	DevOps engineer			15-1250 Software and Web Developers, Programmers, and Testers	
2133/06	IT solutions architect and designer	262 Database and Systems Administrators, and ICT Security Specialists		2514	Applications and Systems Programmers			15-1251	Computer Programmers
2133/99	IT business analyst, architect and systems designers n.e.c.	2621	Database and Systems Administrators, and ICT Security Specialists	25141	Applications/Systems programmer			15-1252	Software Developers
2134	Programmers and software development professionals	262111	Database Administrator	25142	Artificial intelligence applied researcher			15-1253	Software Quality Assurance Analysts and Testers
2134/01	Computer games designer	262113	Systems Administrator	25143	Artificial intelligence/Machine learning engineer			15-1254	Web Developers
2134/02	Computer programmer	262114	Cyber Governance Risk and Compliance Specialist	2515	Information and Communications Technology (ICT) Testing and Quality Assurance Professionals			15-1255	Web and Digital Interface Designers
2134/03	Software developer	262115	Cyber Security Advice and Assessment Specialist	25151	ICT quality assurance specialist			15-1255	Video Game Designers
2134/99	Programmer and software development professionals n.e.c.	262116	Cyber Security Analyst	25152	ICT auditor			15-1290 Miscellaneous Computer Occupations	
2135	Cyber security professionals	262117	Cyber Security Architect	2519	Software and Applications Developers and Analysts Not Elsewhere Classified			15-1299	Web Administrators
2135/01	Cyber operational defence specialists	262118	Cyber Security Operations Coordinator	25190	Software and applications developer/analyst n.e.c.			15-1299	Geographic Information Systems Technologists and Technicians
2135/02	Cybersecurity management and governance specialists	263 ICT Network and Support Professionals		252	DATABASE, NETWORK, INFRASTRUCTURE AND CYBERSECURITY PROFESSIONALS			15-1299	Document Management Specialists
2135/03	Forensic computer specialists	2631	Computer Network Professionals	2521	Database Designers, Administrators and Data Engineers			15-1299	Penetration Testers
2135/04	Secure system development specialists	263111	Computer Network and Systems Engineer	25211	Data/Database administrator			15-1299	Information Security Engineers
2135/99	Cybersecurity professionals n.e.c.	263112	Network Administrator	25212	Data/Database architect			15-1299	Digital Forensic Analysts
2136	IT quality and testing professionals	263113	Network Analyst	25213	Data engineer			15-1299	Blockchain Engineers
2136/00	IT quality and testing professionals	2632 ICT Support and Test Engineers		2522	Network, Servers, Applications and Computer Systems Administrators			15-1299	Computer Systems Engineers/Architects
2137	IT network professionals	263211	ICT Quality Assurance Engineer	25220	Network/Server/Applications/Computer systems administrator			15-1299 Information Technology Project Managers	
2137/00	IT network professionals	263212	ICT Support Engineer	2523	Computer Network, Infrastructure and Platform Professionals				
2139	Information technology professionals n.e.c.	263213	ICT Systems Test Engineer	25231	Cloud specialist				
2139/01	DevOps engineers	263299	ICT Support and Test Engineers n.e.c.	25232	IT infrastructure specialist				
2139/02	IT consultant	2633 Telecommunications Engineering Professionals		25233	Automation and orchestration engineer				
2139/03	Webmaster and website manager	263311	Telecommunications Engineer	25239	Computer network/infrastructure/platform professional n.e.c.				

UK SOC 2020 (extended)		ANZSCO 2022		Singapore S50C 2024		Canada NOC 2021		US SOC 2018	
2131	IT project managers	26 ICT Professionals		2166	Graphic and Multimedia Designer and Artist	2122 Computer and information systems professionals		15-1200 Computer Occupations	
2131#00	IT project manager	261 Business and Systems Analysts, and Programmers		21661	Graphic designer	21220 Cybersecurity specialists		15-1210 Computer and Information Analysts	
2132	IT managers	2611 ICT Business and Systems Analysts		21662	Multimedia designer/artist/animator	21221 Business systems specialists		15-1211 Computer Systems Analysts	
2132#01	IT information manager	2611#11 ICT Business Analyst		21669	Graphic/Multimedia designer/artist n.e.c.	21222 Information systems specialists		15-1211 Health Informatics Specialists	
2132#02	IT product manager	2611#12 Systems Analyst				21223 Database analysts and data administrators		15-1212 Information Security Analysts	
2132#03	IT service delivery manager	2611#13 User Experience Designer (ICT)		25	INFORMATION AND COMMUNICATIONS TECHNOLOGY (ICT) PROFESSIONALS	2123 Computer, software and Web designers and developers		15-1220 Computer and Information Research Scientists	
2132#04	IT systems manager	2612 Multimedia Specialists and Web Developers		251	SOFTWARE AND APPLICATIONS DEVELOPERS AND ANALYSTS	21230 Computer systems developers and programmers		15-1221 Computer and Information Research Scientists	
2132#05	IT test manager	2612#11 Multimedia Specialist		2511	Systems Analysts	21231 Software engineers and designers		15-1230 Computer Support Specialists	
2132#06	Network manager	2612#12 Web Developer		25111	Systems designer/analyst	21232 Software developers and programmers		15-1231 Computer Network Support Specialists	
2132#07	Software development manager	2613 Software and Applications Programmers		25112	ICT business process consultant/business analyst	21233 Web designers		15-1232 Computer User Support Specialists	
2132#08	Technical support manager	2613#11 Analyst Programmer		25113	Enterprise/Solution/Software architect	21234 Web developers and programmers		15-1240 Database and Network Administrators and Architects	
2132#99	IT manager n.e.c.	2613#12 Developer Programmer		25114	ICT product manager			15-1241 Computer Network Architects	
2133	IT business analysts, architects and systems designers	2613#13 Software Engineer		2512	Software, Web, Multimedia and Games Developers and Designers			15-1241 Telecommunications Engineering Specialists	
2133#01	Computer analysts and scientists	2613#14 Software Tester		25121	Software developer			15-1242 Database Administrators	
2133#02	Data architect	2613#15 Cyber Security Engineer		25122	Web/Mobile applications developer			15-1243 Database Architects	
2133#03	Data engineers	2613#16 DevOps Engineer		25123	Multimedia-games developer/designer			15-1243 Data Warehousing Specialists	
2133#04	IT systems architect	2613#17 Penetration Tester		25124	Interaction designer			15-1244 Network and Computer Systems Administrators	
2133#05	IT business analyst	2613#99 Software and Applications Programmers nec		25125	DevOps engineer			15-1250 Software and Web Developers, Programmers, and Testers	
2133#06	IT solution architect and designer	262 Database and Systems Administrators, and ICT Security Specialists		2514	Applications and Systems Programmers			15-1251 Computer Programmers	
2133#99	IT business analysts, architects and systems designers n.e.c.	262#1 Database and Systems Administrators, and ICT Security Specialists		25141	Applications/Systems programmer			15-1252 Software Developers	
2134	Programmers and software development professionals	262#1#1 Database Administrator		25142	Artificial intelligence applied researcher			15-1253 Software Quality Assurance Analysts and Testers	
2134#01	Computer games designer	262#1#3 Systems Administrator		25143	Artificial intelligence/Machine learning engineer			15-1254 Web Developers	
2134#02	Computer programmer	262#1#4 Cyber Governance Risk and Compliance Specialist		2515	Information and Communications Technology (ICT) Testing and Quality Assurance Professionals			15-1255 Web and Digital Interface Designers	
2134#03	Software developer	262#1#5 Cyber Security Advice and Assessment Specialist		25151	ICT quality assurance specialist			15-1255 Video Game Designers	
2134#99	Programmers and software development professionals n.e.c.	262#1#6 Cyber Security Analyst		25152	ICT auditor			15-1290 Miscellaneous Computer Occupations	
2135	Cyber security professionals	262#1#7 Cyber Security Architect		2519	Software and Applications Developers and Analysts Not Elsewhere Classified			15-1299 Web Administrators	
2135#01	Cyber operational defence specialist	262#1#8 Cyber Security Operations Coordinator		25190	Software and applications developer/analyst n.e.c.			15-1299 Geographic Information Systems Technologists and Technicians	
2135#02	Cybersecurity management and governance specialist	263 ICT Network and Support Professionals		252	DATABASE, NETWORK, INFRASTRUCTURE AND CYBERSECURITY PROFESSIONALS			15-1299 Document Management Specialists	
2135#03	Forensic computer specialist	263#1 Computer Network Professionals		2521	Database Designers, Administrators and Data Engineers			15-1299 Penetration Testers	
2135#04	Secure system development specialist	263#1#1 Computer Network and Systems Engineer		25211	Data/Database administrator			15-1299 Information Security Engineers	
2135#99	Cybersecurity professionals n.e.c.	263#1#2 Network Administrator		25212	Data/Database architect			15-1299 Digital Forensic Analysts	
2136	IT quality and testing professionals	263#1#3 Network Analyst		25213	Data engineer			15-1299 Blockchain Engineers	
2136#00	IT quality and testing professionals	263#2 ICT Support and Test Engineers		2522	Network, Servers, Applications and Computer Systems Administrators			15-1299 Computer Systems Engineers/Architects	
2137	IT network professionals	263#2#1 ICT Quality Assurance Engineer		25220	Network/Server/Applications/Computer systems administrator			15-1299 Information Technology Project Managers	
2137#00	IT network professionals	263#2#2 ICT Support Engineer		2523	Computer Network, Infrastructure and Platform Professionals				
2139	Information technology professionals n.e.c.	263#2#3 ICT Systems Test Engineer		25231	Cloud specialist				
2139#01	DevOps engineers	263#2#99 ICT Support and Test Engineers nec		25232	IT infrastructure specialist				
2139#02	IT consultants	263#3 Telecommunications Engineering Professionals		25233	Automation and orchestration engineer				
2139#03	Webmasters and website managers	263#3#1 Telecommunications Engineer		25239	Computer network/infrastructure/platform professional n.e.c.				
2139#99	Information technology professionals n.e.c.	263#3#2 Telecommunications Network Engineer		2524	Cybersecurity Professionals				

ISCO Revision Information Workflow

6





Pilot study:

**Cross mapping
UK SOC 2020 to ISCO-08**

isco08 large ...

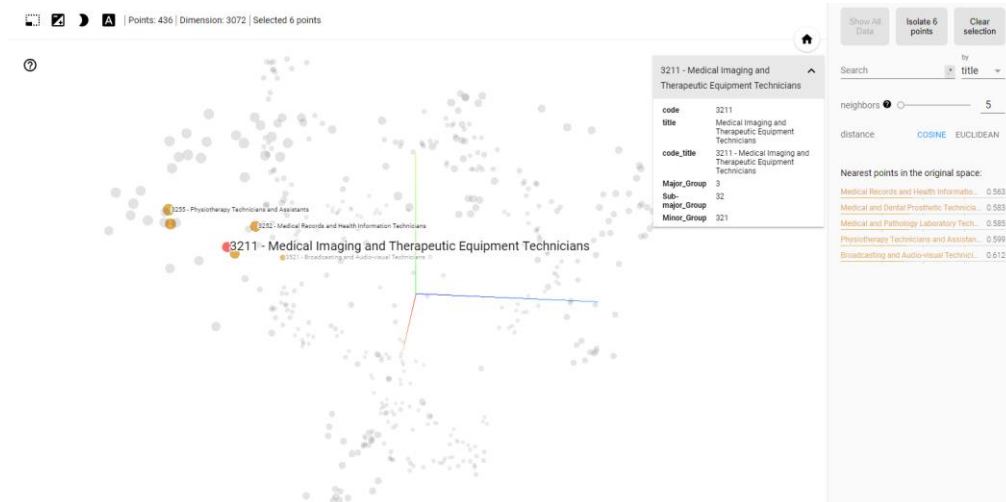


Save Discard Refresh Create demo app Edit JSON Delete Encryption

Search explorer **Fields** CORS Scoring profiles Semantic configurations Vector profiles

 Add field
 Add subfield
 Delete
 Autocomplete settings

Search field names



ISCO-08 structure and index

• ISCO Unit group

- 436 job categories
- aggregated into 130 minor groups, 43 sub-major groups, and 10 major groups
- All variables combined: title, definition, tasks, included/excluded occupations, **index titles**, notes

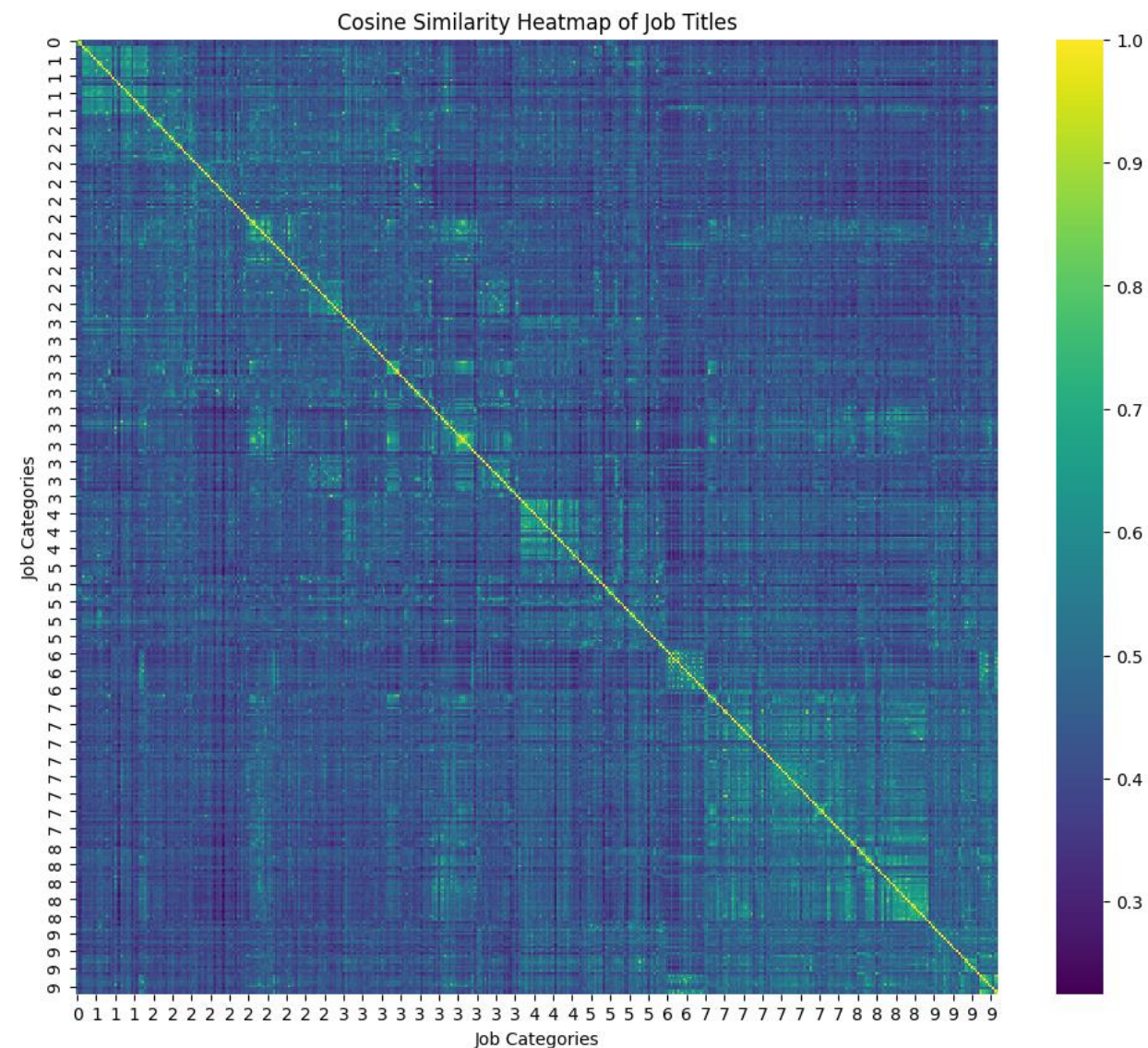
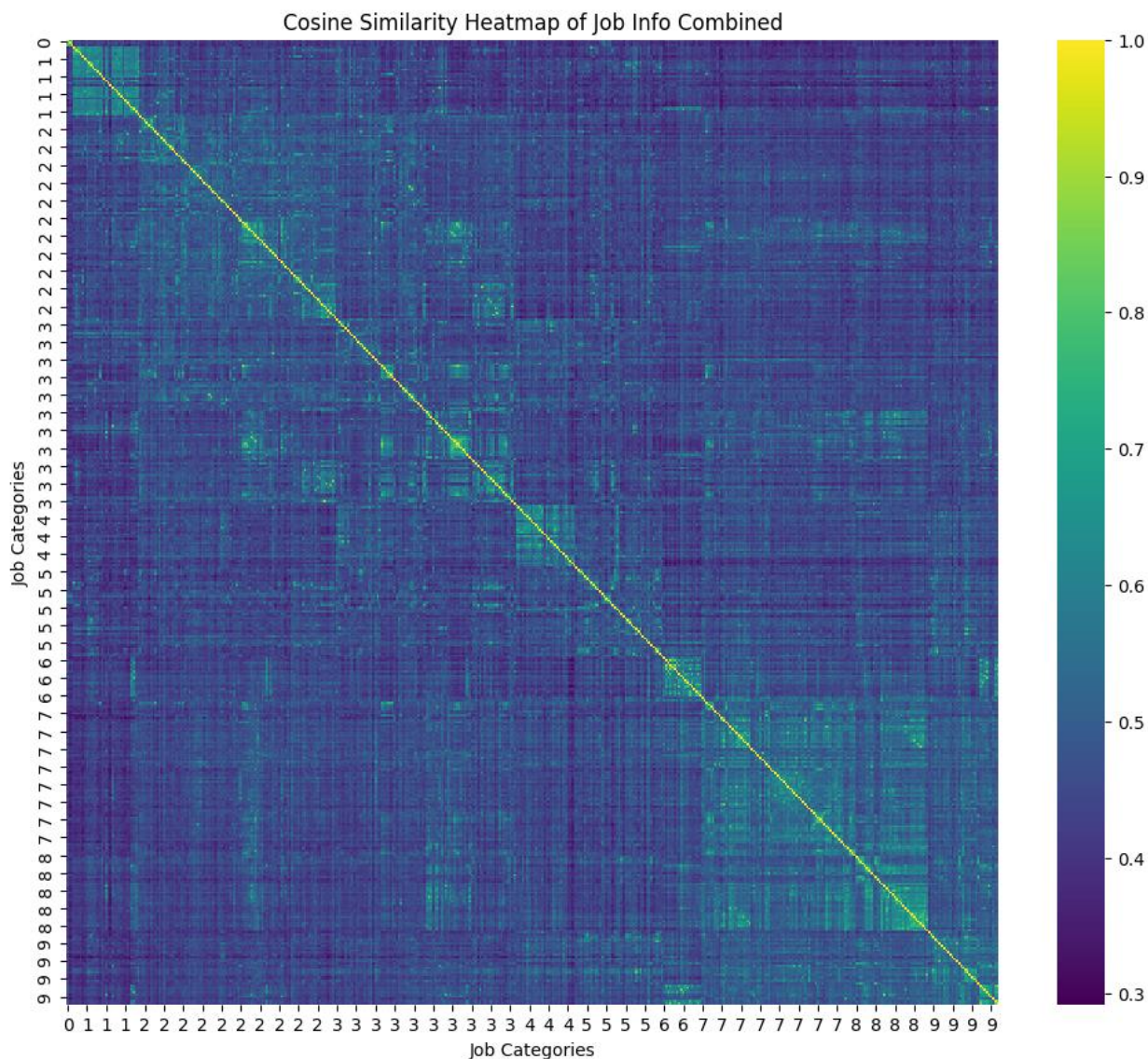
Level	ISCO 08 Code	Title EN	Definition	Tasks include
1	1	Managers	Managers plan, direct, coordinate and evaluate the overall activities of enterprises, governments and other organizations, or of	Tasks performe
2	11	Chief Executives, Senior Officials and Legislators	Chief executives, senior officials and legislators formulate and review the policies,	Tasks performe
3	111	Legislators and Senior Officials	Legislators and senior officials determine, formulate, advise on and direct the implementation of policies of national, state, regional or local governments or	Tasks performe
4	1111	Legislators	Legislators determine, formulate, and direct policies of national, state, regional or local governments and international governmental agencies and make ratify amend or repeal	Tasks include

• ISCO Index

- 7017 job indexes
- English title links to the ISCO-08 code

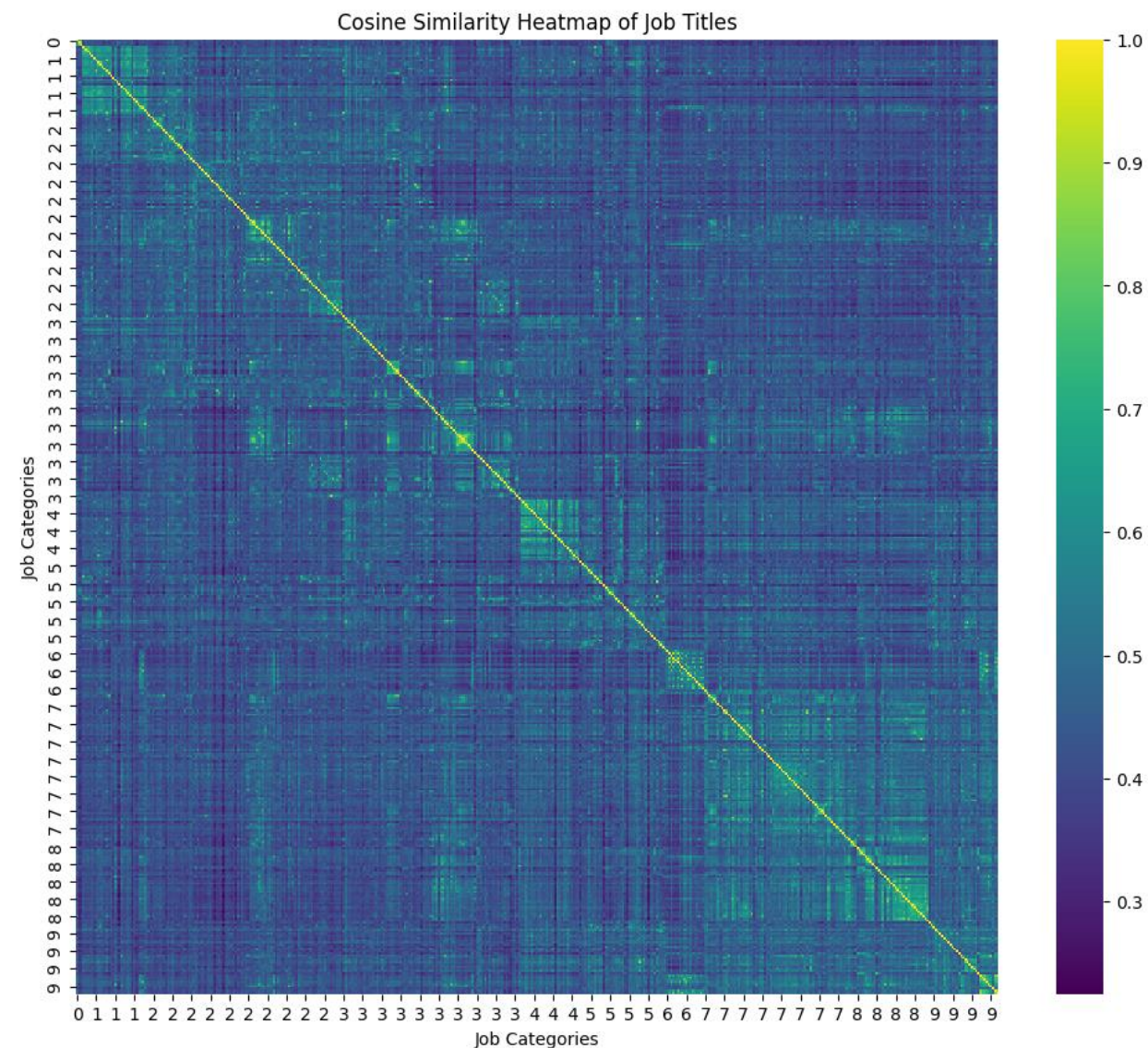
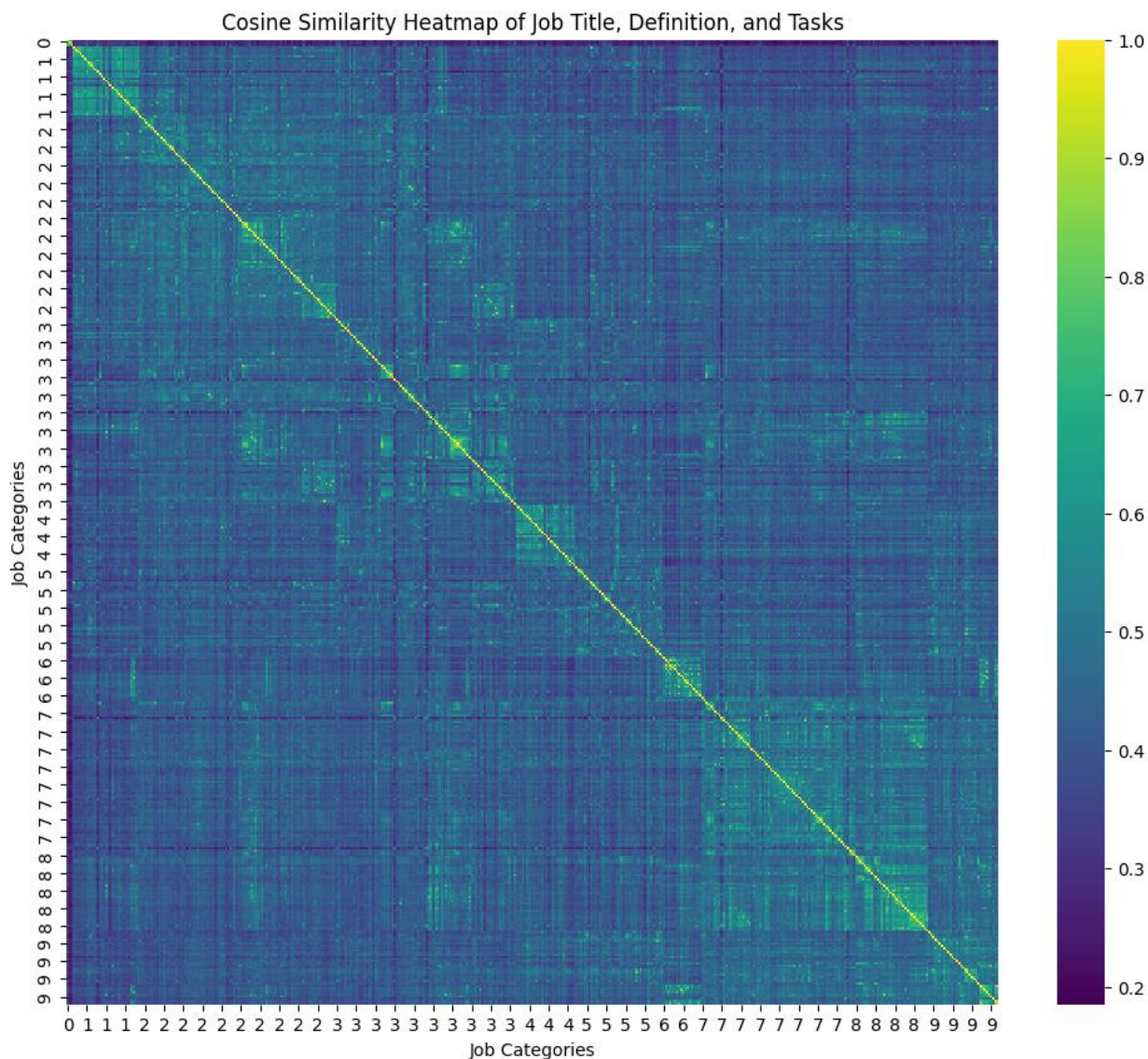
ISCO-08	English title
1111	Alderman
1111	Alderwoman
1111	Chancellor, government
1111	Chief minister, government
1111	Chief whip
1111	Congressman
1111	Congresswoman
1111	Councillor, city

Discriminating Power: Combined Job Information vs. Job Titles¹⁰



Discriminating Power: Title_Definition_Tasks vs. Job Titles

11



UK SOC (Ext2020) – V2 Oct 2023

• SOC Unit group

- 412 job categories (4-digits)
- aggregated into minor groups, sub-major groups, and major groups
- Variables: title, definition, tasks, entry routes and qualifications, **sub-unit group titles**, related job titles

Major Group	Sub-Major Group	Minor Group	Unit Group	Sub-Unit Group	Group Title
1	11	111	1111		MANAGERS, DIRECTORS AND SENIOR OFFICIALS
					CORPORATE MANAGERS AND DIRECTORS
					Chief Executives and Senior Officials
					Chief executives and senior officials
			1111/01		Chairpersons
			1111/02		Chief executives
			1111/03		Diplomats and foreign office officials
			1111/04		Senior public service officials
			1111/99		Chief executives and senior officials n.e.c.
			1112		Elected officers and representatives
			1112/01		Assembly members and Members of Parliament
			1112/02		Councillors
			1112/99		Elected officers and representatives n.e.c.

• SOC Coding Index

- 31570 job indexes
- Index title links to the ISCO-08 code

UNIQUE ID	RECORD	D	SOC_90	SOC_2000	SOC_2010	SOC_2020	SOC_2020_ext	INDEXOCC	ADD	IND	SEE	ISCO 08 code
a1aa	1		103	2441	2429	2439	2439/99	A		Cabinet Office		2422
a1ab	2		103	2441	2429	2439	2439/99	A		Northern Ireland Office		2422
a1ac	3		103	2441	2429	2439	2439/99	A1		Dept for International Development		2422
a1ad	4		400	4112	4112	4111	4111/01	A1		Jobcentre Plus		3353
a1ae	5		400	4112	4112	4111	4111/99	A1		Office for National Statistics		3314
a1af	6		103	2441	2429	2439	2439/99	A2		Dept for International Development		2422
a1aq	7		400	4112	4112	4111	4111/01	A2		Jobcentre Plus		3353
a1ah	8		400	4112	4112	4111	4111/99	A2		Office for National Statistics		3314
a1ai	9		103	3561	3561	3560	3560/99	A3		Dept for International Development		2422
a1aj	10		400	4112	4112	4111	4111/01	A3		Jobcentre Plus		3353
a1ak	11		400	4112	4112	4111	4111/99	A3		Office for National Statistics		3314
a1al	12		400	4112	4112	4111	4111/99	A3		Scottish Government		4110
a1am	13		400	4112	4112	4111	4111/01	A4		Jobcentre Plus		3353
a1an	14		400	4112	4112	4111	4111/99	A4		Scottish Government		4110
a1ao	15		400	4112	4112	4111	4111/01	A5		Jobcentre Plus		3353
a1ap	16		400	4112	4112	4111	4111/01	A6		Jobcentre Plus		3353
a1aq	17		400	4112	4112	4111	4111/01	A7		Jobcentre Plus		3353
a1ar	18		400	4112	4112	4111	4111/99	AA		government		4110
a1as	19		600	3311	3311	3311	3311/99	AB		armed forces		0310
a1at	20		880	8217	8232	8232	8232/00	AB		shipping		8350
f3tf	21		642	6112	6142	6132	6132/00	ACA		ambulance service		3258
f3fq	22		127	1131	1131	2435	2435/00	ACG				1211
a1av	23		127	1131	1131	2435	2435/00	ACIS				1211
a1ax	24		100	1111	1116	1112	1112/01	AM		National Assembly for Wales		1111

Exercise 1: Crosswalk at unit group level

► Batch embeddings

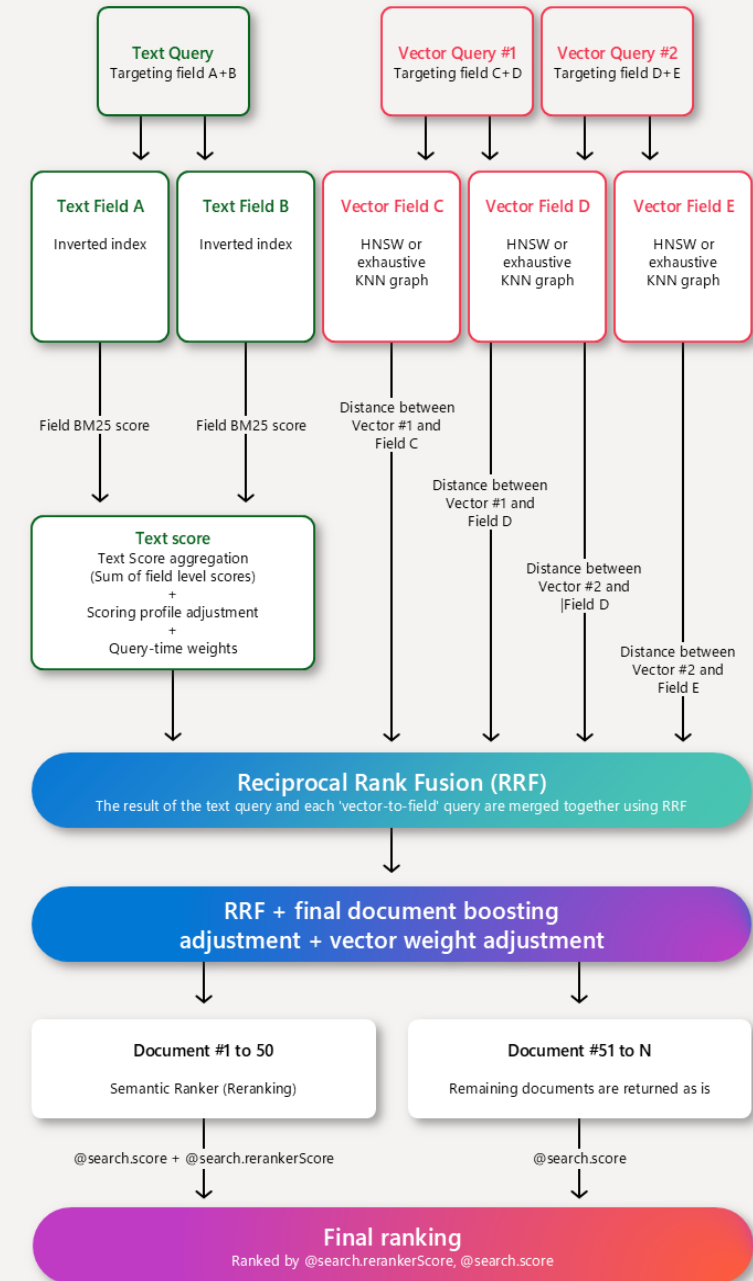
- OpenAI API – Require extra handlers to prepare and retrieve batches

► Search query

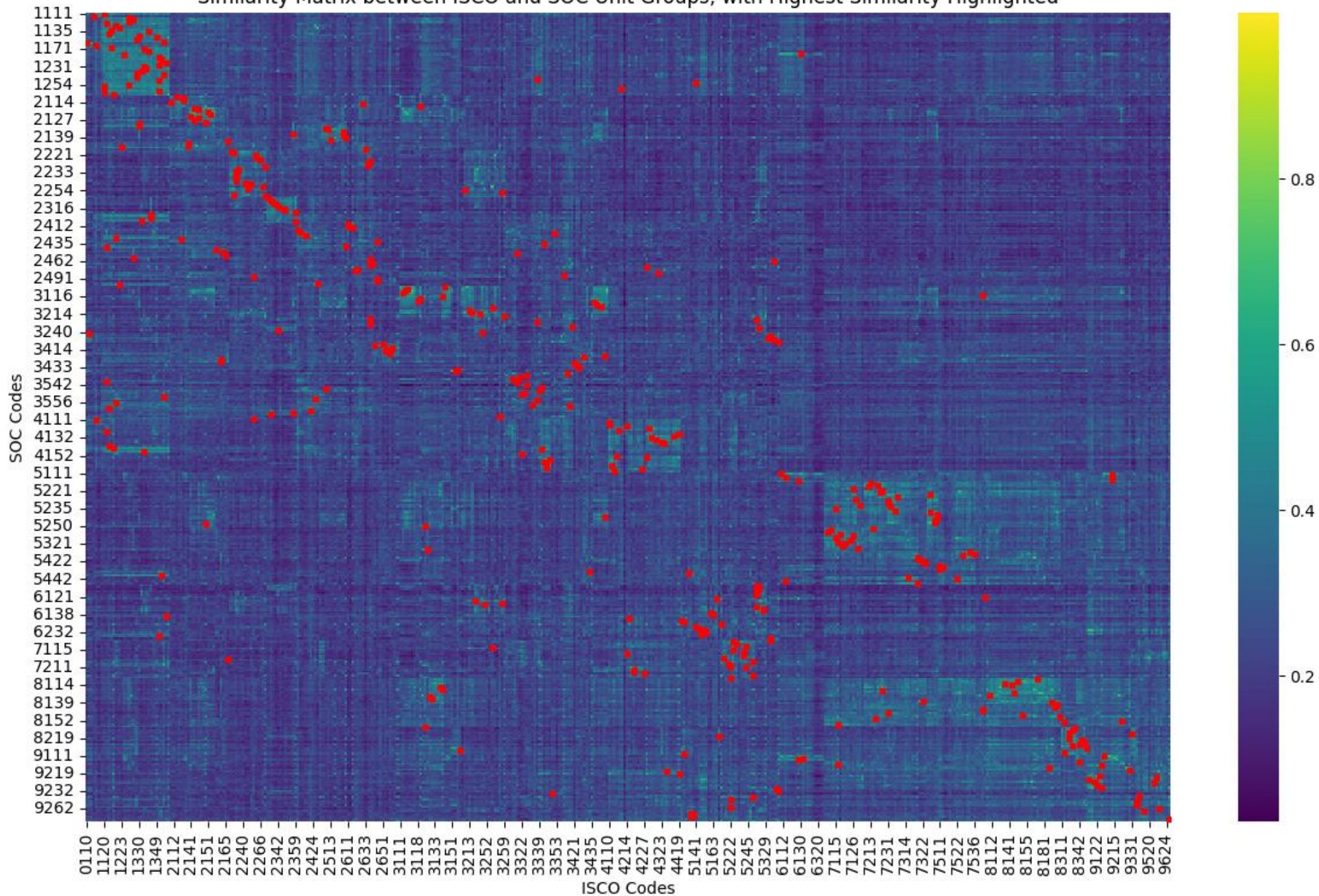
- Vector query -> similarity score
- Hybrid search (text + vector) -> RRF score
- Hybrid search (text + vector) + semantic reranker -> RRF score + Reranker score

► Match – Any unit group identified in ISCO-08?

- Highest search score, or Top_N highest search score



Similarity Matrix between ISCO and SOC Unit Groups, with Highest Similarity Highlighted

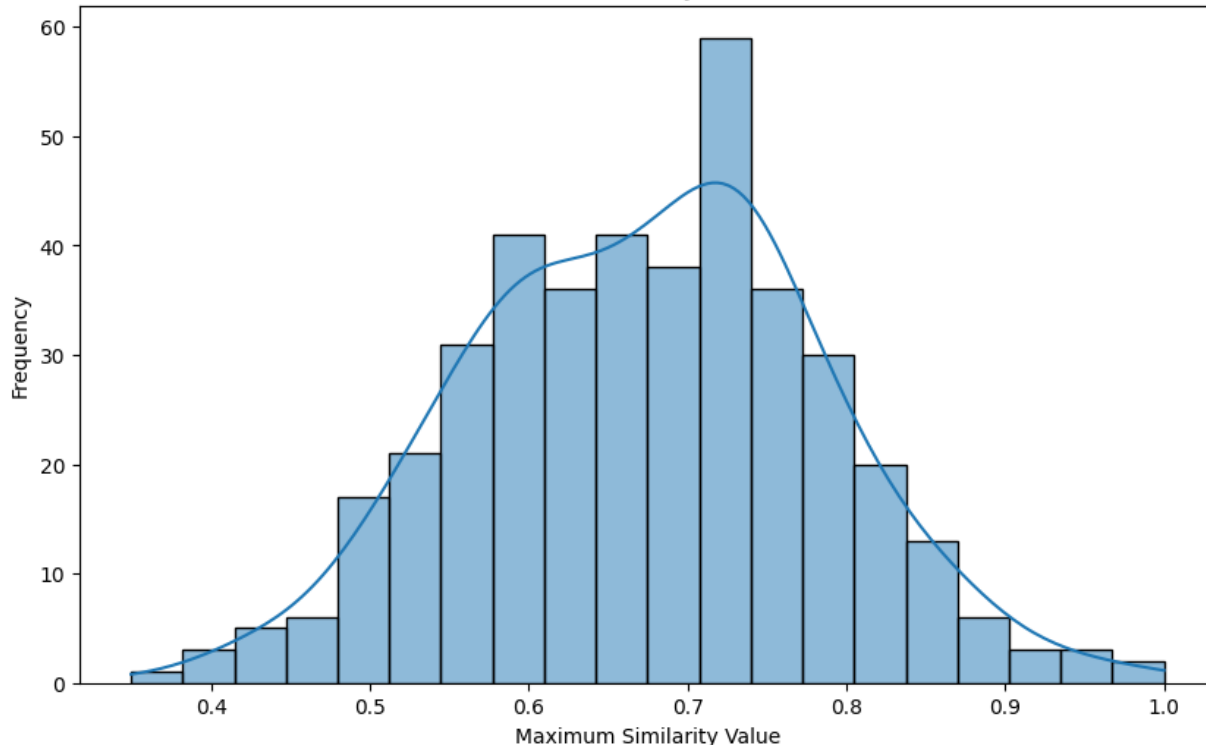


Matching threshold

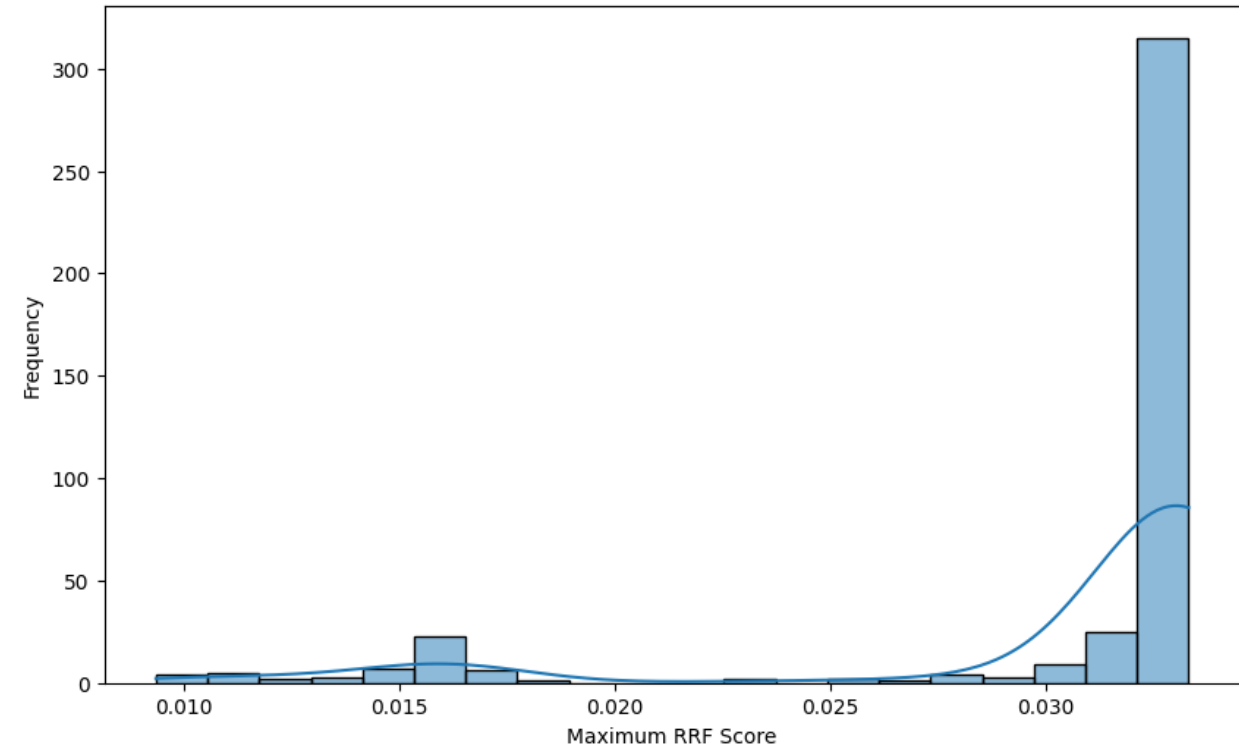
In practice, a difficult match often refer to a `similarity_score < 0.5`, but RRF Score doesn't have a benchmark.

Here we incrementally assess matches by examining matches from the lowest scores upwards to ensure the most challenging matches are considered before setting the threshold.

Distribution of Maximum Similarity Values for Each SOC Code



Distribution of Maximum RRF Scores for Each SOC Code



Matched SOC codes sorted by the lowest RRF Score

SOC_Code	SOC_Title	Predicted_SOC_Code	Predicted_ISCO_title	RRF_Score	Rerank_Score		
1139	Functional managers and directors n.e.c.	2619	Legal Professionals Not Elsewhere Classified	0.00935	2.447		
2482	Quality assurance and regulatory professionals	3115	Mechanical Engineering Technicians	0.00971	1.938		
6136	Senior care workers	5162	Companions and Valets	0.01000	2.493		
5250	Skilled metal, electrical and electronic trades supervisors	2141	Industrial and Production Engineers	0.01075	2.034		
5249	Electrical and electronic trades n.e.c.	3118	Draughtspersons	0.01124	2.495		
2259	Other health professionals n.e.c.	3255	Physiotherapy Technicians and Assistants	0.01149	2.267		
5245	Security system installers and repairers	2152	Electronics Engineers	0.01176	1.783		
8145	Tyre, exhaust and windscreen fitters	5245	Service Station Attendants	0.01515	2.251		
3574	Other vocational and industrial trainers	2423	Personnel and Careers Professionals	0.01563	2.030		
5246	Electrical service and maintenance mechanics and repairers	7411	Building and Related Electricians	0.01563	2.506		
9262	Hospital porters	5329	Personal Care Workers in Health Services Not Elsewhere Classified	0.01563	2.149		
1231	Health care practice managers	3344	Medical Secretaries	0.01587	2.030		
2139	Information technology professionals n.e.c.	3512	Information and Communications Technology User Support	0.01587	2.256		
1111	Chief executives and senior officials	1112	Senior Government Officials	0.01667	2.610		
2126	Aerospace engineers	7232	Aircraft Engine Mechanics and Repairers	0.01667	2.071		
3549	Business associate professionals n.e.c.	2421	Management and Organization Analysts	0.01667	2.110		
3560	Public services associate professionals	1112	Senior Government Officials	0.01667	2.230		
4111	National government administrative occupations	1112	Senior Government Officials	0.01667	2.275		
4143	Customer service managers	4222	Contact Centre Information Clerks	0.01667	1.918		
6137	Care escorts	5162	Companions and Valets	0.02264	1.699		
1253	Hairdressing and beauty salon managers and proprietors	5141	Hairdressers	0.02560	2.063		
1255	Managers and directors in the creative industries	1222	Advertising and Public Relations Managers	0.02618	2.213		
8151	Scaffolders, staggers and riggers	7119	Building Frame and Related Trades Workers Not Elsewhere Classified	0.02889	2.489		

▶ Exercise 2: Crosswalk at index level – Performance Evaluation

▶ Batch embeddings

- ▶ Index title only
- ▶ Index title + unit group title
- ▶ Index title + unit group title with definition, tasks, technical skills, alternative titles generated by GPT-4o

▶ Search query

- ▶ Hybrid search (text + vector) + semantic reranker -> RRF score + Reranker score

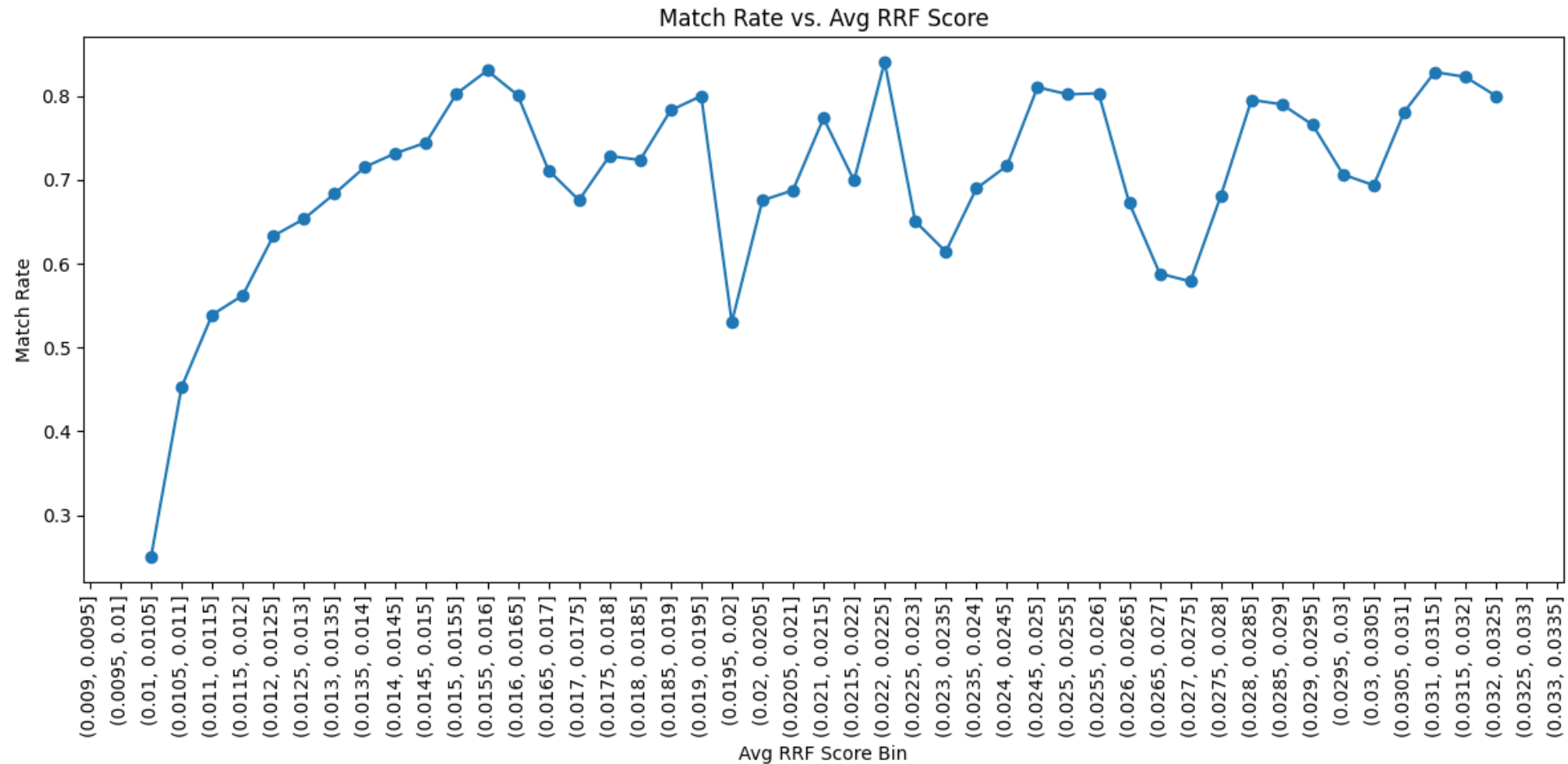
▶ Match Evaluation

- ▶ Highest search score, and Top_N = 5 highest search scores
- ▶ Compare the predicted results against the original mapping

Results - Index title + unit group title with GPT-4o info

SOC Group	Top 1 Match Rate	Top 5 Match Rate	Group Index Count
1	62.49%	77.81%	1762
2	63.97%	79.55%	3683
3	56.87%	74.52%	3571
4	51.41%	67.70%	1910
5	61.63%	80.12%	5403
6	56.71%	76.13%	1148
7	54.08%	74.57%	1054
8	52.54%	76.87%	9718
9	40.98%	61.78%	3072
Overall	55.51%	75.39%	

Lower average RRF scores are associated with lower match rates



SOC Index with the lowest Avg RRF scores within Major Group 4 and 9

SOC_Index	SOC_Code	SOC_Group_Title	ISCO_code	ISCO_title
Rag gatherer woollen mfr	9139	Elementary process plant occupations n.e.c.	9329	Manufacturing Labourers Not Elsewhere Classified
Rights of way officer	4112	Local government administrative occupations n.e.c.	4419	Clerical Support Workers Not Elsewhere Classified
Timber man electricity supplier	9129	Elementary construction occupations n.e.c.	9312	Civil Engineering Labourers
Fitter's Labourer	9139	Elementary process plant occupations n.e.c.	9329	Manufacturing Labourers Not Elsewhere Classified
Leadsman	9139	Elementary process plant occupations n.e.c.	9329	Manufacturing Labourers Not Elsewhere Classified
Pipe Dragger Brickworks	9139	Elementary process plant occupations n.e.c.	9329	Manufacturing Labourers Not Elsewhere Classified
Dockyard Worker	9139	Elementary process plant occupations n.e.c.	9329	Manufacturing Labourers Not Elsewhere Classified
Officer Investigation	9231	Security guards and related occupations n.e.c.	5414	Security Guards
Pontoon Labourer	9259	Elementary storage occupations n.e.c.	9333	Freight Handlers
Ballast Trimmer	9259	Elementary storage occupations n.e.c.	9333	Freight Handlers
Pattern Keeper	9252	Warehouse operatives	9333	Freight Handlers
Breeder Maggot	9119	Fishing and other elementary agriculture occupations n.e.c.	6129	Animal Producers Not Elsewhere Classified
Carver Food	9263	Kitchen and catering assistants	5246	Food Service Counter Attendants
Bellows man shipbuilding	9139	Elementary process plant occupations n.e.c.	9329	Manufacturing Labourers Not Elsewhere Classified
Planning Clerk	4131	Records clerks and assistants n.e.c.	4322	Production Clerks
Ballast man	9259	Elementary storage occupations n.e.c.	9333	Freight Handlers
Writer armed forces	4111	National government administrative occupations n.e.c.	4110	General Office Clerks
Ream tier	9132	Packers, bottlers, canners and fillers	9321	Hand Packers
Syphon man	9139	Elementary process plant occupations n.e.c.	9329	Manufacturing Labourers Not Elsewhere Classified
Street ranger	9231	Security guards and related occupations n.e.c.	5414	Security Guards
Stone mason's labourer	9129	Elementary construction occupations n.e.c.	9313	Building Construction Labourers
Officer British Council	4113	Officers of non-governmental organisations n.e.c.	4110	General Office Clerks
Firing range worker	9267	Leisure and theme park attendants n.e.c.	9629	Elementary Workers Not Elsewhere Classified
Linen checker hotels catering public houses	9259	Elementary storage occupations n.e.c.	9333	Freight Handlers
Glass washer	9139	Elementary process plant occupations n.e.c.	9329	Manufacturing Labourers Not Elsewhere Classified



Concluding remarks

► Sustainability: cost and infrastructure

► Embedding Generation

- Model used: OpenAI text-embedding-3-large, 3072 dimensions, multilingual support
- Cost: \$0.130 / 1M tokens, 50% off with batch processing

► Azure AI Search

- Store ISCO-08 texts and embeddings in AI Search's vector database. Up to 15 indexes, max 15GB
- ~ \$75 / month

► Query & crosswalk execution

- Local PC, Python in VS code, free

Discussion and way forward

▶ Lessons learned

- ▶ Currently it is just **assisted mapping**, still requires human in the loop
- ▶ Establishing thresholds and benchmarks for RRF scores is challenging
- ▶ Embedding generations and large-scale crosswalk on index could involve significant computational resources
- ▶ Be aware of potential information censorship or bias by LLMs

▶ Ongoing process

- ▶ Further refinement and validation are needed (Customized GPT-4o -> search index)
- ▶ Adapting methods for varying NOCs -> broader applicability
- ▶ Voting system based on the candidate lists from all NOCs

► **Thank you!**

 ding@ilo.org



gpt-4o (version:2024-08-06) ▼

Give the model instructions and context ⓘ

You are an expert in the International Standard Classification of Occupations 2008 (ISCO-08). Your task is to classify job titles from national occupational classifications into the appropriate ISCO-08 codes and titles.

Instructions:

Analyze the Job Title: Carefully read the given job title (and description if provided) to understand the nature of the occupation.

Apply changes

Generate prompt



Add section ▼

ⓘ Few shot examples are not used when source data has been added. Any previous examples have been cleared. Once data is removed, then examples can be added.

▼ Add your data PREVIEW

Gain insights into your own data source. Your data is stored securely in your Azure subscription. [Learn more about how your data is protected.](#)

Index:

isco08_large

Search type:

Hybrid + semantic ▼

Select an existing semantic search configuration:

semantic_config ▼