

DiaBizEN

DiaBizEN is a representative sample of the DiaBiz transcriptions localized into English. It includes over 500 high-quality translations of conversations spanning all 9 business domains covered by DiaBiz, amounting to approximately 336,000 words—equivalent to around 10% of the entire DiaBiz corpus. The translations were carried out by experienced translators and proofread by a native English speaker.

Domain	Dialogs	Word count	DiaBiz	Percentage
Banking	127	69 184	773 858	9%
Telecommunications	117	64 805	416 333	16%
Tourism	71	58 626	674 066	9%
Insurance	57	31 009	307 760	10%
Energy services	55	29 740	248 295	12%
Retail	46	25 316	133 702	19%
Medical care	45	22 044	236 057	9%
Debt collection	34	17 776	245 031	7%
Car rental	31	17 199	189 741	9%
Total	583	335 699	3 224 843	10%

A sample of the localised corpus can be [downloaded here](#)

From:

<http://docs.pelcra.pl/> - Group

Permanent link:

http://docs.pelcra.pl/doku.php?id=diabiz_en&rev=1737723344

Last update: **2025/01/24 12:55**

