
WIKIPEDIA – History

Ce renouveau des dons va alors satisfaire la population égyptienne, et les temples, assurant la loyauté de ce contre-pouvoir durant les guerres des Diadoques.

Translation: This renewal of donations will then satisfy the Egyptian population, and the temples, ensuring the loyalty of this counter-power during the wars of the Diadochi.

POLICIES - Contracts

Pour le calcul de la durée de travail effectif hebdomadaire, les heures de présence responsable de jour sont prises en compte après leur conversion en heures de travail effectif.

Translation: To calculate the actual weekly working time, the hours of responsible presence during the day are taken into account after their conversion into actual working hours.

STATEMENTS - Quotes

Les mariages sont écrits dans le ciel.

Translation: Marriages are written in the sky.

LETTERS

Plusieurs fois elle fut recherchée en mariage; mais elle chérissait trop l'indépendance pour contracter un pareil engagement.

Translation: She was sought in marriage several times; but she cherished independence too much to enter into such a commitment.

FORUMS - Reddit

Les syndicats enseignants ont fait part de leurs inquiétudes et de leur surprise face à ces annonces.

Translation: The teachers' unions have expressed their concerns and surprise at these announcements.

RESEARCH - Science & Engineering

Certains champignons (biotrophes) vivent sur la cellule vivante, d'autres (nécrotrophes) la tuent avant de s'en nourrir.

Translation: Some fungi (biotrophs) live on the living cell, others (necrotrophs) kill it before feeding on it.

Figure 14: French sentence examples from README++.

WIKIPEDIA – Mathematics

Несмотря на видимую простоту многих из них, такие доказательства используют свойства площадей фигур, доказательства которых сложнее доказательства самой теоремы Пифагора.

Translation: Despite the seeming simplicity of many of them, these types of proofs use properties of areas of figures, proofs of which are harder than the proofs of the Pythagorean theorem itself

LITERATURE - Poetry

И на суше и водах, веслом и посохом, будут песнь и молитва бездны пасти, и осядет прахом взметенное порохом, и домой вернется пропавший без вести.

Translation: And on the land and sea, oar and staff, there will be songs and prayers at the abyss' mouth, and that which was thrown up by gunpowder will settle into dust, and the missing in action will return home.

ENTERTAINMENT - Jokes

Сколько гостя не корми, он все равно напьется.

Translation: No matter how much you feed a guest, he will still get drunk.

SPEECH - TED Talks

В начале истории Америки характеру лидера придавалось большое значение, и мы ценили людей с богатым внутренним миром и высокой нравственностью.

Translation: At the start of American history, the character of the leader was given more value, and we valued people with rich inner peace and high morality.

GUIDES - Cooking Recipes

Кислота кваса должна приятно дополнять пресный или солоноватый вкус рыбы, а не противоречить ему.

Translation: The acidity of the Kvass should pleasantly add to the fresh or salty taste of the fish, not counteract it.

RESEARCH - Law

Русское общество с огромным интересом следило за первыми шагами нового суда.

Translation: Russian society watched the first steps of the new court with immense interest.

Figure 15: Russian sentence examples from README++.

Domain	# Sentences				
Sub-Domain	ar	en	fr	hi	ru
WIKIPEDIA					
History	50	50	50	22	50
Geography	50	50	50	31	50
Philosophy	49	47	50	34	50
Technology	43	50	50	19	50
Mathematics	43	50	32	23	50
Art & Culture	49	50	50	35	50
Social Sciences	48	50	50	41	50
Natural Sciences	49	49	50	38	50
Health & Fitness	49	49	50	40	50
NEWS ARTICLES					
Sports	46	46	—	—	—
Politics	13	44	—	—	—
Culture	50	50	—	—	—
Economy	41	50	—	—	—
Technology	36	50	—	—	—
RESEARCH					
Law	36	19	—	13	50
Politics	19	22	—	19	50
Medical	—	30	31	—	50
Literature	—	39	—	28	—
Economics	26	46	—	31	50
Science & Engineering	—	30	47	—	50
LITERATURE					
Novels	50	50	50	48	50
History	40	45	50	47	—
Biographies	26	47	—	46	—
Children's Books	50	49	50	44	—
TEXTBOOKS					
Business	35	50	—	47	—
Psychology	—	50	—	47	—
Agriculture	—	50	—	—	—
Engineering	—	50	—	—	—
USER REVIEWS					
Products	50	40	—	33	49
Books	50	47	—	—	—
Movies	—	50	—	43	—
Hotels	50	48	—	—	—
Restaurants	50	47	—	—	—
DICTIONARIES	40	40	—	—	—

Domain	# Sentences				
Sub-Domain	ar	en	fr	hi	ru
FORUMS					
Reddit	39	50	50	49	50
QA Websites	28	48	50	47	50
StackOverflow	—	50	—	—	—
SOCIAL MEDIA					
Twitter	41	47	50	44	49
POLICIES					
Contracts	27	34	45	—	41
Olympic Rules	40	50	50	—	50
Code of Conduct	—	50	—	50	—
GUIDES					
User Manuals	50	46	50	28	50
Online Tutorials	51	47	50	44	50
Cooking Recipes	40	48	50	47	50
Code Documentation	—	49	—	—	—
CAPTIONS					
Images	50	50	47	48	44
Videos	—	50	50	50	—
Movies	27	41	50	46	—
YouTube	—	42	—	—	—
MEDICAL TEXT					
Clinical Reports	—	39	—	—	—
ENTERTAINMENT					
Jokes	50	50	—	46	49
SPEECH					
Ted Talks	49	43	50	48	50
Public Speech	35	47	—	45	30
STATEMENTS					
Rumours	20	40	—	39	—
Quotes	50	50	50	49	50
DIALOGUE					
Open-domain	39	44	50	39	49
Negotiation	—	45	—	—	—
Task-oriented	39	50	50	50	—
LEGAL					
Constitutions	43	30	50	34	50
Judicial Rulings	—	21	—	35	47
UN Parliament	39	43	50	—	50
FINANCE	—	50	50	—	50
POETRY	46	50	50	49	50
LETTERS	—	22	50	—	50

Table 15: Dataset Statistics. (—) denotes that no public resource was found in the particular language.