The Crate at Warehouse 4

Example Author

1.1 Corpus metrics

Linguistic metric	Measured value	Interpretation & reference
Total word count (full text incl. appendix)	1,748 words (9,608 characters)	Full text including the appendix. (~7.0 norm pages à 250 words / ~6.4 norm pages à 1,500 characters).
Main text word count (prose only)	1,718 words (9,466 characters)	Prose only. (6 chapters).
Total sentences (main text)	193 sentences	Total sentences of the main text.
Average sentence length (ASL)	8.90 words / sentence	Average sentence length of the main text.
Median sentence length	7 words	Median sentence length.
Sentence-length standard deviation	7.83 words	Spread of sentence lengths.
Lexical diversity (TTR)	0.3174 (V = 545 / N = 1,717)	Lexical diversity (type-token ratio).
Guiraud index R (V/√N)	13.15	V/√N; still sensitive to text length and topic.
Yule's Characteristic K	169.64	Vocabulary stability across text length.
MTLD (Textual Lexical Diversity)	91.2	Length-robust lexical diversity; short samples remain unstable.
MATTR (Moving-Average TTR)	0.806	Window-stable type-token ratio (Covington & McFall 2010).
Maas a^2	0.0476	Vocabulary concentration (Maas 1972; lower = richer).
Flesch Reading Ease	86.9	Language-calibrated readability formula of the active profile.
LIX (readability index)	24.3	Readability index (LIX).
Dialogue share (quoted speech)	310 words (18.05%)	Share of quoted speech in the prose.
Prose paragraphs	65 (26.4 words/paragraph)	Paragraph economy of the main text.

1.2 Sentence-length architecture & rhythm

The empirical sentence distribution of the main text:

Sentence-length category	Criterion	Sentences	Share	Interpretation limit
Short sentences (staccato)	≤ 6 words	89	46.1%	Short sentences; pace requires contextual review.
Medium sentences	7–15 words	79	40.9%	Sentence-length category, not a narrative judgement.
Long sentences	16–25 words	23	11.9%	Sentence-length category, not a narrative judgement.
Very long sentences	> 25 words	2	1.0%	Length does not establish syntactic nesting.

1.3 Punctuation as a stylistic seismograph

Punctuation mark	Frequency	Density (per 1,000 words)	Interpretation limit
Periods (.)	185	107.7	Basic sentence rhythm.
Commas (,)	56	32.6	Enumerations and appositions.
Dashes (–/—)	0	0.0	Afterthought, self-correction, insertion.
Colons (:)	2	1.2	Announcing lists, quotes, insights.
Semicolons (;)	1	0.6	Rare; avoids a lecturing tone.
Question marks (?)	8	4.7	Self-questioning and dialogue.
Exclamation marks (!)	0	0.0	Used sparingly.
Ellipses (…/...)	0	0.0	Breaking off a thought.

1.4 Chapter comparison matrix (linguistic deep profiles)

Ch.	Title	Words	Sentences	ASL	Dialogue %	TTR	Filter verbs	Dominant tense

Ch.	Title	Words	Sentences	ASL	Dialogue %	TTR	Filter verbs	Dominant tense
01	The Eastern Gate	203	17	11.9	0.0%	0.631	1	Past
02	The Loose Board	233	41	5.7	1.7%	0.618	0	Past
03	Questions on the Quay	401	48	8.4	64.6%	0.489	1	Past
04	The Customs Note	184	17	10.8	0.0%	0.652	0	Past
05	The Blank Line	386	40	9.7	3.6%	0.503	3	Past
06	Morning Handover	310	30	10.3	10.6%	0.552	0	Past