

Объём исследования

Анализ включает 63 романа Стивена Кинга на английском языке.

1974	Carrie	1984	Thinner	1997	Wizard and Glass	2012	The Wind Through
1975	'Salem's Lot	1986	It	1998	Bag of Bones	the Ke	yhole
1977	The Shining	1987	The Drawing of the	1999	The Girl Who	2013	Joyland
1977	Rage	Three		Loved	Tom Gordon	2013	Doctor Sleep
1979	The Long Walk	1987	Misery	2001	Dreamcatcher	2014	Mr. Mercedes
1979	The Dead Zone	1987	The	2001	Black House	2014	Revival
1980	Firestarter		yknockers	2002	From a Buick 8	2015	Finders Keepers
1981	Roadwork	1989	The Dark Half	2003	Wolves of the Calla	2016	End of Watch
1981	Cujo	1990	The Stand	2004	Song of Susannah	2017	Gwendy's Button
1982	The Running Man	1991	The Waste Lands	2004	The Dark Tower	Box	
1982	The Gunslinger	1991	Needful Things	2005	The Colorado Kid	2017	Sleeping Beauties
1983	Christine	1992	Gerald's Game	2006	Cell	2018	The Outsider
1983	Pet Sematary	1992	Dolores Claiborne	2006	Lisey's Story	2018	Elevation
1983	Cycle of the	1994	Insomnia	2007	Blaze	2019	The Institute
Werev	•	1995	Rose Madder	2008	Duma Key	2021	Later
1984	The Talisman	1996	The Green Mile	2009	Under the Dome	2021	Billy Summers
1984	The Eyes of the	1996	Desperation	2011	11/22/63	2022	Gwendy's Final
Dragon		1996	The Regulators		. ,	Task	

Метод сравнения

Для анализа использовался инструмент **TfidfVectorizer** из библиотеки sklearn. Он превращает документы в матрицу чисел **TF-IDF** (term frequency — inverse document frequency; частота термина — обратная частота документа).

TF-IDF — мера оригинальности (веса) слова, при которой частота слова в документе сравнивается с количеством документов, в которых встречается слово. Оценивается только частота слов в документе; при этом информация об относительном положении слов игнорируется. В итоге для каждой пары книг получается число от 0 (наименьшее сходство) до 1 (наибольшее сходство).

Самые похожие романы

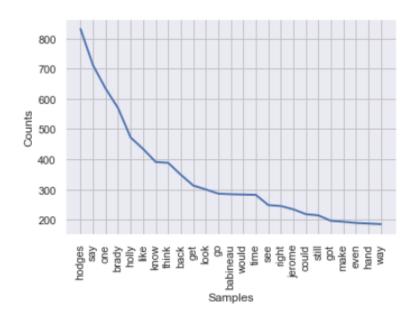
Оказалось, что все из наиболее похожих романов в парах похожих произведений (степень сходства ≥ 0,95) **написаны после 2009 года** (кроме книги «Мешок с костями»).

			Similarity
Pair 1	End Of Watch (2016)	Finders Keepers (2015)	0.966924
Pair 2	Gwendy's Button Box (2017)	Gwendy's Final Task (2022)	0.965771
Pair 3	Joyland (2013)	Under The Dome (2009)	0.964749
Pair 4	Joyland (2013)	Revival (2014)	0.956291
Pair 5	Joyland (2013)	Bag Of Bones (1998)	0.951275
Pair 6	Joyland (2013)	The Outsider (2018)	0.950588
Pair 7	Revival (2014)	Under The Dome (2009)	0.950219

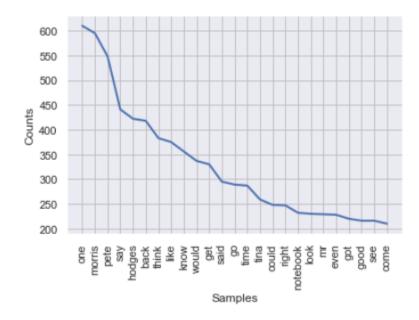
Самые похожие романы

End Of Watch и **Finders Keepers** — это две части трилогии о Билле Ходжесе, которые вышли одна за другой. Логично, что они похожи.

Word frequency in the novel End Of Watch



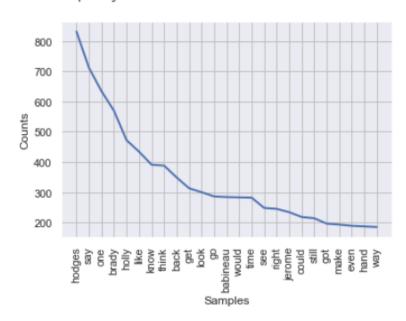
Word frequency in the novel Finders Keepers



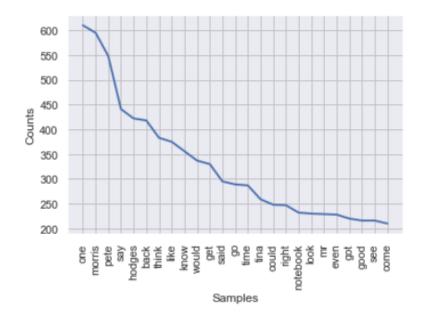
Самые похожие романы

Gwendy's Final Task — продолжение романа **Gwendy's Button Box.** Тоже неудивительно, что у них много сходства по частоте слов.

Word frequency in the novel End Of Watch



Word frequency in the novel Finders Keepers



Однако **Joyland** сильно похожа сразу на четыре романа, с которыми не входит в один цикл: Bag Of Bones, Under The Dome, Revival, Outsider. При этом последние три также все похожи между собой.

Сходство внутри цикла The Dark Tower

- 0.9

- 0.8

- 0.7

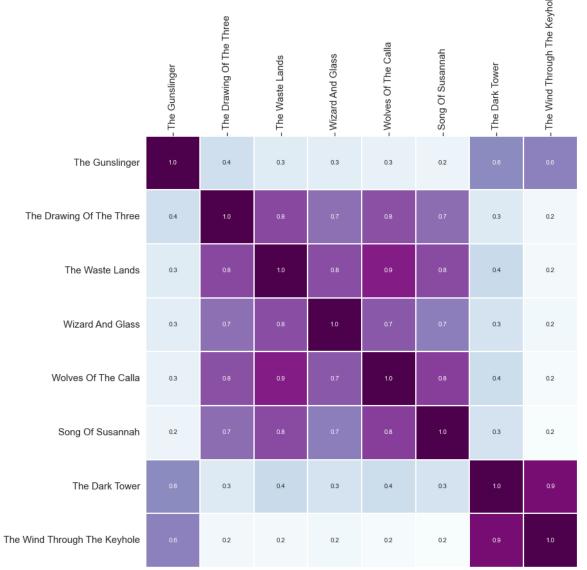
- 0.6

- 0.5

- 0.4

- 0.3

- 0.2



Сходство внутри цикла The Dark Tower

Выделяются две группы книг:

- **1)** The Drawing Of The Three / «Извлечение троих», The Waste Lands / «Бесплодные земли», Wizard and Glass / «Колдун и кристалл», Wolves Of The Calla / «Волки Кальи» и Song of Susannah / «Песнь Сюзанны» оказались довольно похожи между собой: их пары имеют степень сходства 0,7–0,9.
- 2) The Gunslinger / «Стрелок», The Dark Tower / «Тёмная башня» и The Wind Through The Keyhole / «Ветер сквозь замочную скважину» стоят особняком: они похожи друг на друга, но не похожи на остальные книги цикла. «Стрелок» самый первый роман цикла, «Тёмная башня» последний, а «Ветер сквозь замочную скважину» был написан вообще после завершения основного цикла, действие этого произведения происходит в промежутке между концом четвёртой и началом пятой книг.

Романы, написанные в соавторстве

Пять романов были написаны в соавторстве:

- The Talisman / «Талисман» (1984, с Питером Страубом)
- Black House / «Чёрный дом» (2001, с Питером Страубом)
- Gwendy's Button Box / «Гвенди и её шкатулка» (2017, с Ричардом Чизмаром)
- Sleeping Beauties / «Спящие красавицы» (2017, с Оуэном Кингом)
- Gwendy's Final Task / «Последняя задача Гвенди» (2022, с Ричардом Чизмаром)

Анализ показал следующее:

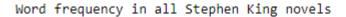
- Как и ожидалось, «Талисман» и «Чёрный дом» оказались довольно отличными от других романов Кинга (степень сходства < 0,5)
- «Гвенди и её шкатулка» и «Последняя задача Гвенди» похожи в основном только друг на друга
- Однако роман «Спящие красавицы», написанный в соавторстве с сыном писателя, похож на сразу множество произведений Стивена Кинга

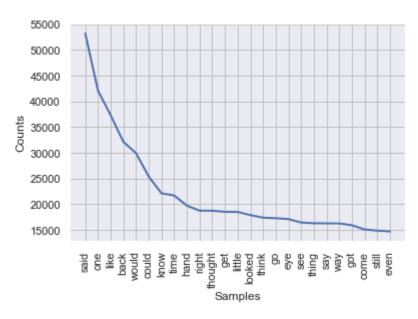
Романы, похожие на Sleeping Beauties

Следующие романы Стивена Кинга очень похожи на Sleeping Beauties (степень сходства ≥ 0.85):

Joyland (2013)	0.949744	The Wind Through The Keyhole (2012)	0.914879
Under The Dome (2009)	0.945271	The Colorado Kid (2005)	0.913519
Revival (2014)	0.938211	Cell (2006)	0.913498
The Outsider (2018)	0.935828	Billy Summers (2021)	0.911454
Bag Of Bones (1998)	0.933672	The Regulators (1996)	0.910473
Needful Things (1991)	0.928754	The Dark Tower (2004)	0.910413
Later (2021)	0.928649	Elevation (2018)	0.907241
Duma Key (2008)	0.928171	Blaze (2007)	0.895995
Doctor Sleep (2013)	0.923030	Mr Mercedes (2014)	0.887420
The Institute (2019)	0.916117	Insomnia (1994)	0.874596

Частота слов во всех романах Стивена Кинга





Контакты

■ GITHUB: github.com/kooznitsa

■ EMAIL: jukuznetswork@gmail.com