

Отчет по заданию “Robot”

Кукуев Михаил, АИСОБОИ-2

Реализация обхода графа находится в файле `crawl.py`, вычисление PageRank и HITS в `rank.py`. Файлы запускаются без каких-либо аргументов. Для выполнения `rank.py` необходимо наличие в корневой директории файла `wiki_pages.bin` с сохраненными данными для веб-страниц википедии.

1. Обход графа выполнялся с помощью `scrapy`, за основу был взят пример из `spider.ipynb`, для удобства класс спайдера реализовывался как наследник от `CrawlSpider`, допустимые и запрещенные адреса для перехода по ссылкам задаются при создании `extractor-a`, игнорируются ссылки с якорем. В качестве сниппета брались как в примере первые 255 символов из первого абзаца в `mw-content-text`.

Для хранения данных страницы служит класс `WikiPage`, где для каждой страницы сохраняются адрес, заголовок, сниппет, и список ссылок на другие страницы. При обходе постоянно обновляется множество адресов просмотренных страниц `visited_urls`, и если адреса очередной страницы там нет, то она просматривается и затем помещается в словарь `pages_dict`, где ключ – адрес страницы, значение – экземпляр класса `WikiPage`. Этот словарь по окончании обхода сохраняется в бинарный файл `wki_pages.bin` как сериализованный объект.

В качестве дополнительных настроек через `custom_settings` было добавлено ограничение на общее количество страниц и заданы параметры для обхода в ширину, а не в глубину, как это делает `CrawlSpider` по умолчанию, чтобы достигнуть примерно одинаковой глубины обхода для стартового набора из 5 страниц.

2. С помощью библиотеки `networkx` строился ориентированный граф путем последовательного добавления дуг, выходящих из вершин-адресов страниц и ведущих к вершинам-ссылкам с этой страницы, на основе загруженного из файла словаря с данными. Ссылки на страницы, которых нет в сохраненном словаре, игнорируются и в граф не попадают, т.к. такая страница не была просмотрена, и будет висячей вершиной в графе, что негативно скажется на выдаваемом результате алгоритма и времени его работы (в этом случае количество вершин в графе будет порядка 10^5 вместо предполагаемой 10^4).

3. Результат подсчета Pagerank с $\alpha=0.85$:

```
United States < 0.00644868271352 >  
https://en.wikipedia.org/wiki/United_States  
Coordinates: 40°N 100°W / 40°N 100°W / 40; -100...
```

```
International Standard Book Number < 0.00628171718528 >  
https://en.wikipedia.org/wiki/International_Standard_Book_Number  
The International Standard Book Number (ISBN) is a unique[a][b] numeric commercial book  
identifier. Publishers purchase ISBNs from an affiliate of the International ISBN Agency.[1]...
```

United Kingdom < 0.00508372878615 >

https://en.wikipedia.org/wiki/United_Kingdom

The United Kingdom of Great Britain and Northern Ireland, commonly known as the United Kingdom (UK) or Britain, is a sovereign country in western Europe. Lying off the north-western coast of the European mainland, the UK includes the island of Great Brita...

Geographic coordinate system < 0.00395316787027 >

https://en.wikipedia.org/wiki/Geographic_coordinate_system

A geographic coordinate system is a coordinate system used in geography that enables every location on Earth to be specified by a set of numbers, letters or symbols.[n 1] The coordinates are often chosen such that one of the numbers represents a vertical ...

Digital object identifier < 0.00386124549609 >

https://en.wikipedia.org/wiki/Digital_object_identifier

In computing, a Digital Object Identifier or DOI is a persistent identifier or handle used to uniquely identify objects, standardized by the International Organization for Standardization (ISO).[1] An implementation of the Handle System,[2][3] DOIs are in...

Daylight saving time < 0.00348884138144 >

https://en.wikipedia.org/wiki/Daylight_saving_time

Daylight saving time (abbreviated DST), sometimes referred to as daylight savings time in US, Canadian and Australian speech,[1][2] and known as British Summer Time (BST) in the UK and just summer time in some countries, is the practice of advancing clock...

Demonym < 0.0034758132573 >

<https://en.wikipedia.org/wiki/Demonym>

A demonym (/ˈdɛmənɪm/; δῆμος dēmos "people, tribe", ὄνομα ónoma "name") is a word that identifies residents or natives of a particular place, which is derived from the name of that particular place.[1]...

Europe < 0.00305316037764 >

<https://en.wikipedia.org/wiki/Europe>

Europe is a continent located entirely in the Northern Hemisphere and mostly in the Eastern Hemisphere. It is bordered by the Arctic Ocean to the north, the Atlantic Ocean to the west, and the Mediterranean Sea to the south. It comprises the westernmost p...

Greek language < 0.00287913339989 >

https://en.wikipedia.org/wiki/Greek_language

Greek (Modern Greek: ελληνικά [eliniˈka], elliniká, "Greek", ελληνική γλώσσα [eliniˈci ˈɣlosa] (listen), ellinikí glóssa, "Greek language") is an independent branch of the Indo-European family of languages, native to Greece and other parts of the Eastern...

English language < 0.00274104953906 >

https://en.wikipedia.org/wiki/English_language

English is a West Germanic language that was first spoken in early medieval England and is now a global lingua franca.[4][5] Named after the Angles, one of the Germanic tribes that migrated to England, it ultimately derives its name from the Anglia (Angel...

4. Результат подсчета Pagerank с разными значениями alpha размещены в Приложении А в конце отчета. Здесь приведена для наглядности только таблица с заголовками страниц из топ-10 для разных alpha.

Alpha=0.85	Alpha=0.95	Alpha=0.5	Alpha=0.3
1. United States	1. United States	1. International Standard Book Number	1. International Standard Book Number
2. International Standard Book Number	2. United Kingdom	2. United States	2. United States
3. United Kingdom	3. International Standard Book Number	3. United Kingdom	3. Digital object identifier
4. Geographic coordinate system	4. Demonym	4. Digital object identifier	4. United Kingdom
5. Digital object identifier	5. Daylight saving time	5. Geographic coordinate system	5. Geographic coordinate system
6. Daylight saving time	6. Geographic coordinate system	6. International Standard Serial Number	6. Virtual International Authority File
7. Demonym	7. Europe	7. Virtual International Authority File	7. England
8. Europe	8. Greek language	8. England	8. London
9. Greek language	9. Unicode	9. London	9. International Standard Serial Number
10. English language	10. Gross domestic product	10. World War II	10. Rock music

Результаты не слишком сильно различаются, в основном меняется порядок следования страниц. Судя по заголовкам, понятно что такие страницы авторитетные, и на них ссылается множество других страниц википедии. Интересно, что даже с сильным изменением alpha с 0.3 до 0.95, когда значительно возросла вероятность случайного перехода по страницам и уменьшилась вероятность перехода по ссылкам, в топ-10 все равно неизменно остались одни и те же 6 страниц.

Стоит отметить, что изменения коэффициента затухания alpha повлияло также на значения pagerank-а каждой страницы. При меньшем значении коэффициента alpha, получалось меньшее значение pagerank-а, что очевидно из формулы расчета pagerank, т.к. слагаемое, отвечающее за вклад в pagerank следующей страницы от текущей, умножается на alpha.

5. Алгоритм HITS, реализованный в библиотеке networkx не сходиллся для построенного графа при параметрах по умолчанию: максимальное количество итераций 100, допустимая ошибка $1.0e-8$. После постепенного увеличения количества итераций до 500 и ошибки до 0.1, алгоритм сошелся.

К сожалению, ожидания, что топ-10 авторитетности алгоритма HITS и PageRank будут хотя бы похожи, если не совпадут, не оправдались. Более правдивыми кажутся результаты PageRank, так как на страницы из его топ-10 потенциально могут ссылаться множество страниц, независимо от тематики. В то время как на страницы из списка авторитетности HITS скорее будут ссылаться статьи из квантовой физики, возможно и многие, но все же это некая локальная авторитетность, и не понятно почему алгоритм HITS выдал в качестве результатов именно эти страницы.

Ради интереса была вычислена средняя степень входа и выхода для вершин, соответствующих страницам из топ-10 для каждого алгоритма:

	PageRank	HITS, Авторитетность	HITS, Хабовость
Степень входа:	611	233	50
Степень выхода:	66	66	98

Для хабовости эти данные в некотором смысле могут быть правдивыми, т.к. степень выхода в 2 раза больше степени входа, но данные для авторитетности HITS все же далеки от ожидаемых по сравнению с PageRank. Возможно, алгоритм HITS показывает хорошие результаты на более разреженных графах, а для данного случая ($|V|=10006$, $|E|=259873$) обладает плохой сходимостью и поэтому выдает скорее всего ошибочные результаты.

Результаты работы алгоритма HITS:

Хабовость:

Hubs results:

Quantum chaos < 0.00370242539535 >

https://en.wikipedia.org/wiki/Quantum_chaos

Quantum chaos is a branch of physics which studies how chaotic classical dynamical systems can be described in terms of quantum theory. The primary question that quantum chaos seeks to answer is: "What is the relationship between quantum mechanics and cla...

EPR paradox < 0.00370242539535 >

https://en.wikipedia.org/wiki/EPR_paradox

The Einstein–Podolsky–Rosen paradox or the EPR paradox[1] of 1935 is a thought experiment in quantum mechanics with which Albert Einstein and his colleagues Boris Podolsky and Nathan Rosen (EPR) claimed to demonstrate that the wave function does not provi...

Atomic, molecular, and optical physics < 0.00370242539535 >

https://en.wikipedia.org/wiki/Atomic,_molecular,_and_optical_physics

Atomic, molecular, and optical physics (AMO) is the study of matter-matter and light-matter interactions; at the scale of one or a few atoms [1] and energy scales around several electron volts. [2]:1356[3] The three areas are closely interrelated. AMO theo...

EPR paradox < 0.00370242539535 >

https://en.wikipedia.org/wiki/Einstein%E2%80%93Podolsky%E2%80%93Rosen_paradox

The Einstein–Podolsky–Rosen paradox or the EPR paradox[1] of 1935 is a thought experiment in quantum mechanics with which Albert Einstein and his colleagues Boris Podolsky and Nathan Rosen (EPR) claimed to demonstrate that the wave function does not provi...

Quantum information science < 0.00370242539535 >

https://en.wikipedia.org/wiki/Quantum_information_science

Quantum information science is an area of study based on the idea that information science depends on quantum effects in physics. It includes theoretical issues in computational models as well as more experimental topics in quantum physics including what ...

Relativistic quantum mechanics < 0.0036942619618 >

https://en.wikipedia.org/wiki/Relativistic_quantum_mechanics

In physics, relativistic quantum mechanics (RQM) is any Poincaré covariant formulation of quantum mechanics (QM). This theory is applicable to massive particles propagating at all velocities up to those comparable to the speed of light c , and can accommod...

Fractional quantum mechanics < 0.0036911738838 >

https://en.wikipedia.org/wiki/Fractional_quantum_mechanics

In physics, fractional quantum mechanics is a generalization of standard quantum mechanics, which naturally comes out when the Brownian-like quantum paths substitute with the Lévy-like ones in the Feynman path integral. It has been discovered by Nick Lask...

Transactional interpretation < 0.00369065666309 >

https://en.wikipedia.org/wiki/Transactional_interpretation

The transactional interpretation of quantum mechanics (TIQM) takes the ψ and ψ^* wave functions of the standard quantum formalism to be retarded (forward in time) and advanced (backward in time) waves that form a quantum interaction as a Wheeler–Feynma...

Stochastic quantum mechanics < 0.00368982538328 >

https://en.wikipedia.org/wiki/Stochastic_quantum_mechanics

Stochastic quantum mechanics (or the stochastic interpretation) is an interpretation of quantum mechanics....

Objective collapse theory < 0.00368978818559 >

https://en.wikipedia.org/wiki/Objective_collapse_theory

Objective collapse theories, also known as quantum mechanical spontaneous localization models (QMSL), are an approach to the interpretational problems of quantum mechanics. They are realistic and indeterministic and reject hidden variables. The approach i...

Авторитетность:

Authorities results:

Classical mechanics < 0.00634635253229 >

https://en.wikipedia.org/wiki/Classical_mechanics

Classical mechanics describes the motion of macroscopic objects, from projectiles to parts of machinery, and astronomical objects, such as spacecraft, planets, stars and galaxies....

Energy < 0.00506277449584 >

<https://en.wikipedia.org/wiki/Energy>

In physics, energy is the quantitative property that must be transferred to an object in order to perform work on, or to heat, the object.[note 1] Energy is a conserved quantity; the law of conservation of energy states that energy can be converted in for...

Quantum mechanics < 0.00495410428075 >

https://en.wikipedia.org/wiki/Quantum_mechanics

Quantum mechanics (QM; also known as quantum physics or quantum theory), including quantum field theory, is a fundamental theory in physics which describes nature at the smallest scales of energy levels of atoms and subatomic particles.[2]...

Isaac Newton < 0.0044317941402 >

https://en.wikipedia.org/wiki/Isaac_Newton

Sir Isaac Newton PRS (/ˈnjuːtən/;[6] 25 December 1642 – 20 March 1726/27[1]) was an English

mathematician, astronomer, theologian, author and physicist (described in his own day as a "natural philosopher") who is widely recognised as one of the most influ...

Galileo Galilei < 0.00442961892101 >

https://en.wikipedia.org/wiki/Galileo_Galilei

Galileo Galilei (Italian: [ɡaliˈlɛːo ɡaliˈlɛi]; 15 February 1564[3] – 8 January 1642) was an Italian polymath. Galileo is a central figure in the transition from natural philosophy to modern science and in the transformation of the scientific Renaissance ...

Spin (physics) < 0.00415477815261 >

[https://en.wikipedia.org/wiki/Spin_\(physics\)](https://en.wikipedia.org/wiki/Spin_(physics))

In quantum mechanics and particle physics, spin is an intrinsic form of angular momentum carried by elementary particles, composite particles (hadrons), and atomic nuclei.[1][2]...

Albert Einstein < 0.00414257458214 >

https://en.wikipedia.org/wiki/Albert_Einstein

Albert Einstein (14 March 1879 – 18 April 1955) was a German-born theoretical physicist[5] who developed the theory of relativity, one of the two pillars of modern physics (alongside quantum mechanics).[4][6]:274 His work is also known for its influence o...

Spontaneous symmetry breaking < 0.00413263799591 >

https://en.wikipedia.org/wiki/Spontaneous_symmetry_breaking

Spontaneous symmetry breaking is a spontaneous process of symmetry breaking, by which a physical system in a symmetric state ends up in an asymmetric state.[1][2][3] In particular, it can describe systems where the equations of motion or the Lagrangian ob...

Schrödinger equation < 0.00411766586431 >

https://en.wikipedia.org/wiki/Schr%C3%B6dinger_equation

In quantum mechanics, the Schrödinger equation is a mathematical equation that describes the changes over time of a physical system in which quantum effects, such as wave–particle duality, are significant. The equation is a mathematical formulation for st...

Vacuum state < 0.00411650574395 >

https://en.wikipedia.org/wiki/Vacuum_state

In quantum field theory, the quantum vacuum state (also called the quantum vacuum or vacuum state) is the quantum state with the lowest possible energy. Generally, it contains no physical particles. Zero-point field is sometimes used as a synonym for the ...

Среднее хабовости и авторитетности:

Mean hits results:

Quantum mechanics < 0.00427665196592 >

https://en.wikipedia.org/wiki/Quantum_mechanics

Quantum mechanics (QM; also known as quantum physics or quantum theory), including quantum field theory, is a fundamental theory in physics which describes nature at the smallest scales of energy levels of atoms and subatomic particles.[2]...

Classical mechanics < 0.00409983256894 >

https://en.wikipedia.org/wiki/Classical_mechanics

Classical mechanics describes the motion of macroscopic objects, from projectiles to parts of

machinery, and astronomical objects, such as spacecraft, planets, stars and galaxies....

Schrödinger equation < 0.00390203766551 >

https://en.wikipedia.org/wiki/Schr%C3%B6dinger_equation

In quantum mechanics, the Schrödinger equation is a mathematical equation that describes the changes over time of a physical system in which quantum effects, such as wave–particle duality, are significant. The equation is a mathematical formulation for st...

Quantum entanglement < 0.00385189808878 >

https://en.wikipedia.org/wiki/Quantum_entanglement

Quantum entanglement is a physical phenomenon which occurs when pairs or groups of particles are generated or interact in ways such that the quantum state of each particle cannot be described independently of the state of the other(s), even when the parti...

Introduction to quantum mechanics < 0.0038512755025 >

https://en.wikipedia.org/wiki/Introduction_to_quantum_mechanics

Quantum mechanics is the science of the very small. It explains the behavior of matter and its interactions with energy on the scale of atoms and subatomic particles....

Symmetry in quantum mechanics < 0.00384958179571 >

https://en.wikipedia.org/wiki/Symmetry_in_quantum_mechanics

Symmetries in quantum mechanics describe features of spacetime and particles which are unchanged under some transformation, in the context of quantum mechanics, relativistic quantum mechanics and quantum field theory, and with applications in the mathemat...

Quantum decoherence < 0.0038448121285 >

https://en.wikipedia.org/wiki/Quantum_decoherence

Quantum decoherence is the loss of quantum coherence. In quantum mechanics, particles such as electrons are described by a wavefunction, a mathematical description of the quantum state of a system; the probabilistic nature of the wavefunction gives rise t...

Old quantum theory < 0.00384388990513 >

https://en.wikipedia.org/wiki/Old_quantum_theory

The old quantum theory is a collection of results from the years 1900–1925[1] which predate modern quantum mechanics. The theory was never complete or self-consistent, but was rather a set of heuristic corrections to classical mechanics.[2] The theory is ...

Quantum state < 0.00384339004381 >

https://en.wikipedia.org/wiki/Quantum_state

In quantum physics, quantum state refers to the state of an isolated quantum system. A quantum state provides a probability distribution for the value of each observable, i.e. for the outcome of each possible measurement on the system. Knowledge of the qu...

Bra–ket notation < 0.00384132768138 >

https://en.wikipedia.org/wiki/Bra%E2%80%93ket_notation

In quantum mechanics, bra–ket notation is a standard notation for describing quantum states. It can also be used to denote abstract vectors and linear functionals in mathematics. The notation begins with using angle brackets, \langle and \rangle , and a vertical bar, ...

ПРИЛОЖЕНИЕ А

PageRank results, alpha=0.85:

United States < 0.00644868271352 >

https://en.wikipedia.org/wiki/United_States

Coordinates: 40°N 100°W / 40°N 100°W / 40; -100...

International Standard Book Number < 0.00628171718528 >

https://en.wikipedia.org/wiki/International_Standard_Book_Number

The International Standard Book Number (ISBN) is a unique[a][b] numeric commercial book identifier. Publishers purchase ISBNs from an affiliate of the International ISBN Agency.[1]...

United Kingdom < 0.00508372878615 >

https://en.wikipedia.org/wiki/United_Kingdom

The United Kingdom of Great Britain and Northern Ireland, commonly known as the United Kingdom (UK) or Britain, is a sovereign country in western Europe. Lying off the north-western coast of the European mainland, the UK includes the island of Great Brita...

Geographic coordinate system < 0.00395316787027 >

https://en.wikipedia.org/wiki/Geographic_coordinate_system

A geographic coordinate system is a coordinate system used in geography that enables every location on Earth to be specified by a set of numbers, letters or symbols.[n 1] The coordinates are often chosen such that one of the numbers represents a vertical ...

Digital object identifier < 0.00386124549609 >

https://en.wikipedia.org/wiki/Digital_object_identifier

In computing, a Digital Object Identifier or DOI is a persistent identifier or handle used to uniquely identify objects, standardized by the International Organization for Standardization (ISO).[1] An implementation of the Handle System,[2][3] DOIs are in...

Daylight saving time < 0.00348884138144 >

https://en.wikipedia.org/wiki/Daylight_saving_time

Daylight saving time (abbreviated DST), sometimes referred to as daylight savings time in US, Canadian and Australian speech,[1][2] and known as British Summer Time (BST) in the UK and just summer time in some countries, is the practice of advancing clock...

Demonym < 0.0034758132573 >

<https://en.wikipedia.org/wiki/Demonym>

A demonym (/ˈdɛmənɪm/; δῆμος dēmos "people, tribe", ὄνομα ónoma "name") is a word that identifies residents or natives of a particular place, which is derived from the name of that particular place.[1]...

Europe < 0.00305316037764 >

<https://en.wikipedia.org/wiki/Europe>

Europe is a continent located entirely in the Northern Hemisphere and mostly in the Eastern Hemisphere. It is bordered by the Arctic Ocean to the north, the Atlantic Ocean to the west, and the Mediterranean Sea to the south. It comprises the westernmost p...

Greek language < 0.00287913339989 >

https://en.wikipedia.org/wiki/Greek_language

Greek (Modern Greek: ελληνικά [eliniˈka], elliniká, "Greek", ελληνική γλώσσα [eliniˈci

'ɣlosa] (listen), ellinikí glóssa, "Greek language") is an independent branch of the Indo-European family of languages, native to Greece and other parts of the Eastern...

English language < 0.00274104953906 >

https://en.wikipedia.org/wiki/English_language

English is a West Germanic language that was first spoken in early medieval England and is now a global lingua franca.[4][5] Named after the Angles, one of the Germanic tribes that migrated to England, it ultimately derives its name from the Anglia (Angel...

PageRank results, alpha=0.95:

United States < 0.0070585692626 >

https://en.wikipedia.org/wiki/United_States

Coordinates: 40°N 100°W / 40°N 100°W / 40; -100...

United Kingdom < 0.00566321659744 >

https://en.wikipedia.org/wiki/United_Kingdom

The United Kingdom of Great Britain and Northern Ireland, commonly known as the United Kingdom (UK) or Britain, is a sovereign country in western Europe. Lying off the north-western coast of the European mainland, the UK includes the island of Great Brita...

International Standard Book Number < 0.00559048449347 >

https://en.wikipedia.org/wiki/International_Standard_Book_Number

The International Standard Book Number (ISBN) is a unique[a][b] numeric commercial book identifier. Publishers purchase ISBNs from an affiliate of the International ISBN Agency.[1]...

Demonym < 0.00457232428843 >

<https://en.wikipedia.org/wiki/Demonym>

A demonym (/ˈdɛmənɪm/; δῆμος dēmos "people, tribe", ὄνομα ónoma "name") is a word that identifies residents or natives of a particular place, which is derived from the name of that particular place.[1]...

Daylight saving time < 0.00447955542535 >

https://en.wikipedia.org/wiki/Daylight_saving_time

Daylight saving time (abbreviated DST), sometimes referred to as daylight savings time in US, Canadian and Australian speech,[1][2] and known as British Summer Time (BST) in the UK and just summer time in some countries, is the practice of advancing clock...

Geographic coordinate system < 0.00428932063581 >

https://en.wikipedia.org/wiki/Geographic_coordinate_system

A geographic coordinate system is a coordinate system used in geography that enables every location on Earth to be specified by a set of numbers, letters or symbols.[n 1] The coordinates are often chosen such that one of the numbers represents a vertical ...

Europe < 0.0042170139037 >

<https://en.wikipedia.org/wiki/Europe>

Europe is a continent located entirely in the Northern Hemisphere and mostly in the Eastern Hemisphere. It is bordered by the Arctic Ocean to the north, the Atlantic Ocean to the west, and the Mediterranean Sea to the south. It comprises the westernmost p...

Greek language < 0.00398926704731 >

https://en.wikipedia.org/wiki/Greek_language

Greek (Modern Greek: ελληνικά [elini'ka], elliniká, "Greek", ελληνική γλώσσα [elini'ci 'ɣlosa] (listen), ellinikí glóssa, "Greek language") is an independent branch of the Indo-European family of languages, native to Greece and other parts of the Eastern...

Unicode < 0.00376002971689 >

<https://en.wikipedia.org/wiki/Unicode>

Unicode is a computing industry standard for the consistent encoding, representation, and handling of text expressed in most of the world's writing systems. The latest version contains a repertoire of 136,755 characters covering 139 modern and historic sc...

Gross domestic product < 0.0037187949152 >

https://en.wikipedia.org/wiki/Gross_domestic_product

Gross domestic product (GDP) is a monetary measure of the market value of all final goods and services produced in a period (quarterly or yearly) of time. Nominal GDP estimates are commonly used to determine the economic performance of a whole country or ...

PageRank results, alpha=0.5:

International Standard Book Number < 0.00519926329786 >

https://en.wikipedia.org/wiki/International_Standard_Book_Number

The International Standard Book Number (ISBN) is a unique[a][b] numeric commercial book identifier. Publishers purchase ISBNs from an affiliate of the International ISBN Agency.[1]...

United States < 0.00346570028408 >

https://en.wikipedia.org/wiki/United_States

Coordinates: 40°N 100°W / 40°N 100°W / 40; -100...

United Kingdom < 0.00282418503819 >

https://en.wikipedia.org/wiki/United_Kingdom

The United Kingdom of Great Britain and Northern Ireland, commonly known as the United Kingdom (UK) or Britain, is a sovereign country in western Europe. Lying off the north-western coast of the European mainland, the UK includes the island of Great Brita...

Digital object identifier < 0.00281340743128 >

https://en.wikipedia.org/wiki/Digital_object_identifier

In computing, a Digital Object Identifier or DOI is a persistent identifier or handle used to uniquely identify objects, standardized by the International Organization for Standardization (ISO).[1] An implementation of the Handle System,[2][3] DOIs are in...

Geographic coordinate system < 0.00259962187457 >

https://en.wikipedia.org/wiki/Geographic_coordinate_system

A geographic coordinate system is a coordinate system used in geography that enables every location on Earth to be specified by a set of numbers, letters or symbols.[n 1] The coordinates are often chosen such that one of the numbers represents a vertical ...

International Standard Serial Number < 0.00160713325491 >

https://en.wikipedia.org/wiki/International_Standard_Serial_Number

An International Standard Serial Number (ISSN) is an eight-digit serial number used to uniquely

identify a serial publication.[1] The ISSN is especially helpful in distinguishing between serials with the same title. ISSN are used in ordering, cataloging, ...

Virtual International Authority File < 0.0015644044491 >

https://en.wikipedia.org/wiki/Virtual_International_Authority_File

The Virtual International Authority File (VIAF) is an international authority file. It is a joint project of several national libraries and operated by the Online Computer Library Center (OCLC). [1]...

England < 0.00149138562686 >

<https://en.wikipedia.org/wiki/England>

England is a country that is part of the United Kingdom.[6][7][8] It shares land borders with Scotland to the north and Wales to the west. The Irish Sea lies northwest of England and the Celtic Sea lies to the southwest. England is separated from continen...

London < 0.00146401868738 >

<https://en.wikipedia.org/wiki/London>

London (/ˈlʌndən/ (listen)) is the capital and most populous city of England and the United Kingdom.[7][8] Standing on the River Thames in the south east of the island of Great Britain, London has been a major settlement for two millennia. It was founded...

World War II < 0.00145200270502 >

https://en.wikipedia.org/wiki/World_War_II

World War II (often abbreviated to WWII or WW2), also known as the Second World War, was a global war that lasted from 1939 to 1945, although related conflicts began earlier. The vast majority of the world's countries—including all of the great powers—eve...

PageRank results, alpha=0.3:

International Standard Book Number < 0.00349851182446 >

https://en.wikipedia.org/wiki/International_Standard_Book_Number

The International Standard Book Number (ISBN) is a unique[a][b] numeric commercial book identifier. Publishers purchase ISBNs from an affiliate of the International ISBN Agency.[1]...

United States < 0.00202676066969 >

https://en.wikipedia.org/wiki/United_States

Coordinates: 40°N 100°W / 40°N 100°W / 40; -100...

Digital object identifier < 0.00173697250877 >

https://en.wikipedia.org/wiki/Digital_object_identifier

In computing, a Digital Object Identifier or DOI is a persistent identifier or handle used to uniquely identify objects, standardized by the International Organization for Standardization (ISO).[1] An implementation of the Handle System,[2][3] DOIs are in...

United Kingdom < 0.00169132359152 >

https://en.wikipedia.org/wiki/United_Kingdom

The United Kingdom of Great Britain and Northern Ireland, commonly known as the United Kingdom (UK) or Britain, is a sovereign country in western Europe. Lying off the north-western coast of the European mainland, the UK includes the island of Great Brita...

Geographic coordinate system < 0.00167002772983 >

https://en.wikipedia.org/wiki/Geographic_coordinate_system

A geographic coordinate system is a coordinate system used in geography that enables every location on Earth to be specified by a set of numbers, letters or symbols.[n 1] The coordinates are often chosen such that one of the numbers represents a vertical ...

Virtual International Authority File < 0.0011202009946 >

https://en.wikipedia.org/wiki/Virtual_International_Authority_File

The Virtual International Authority File (VIAF) is an international authority file. It is a joint project of several national libraries and operated by the Online Computer Library Center (OCLC). [1]...

England < 0.000987222297629 >

<https://en.wikipedia.org/wiki/England>

England is a country that is part of the United Kingdom.[6][7][8] It shares land borders with Scotland to the north and Wales to the west. The Irish Sea lies northwest of England and the Celtic Sea lies to the southwest. England is separated from continen...

London < 0.000941075323188 >

<https://en.wikipedia.org/wiki/London>

London (/ˈlʌndən/ (listen)) is the capital and most populous city of England and the United Kingdom.[7][8] Standing on the River Thames in the south east of the island of Great Britain, London has been a major settlement for two millennia. It was founded...

International Standard Serial Number < 0.000921733679076 >

https://en.wikipedia.org/wiki/International_Standard_Serial_Number

An International Standard Serial Number (ISSN) is an eight-digit serial number used to uniquely identify a serial publication.[1] The ISSN is especially helpful in distinguishing between serials with the same title. ISSN are used in ordering, cataloging, ...

Rock music < 0.000915759119176 >

https://en.wikipedia.org/wiki/Rock_music

Rock music is a broad genre of popular music that originated as "rock and roll" in the United States in the early 1950s, and developed into a range of different styles in the 1960s and later, particularly in the United Kingdom and in the United States.[1]...