

GALE PRIMARY SOURCES

Eighteenth Century Collections Online

《18世紀英語・英国出版物集成 ECCO》

データベース講習会



2022年2月22日

センゲージラーニング株式会社
Gale部門 森澤 正樹

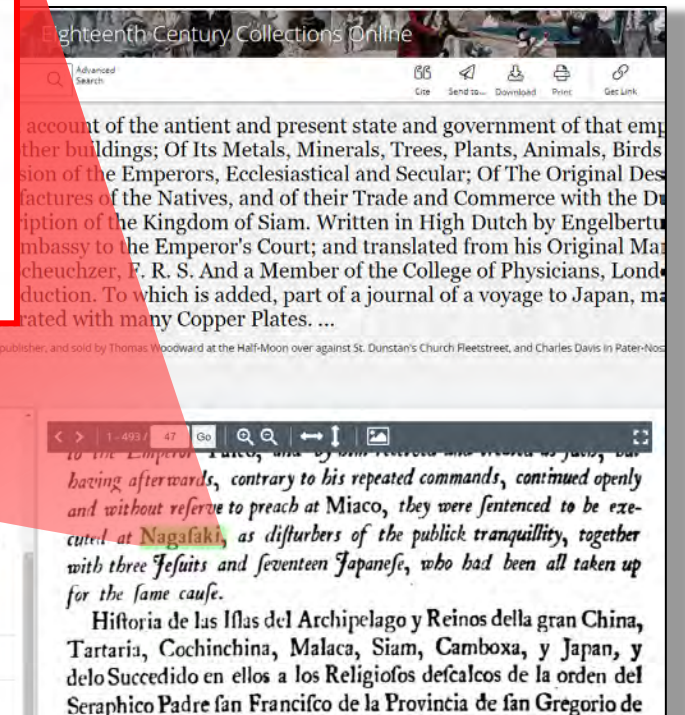
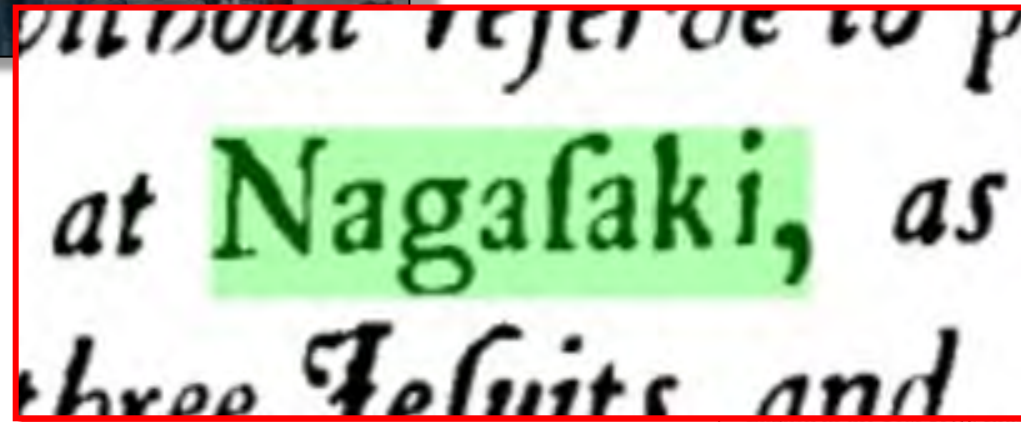
会社紹介

GALE

- 米国の教育出版社センゲージ ラーニングの図書館部門
- 本部: 米国ミシガン州ファーミントン・ヒルズ市
- 出版形態
 - ✓ 百科事典・年鑑などのレファレンス書籍
 - ✓ 電子ブック・文学情報・雑誌記事などのデータベース
 - ✓ 一次資料・新聞などのマイクロフォーム
 - ✓ 一次資料・新聞などのデジタル・アーカイブ **Gale Primary Sources**



OCR（光学文字認識）によるフルテキスト検索



OCR = Optical Character Recognition

本日のアジェンダ

- 18世紀英語・英国出版物集成 (ECCO)
Eighteenth Century Collections Online について
- 3つのプラットフォームについて
(Gale Primary Sources、NII-REO、Gale Digital Scholar Lab)
- デモンストレーション
(Gale Primary Sources プラットフォーム)
- Gale Digital Scholar Lab について
- 関連資料のご案内

The screenshot shows the Gale Primary Sources website interface. At the top, there is a navigation bar with the Gale logo and options for language (English) and sign-in (Google, Microsoft). The main header features the text "GALE PRIMARY SOURCES" and "Eighteenth Century Collections Online" with a search bar. Below the header, there is a navigation menu with icons for Authors, About, Research Tools, Search History, and Get Link. The main content area is titled "ABOUT THIS RESOURCE" and contains two paragraphs of text describing the ECCO database. The first paragraph states that ECCO is a vast eighteenth-century library with over 180,000 titles and 32 million pages. The second paragraph explains that ECCO is a digitization of the ESTC project, which records works published in Britain, Ireland, and the United States. Below the text, there is a "SEARCH TOOLS" section with two options: "Topic Finder" and "Term Frequency". The "Topic Finder" option includes a pie chart icon and a description: "Visualize connections between search terms and topics and view relevant articles for those topics." The "Term Frequency" option includes a line graph icon and a description: "Explore the coverage of your search terms, graphed over time and note trends in coverage of multiple topics." At the bottom of the page, there is a footer with the Gale logo and navigation links: ABOUT | HELP | CONTACT US | TERMS OF USE | PRIVACY POLICY | ACCESSIBILITY | END SESSION. The tagline "Gale, here for everyone." is also present.

18世紀英語・英国出版物集成 (ECCO) Eighteenth Century Collections Online

18世紀の英語・英語圏刊行物の最も包括的なデータベース

- 18世紀刊行の英語・英語圏刊行物が収録対象
- 約18.5万タイトル(20.7万冊)を収録、総ページ数 約3,300万
- 原本は大英図書館他、世界各国の所蔵機関
- マイクロフィルム版「The Eighteenth Century」からスキャニング
- 優れたメタデータ: LC主題分類・目次・図版一覧など
- OCR(光学文字認識)によるフルテキスト検索



18世紀英語・英国出版物集成 (ECCO) Eighteenth Century Collections Online

収録基準

- 初期英語刊本所蔵目録『[The English Short Title Catalogue 1473-1800 \(ESTC\)](#)』記載の18世紀文献が対象
- 年代: 1701年から1800年
- イギリスおよびイギリス植民地の刊行物(全言語)
- 他の地域で刊行された英語刊行物
- 重要な著者・著作は複数の版を収録
- ただし、一枚刷り凹版印刷物、小型の消耗印刷物、新聞等は除く



18世紀英語・英国出版物集成 (ECCO) Eighteenth Century Collections Online

いわゆる「長いs」について

- 18世紀には小文字の「s」において、「長いs (long s)」と呼ばれる活字がよく用いられた
- 主に語頭・語中の小文字で用いられ、語尾および大文字の「s」は現代と同形
- 18世紀後半から19世紀初めにかけて次第に用いられなくなった
- ECCOのOCRでは可能な限り正確にとらえるようにしている

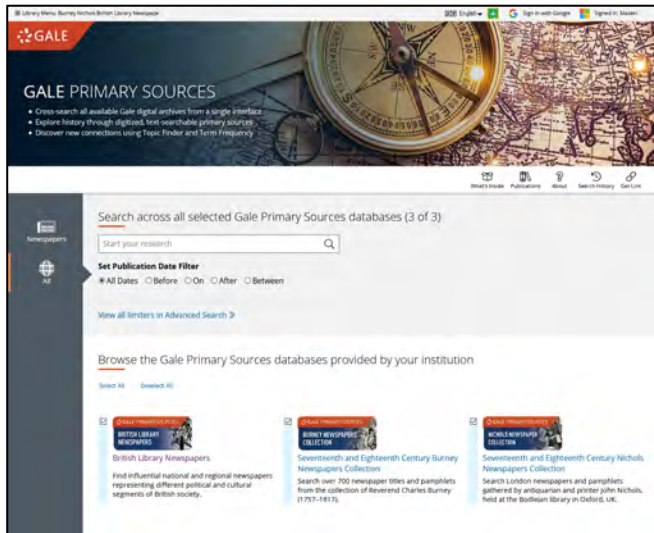


sinfulss



Nagasaki

3つのプラットフォーム



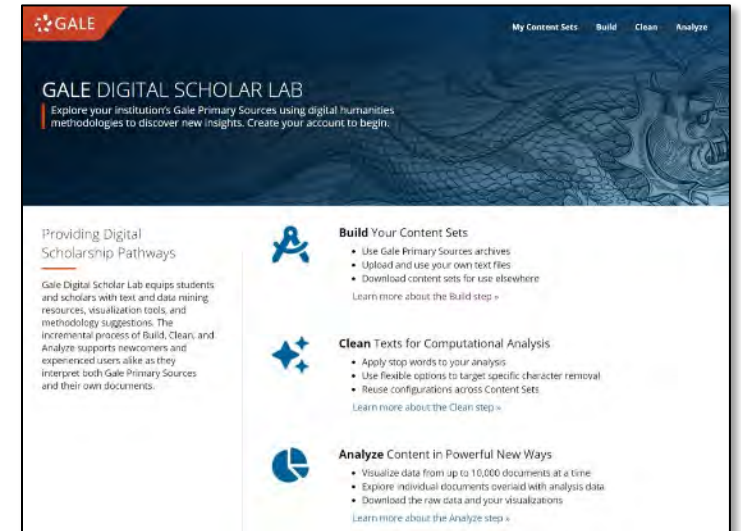
Gale Primary Sources

横断検索・閲読用



NII-REO

JUSTICE会員館のみ

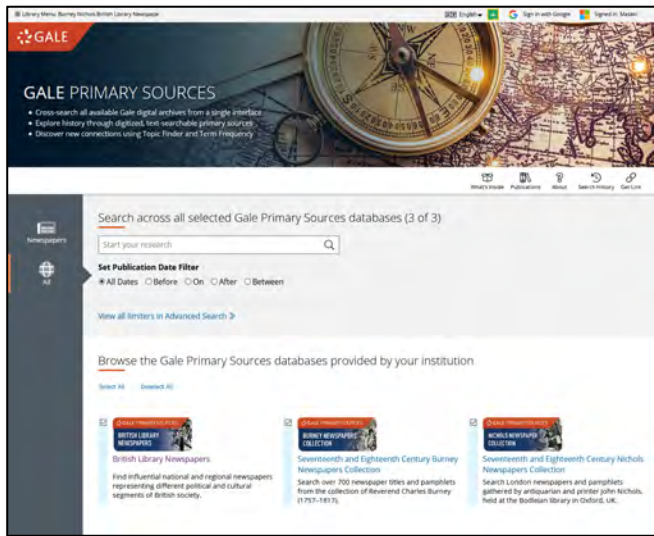


Gale Digital Scholar Lab

テキストマイニング用(別売)

デモンストレーション

3つのプラットフォーム



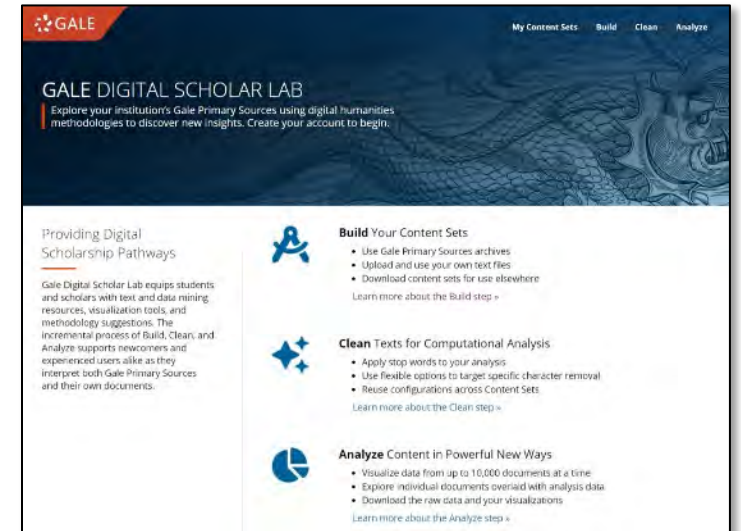
Gale Primary Sources

横断検索・閲読用



NII-REO

JUSTICE会員館のみ



Gale Digital Scholar Lab

テキストマイニング用(別売)

Gale Digital Scholar Lab による分析の例 (クラスタリング分析)

GALE DIGITAL SCHOLAR LAB

My Content Sets Build Clean Analyze

Document Clustering
Women in ECCO

Back Download Help

Legend

Document Clusters

- Cluster-1
242 Documents
- Cluster-2
220 Documents
- Cluster-3
189 Documents
- Cluster-4
57 Documents
- Cluster-5
211 Documents
- Cluster-6
211 Documents

Inspect

Document

Advice to the fair: an Epistolary essay, in three parts: on dress, converse, and marriage: address'd to a sister

Open Document

Top 20 Most Similar Documents

The smaller the distance, the greater the similarity to the selected document. Download Full List

<input type="checkbox"/>	Distance	<input type="checkbox"/>	Title
<input type="checkbox"/>	0.0067	▲	The young woman's monitor; shewing the great happiness of early piety, and the...
<input type="checkbox"/>	0.0089	▲	The young woman's monitor; shewing the great happiness of early piety. And the...
<input type="checkbox"/>	0.0096	●	Advice from farmer truman

Gale Digital Scholar Lab による分析の例（トピックモデリング）

The screenshot shows the Gale Digital Scholar Lab interface for a topic modeling analysis of 'Women in ECCO'. The interface includes a navigation bar with 'Back', 'Tool Setup', and 'Run History' options. The main content area displays the 'Average Word Length' for various topics, with a horizontal bar chart showing the length of the top words for each topic. The x-axis represents the average word length in characters, ranging from 0 to 6. The y-axis lists topics, with Topic 0 highlighted in red, Topic 1 in pink, and Topic 10 in light blue. The text for each topic is displayed below the bar, with the top words highlighted in the same color as the bar. For example, Topic 0 has the top words 'her him little thing did way think lady fine great'. The right sidebar shows a table of statistics for the analysis, including Document Count (100), Document Entropy (4.925), Exclusivity (0.195), Tokens (618283), and Uniform Distance (5.3769). Below the statistics is the 'Topic Composition' section, which is currently set to 'Terms' and shows a list of terms with their counts, probabilities, and document frequencies.

GALE DIGITAL SCHOLAR LAB My Content Sets Build Clean Analyze

Back

Topic Modeling
Women in ECCO

Download Help

Topics

- Topic 0
- Topic 1
- Topic 2
- Topic 3
- Topic 4
- Topic 5
- Topic 6
- Topic 7
- Topic 8
- Topic 9
- Topic 10
- Topic 11
- Topic 12
- Topic 13
- Topic 14
- Topic 15
- Topic 16
- Topic 17
- Topic 18

Average Word Length

The average length, in characters, of the top words. MALLET suggests that longer words often have more specific meaning and therefore topics with a higher average word length may be more specific.

0 1 2 3 4 5 6

Topic 0
her him little thing did way think lady fine great

Topic 1
said Mrs Lady great him dear good you God king

Topic 10
che Rome Emperor did Mamma hath non Death him Prin

Topic 11
the and that for have not with this all big

Statistics

Document Count	100
Document Entropy	4.925
Exclusivity	0.195
Tokens	618283
Uniform Distance	5.3769

Topic Composition

Terms Phrases Documents (457)

Terms	Count	Probability	Docs
thy	4463	0.0072	279
love	2966	0.0048	311
fair	2391	0.0039	286
heart	2298	0.0037	285
like	1997	0.0032	293
eyes	1863	0.003	245
thou	1711	0.0028	218
foul	1579	0.0026	254
o'er	1540	0.0025	179

18世紀英語・英国出版物集成(ECCO) Eighteenth Century Collections Online

研究トピックの例

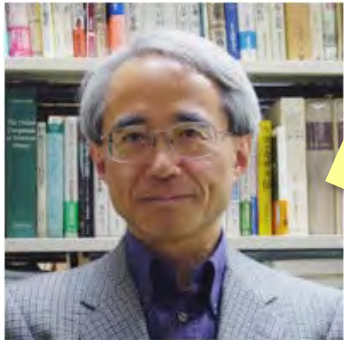
- 18世紀 イギリス史
- 18世紀 ヨーロッパ各国史
- 18世紀 イギリス文学史
- 近代西洋思想史・宗教史
- 大英帝国史・植民地史
- 科学史・産業革命史
- 資本主義の歴史
- 近代芸術史
- 近代英語史
- 書物史・印刷史
- など、など

関連するGALE商品



先生方へのユーザー・インタビュー集もぜひご覧ください！

インタビュー集：www.gale.com/jp/interviews



近藤和彦先生
(東京大学)

(...)そういう点で学問がデモクラティックになって来ているのです。(...)研究者の側も、偉い先生、ミルトンならミルトンについて何十年も研究していた先生もそうでない人も、ミルトンについていっぱしのことが言えるようになってしまう。それは権威者にとっては悲しいことですけどね。でも、学問全体のためには、そのほうが良いのではないかという気がします。(...)逆に、次の世代の研究者たちも、またその上の世代を乗り越えてキャンオンを壊していくようなことになっていくのかも知れないですね。

ECCOが出てきてからは、ECCOで調べても見つからないのであれば、それ以上調べても無理かも知れないという感覚が出てきたように思われます。逆に言えば、昔であれば誰も目にすることがなかったようなマイナーな文献でも、ECCOに収録されていれば、誰かが読んでいるかも知れないし、押さえて置く必要がある、そうしないと学会発表をするとき、指摘されるかも知れない、という感じになりつつあるような気がします。



南井正廣先生
(同志社大学)

ご清聴ありがとうございました。

無料トライアルのお申込み・お問い合わせは下記まで

GaleJapan@cengage.com

www.gale.com/jp



twitter.com/gale_japan