



一次史料テキストマイニング・プラットフォーム  
Gale Digital Scholar Lab

Galeウェビナー

2021年4月28日

センゲージラーニング株式会社  
Gale部門 商品担当

森澤 正樹

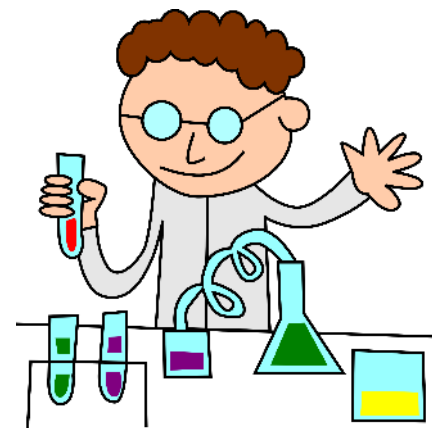
[masaki.morisawa@cengage.com](mailto:masaki.morisawa@cengage.com)

DIGITAL SCHOLAR LAB

Explore your institution's holdings from Gale Primary Sources with new eyes,  
using digital humanities to discover new insights on historical materials

# 本日のアジェンダ

- Gale Primary Sources について
- デジタル人文学: 3つの障害
- Gale Digital Scholar Lab とは
- Gale Digital Scholar Lab デモンストレーション
- インターフェースの更新予定について



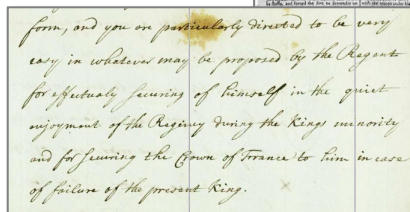
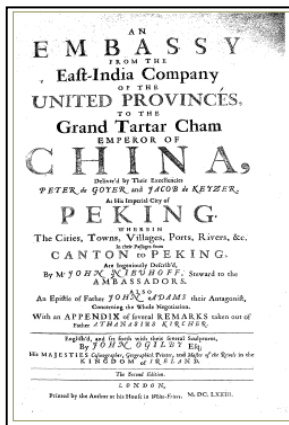
# 会社紹介



- 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

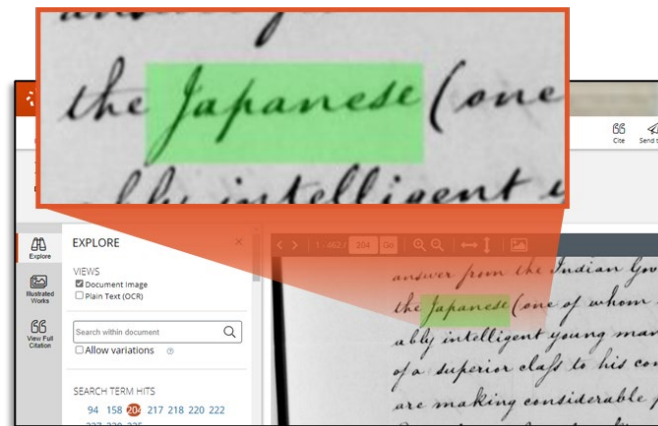
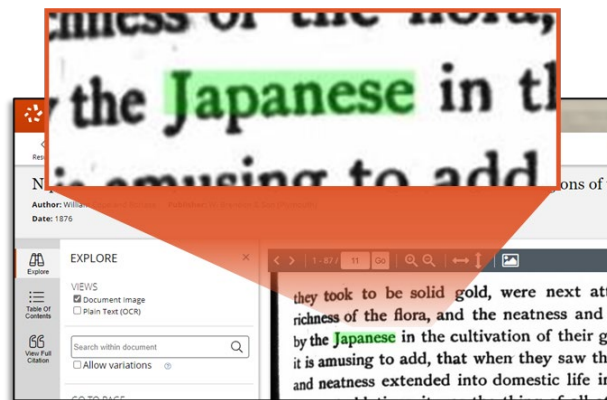
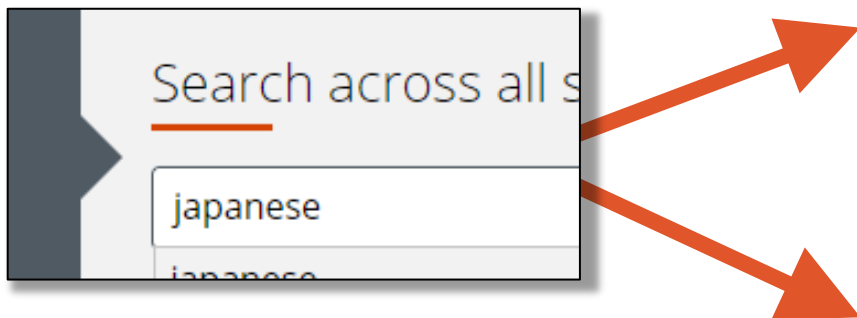


# Gale Primary Sources



- Gale のデジタル一次史料データベース群の総称
- 同データベース群の横断検索プラットフォームの名称
- データベース数400種以上、総ページ数2億ページ以上
- 主な商品の例:
  - Eighteenth Century Collections Online (ECCO)
  - Nineteenth Century Collections Online (NCCO)
  - The Times Digital Archive
  - The Making of the Modern World (MOMW)
  - U.S. Declassified Documents Online
  - ほか、多数

# 歴史資料のフルテキスト検索



OCR = 光学文字認識 (Optical Character Recognition)  
HTR = 手書き文字認識 (Handwritten Text Recognition)

\* HTRIは一部の最新データベースのみ



# Gale Primary Sources 商品群



# OCRテキスト

## Kakemono: Japanese Sketches: by A. Herbage Edwards

**Author:** A. Herbage Edwards **Publisher:** A. C. McClurg & Co.; William Heinemann (Chicago; London)

**Date:** 1906



Explore



Table Of Contents



Illustrated Works



View Full Citation

### EXPLORE



#### VIEWS

- Document Image
- Plain Text (OCR)

Search within document



Allow variations

#### GO TO PAGE

21

Go

#### > Search Term Hits (click for more)

2 4 8 10 14 29 33 35  
42 50 51 58 62 65 66 67

#### Notes

< > | 1 - 308 / 29 | Go | 🔍 🔍 | ↔️ ⬆️ ⬆️ | 🖼️

#### TWO CREEDS

21

with all its contents to fishermen along the coast. The story of the finding of the corpses of the drowned sailors, all entangled among the wreckage, and of how the **Japanese** fishermen collected them reverently, saying, with the faith of the ancient Greeks, that their souls would wander restless and distressed unless they were laid in their graves and the funeral prayers sung over them. So they sent a petition to the great *Ijin San* in Tokyo praying him to come to the temple of Ikkégami, that his dead brothers might have some one of their own race, if not of their own family, to perform the last solemn rites. And the Ambassador came to Ikkégami, and the long line of weather-beaten **Japanese** fishermen bore the western sailors up the hill to the temple, and buried them in the courtyard, under the silent trees, with all the rites of the Buddhist church. And they set up the wooden lath as over the grave of a brother, among the long lines of the tombstones of their fathers; but they wrote on it in the tongue of the stranger so that God and their countrymen might know their own again.

And all this they did out of their own hearts, and

### DOCUMENT TEXT

OCR Confidence: **63%** [Learn how this text was created](#)

with all its contents to fishermen along the coast. The story of the finding of the corpses of the drowned sailors, all entangled among the wreckage, and of how the Japanese fishermen collected them reverently, saying, with the faith of the ancient Greeks, that their souls would wander restless and distressed unless they were laid in their graves and the funeral prayers sung over them. So they sent a petition to the great *Ijin San* in Tokyo praying him to come to the temple of Ikkégami, that his dead brothers might have some one of their own race, if not of their own family, to perform the last solemn rites. And the Ambassador came to Ikkégami, and the long line of weather-beaten Japanese fishermen bore the western sailors up the hill to the temple, and buried them in the courtyard

# 『タイムズ』紙と婦人参政権運動



TO THE EDITOR OF THE TIMES.

Sir,—May I express my entire agreement with the letter of Miss Milner in your issue of this morning? If the recent scenes of rowdiness associated with women's franchise only served to bring ridicule on the self-appointed champions of that cause other women might be well content to let the matter rest there. Unfortunately, such behaviour can only have the most mischievous effect in prejudicing the influence of women in those branches of public life where the beneficial character of their work is universally recognized.

It is often said of women that neither logic nor humour counts among their strongest points. The recent behaviour of the **suffragettes** would appear to support this contention. Mrs. Fenwick Miller's letter in *The Times* this morning is in every way a remarkable document. It opens up an attractive vista of the public results we might expect to follow from the establishment of feminine rule marked by such a judicious and temperate spirit, say, at the Board of Trade or India Office. As an onlooker nothing strikes me as more curious in this controversy than the unreasonable but most feminine desire of the **suffragettes** both to eat and to keep their political and domestic cake. Women cannot expect to have it both ways. They cannot at one and the same time

```
<pubDateStart>19110000</pubDateStart>
<pubDateEnd>19140000</pubDateEnd>
</pubDate>
<releaseDate>20120601</releaseDate>
<sourceLibrary>
  <libraryName>Columbia University Law Library</libraryName>
  <libraryLocation>New York, NY, United States</libraryLocation>
</sourceLibrary>
<language ocr="English" primary="T"/>English</language>
<notes>Vol. 11, 12, 23, 26, 27, 29-31 never published; so more published after v. 32 American edition
Oskar Borchart Danish, English, French, German, Spanish, or Swedish.</notes>
<comments>Includes bibliographical references and indexes</comments>
<locSubjectHead type="topicalItem">
  <locSubject subField="a">Commercial law</locSubject>
</locSubjectHead>
<authorGroup role="author">
  <author>
    <composed>Honourable Sir Thomas Edward Scrutton, Judge</composed>
    <prefix>Judge</prefix>
    <first>Thomas</first>
    <middle>Edward</middle>
    <last>Scrutton</last>
    <birthDate>1856</birthDate>
    <deathDate>1934</deathDate>
  </author>
</authorGroup>
<authorGroup role="author" terms="added">
  <author>
    <composed>W
    <first>William
    <last>Boyste
    <deathDate>d
  </author>
</authorGroup>
<titleGroup>
  <fullTitle>
    Translation</fullTitle>
    <displayTitle>
    Nations Together ..
  </titleGroup>
  <editionStatement>M
  <imprint>
    <imprintFull>
    <imprintPub>
  </imprint>
  <publicationPlace>
    <publication>
    <publishing>
  </publication>
  <totalPage>
</title>
```

ノッティンガム大学 キャット・グプタ先生



# 『タイムズ』紙と婦人参政権運動



TO THE EDITOR OF THE TIMES.

Sir,—May I express my entire agreement with the letter of Miss Milner in your issue of this morning? If the recent scenes of rowdiness associated with women's franchise only served to bring ridicule on the self-appointed champions of that cause other women might be well content to let the matter rest there. Unfortunately, such behaviour can only have the most mischievous effect in prejudicing the influence of women in those branches of public life where the beneficial character of their work is universally recognized.

It is often said of women that neither logic nor humour counts among their strongest points. The recent behaviour of the **suffragette** would appear to support this contention. Mrs. Fenwick Miller's letter in *The Times* this morning is in every way a remarkable document. It opens up an attractive vista of the public results we might expect to follow from the establishment of feminine rule marked by such a judicious and temperate spirit, say, at the Board of Trade or India Office. As an onlooker nothing strikes me as more curious in this controversy than the unreasonable but most feminine desire of the **suffragettes** both to eat and to keep their political and domestic cake. Women cannot expect to have it both ways. They cannot at one and the same time

1913\_ML.mut

| N  | Word 1      | Freq  | Word 2       | Freq  | Texts | Gap | Join | M    | Z      |
|----|-------------|-------|--------------|-------|-------|-----|------|------|--------|
| 1  | SUFFRAGISTS | 615   | ATTRIBUTED   | 56    | 20    | 2   | 22   | 9.83 | 44.29  |
| 2  | SUFFRAGISTS | 527   | COLOURS      | 39    | 10    | 1   | 11   | 9.58 | 28.61  |
| 3  | SUFFRAGISTS | 615   | MILITANT     | 353   | 93    | 1   | 116  | 9.58 | 92.82  |
| 4  | SUFFRAGIST  | 527   | BOMBS        | 25    | 4     | 1   | 7    | 9.57 | 22.73  |
| 5  | SUFFRAGE    | 1,386 | WVOMAN       | 30    | 18    | 1   | 22   | 9.56 | 40.21  |
| 6  | SUFFRAGIST  | 527   | LITERATURE   | 87    | 21    | 1   | 23   | 9.48 | 40.01  |
| 7  | SUFFRAGIST  | 527   | BREAKERS     | 18    | 3     | 2   | 5    | 9.56 | 19.14  |
| 8  | SUFFRAGE    | 1,386 | WOMAN        | 1,178 | 253   | 1   | 683  | 9.22 | 198.56 |
| 9  | SUFFRAGE    | 1,386 | ADULT        | 63    | 25    | 1   | 43   | 9.46 | 54.17  |
| 10 | SUFFRAGISTS | 615   | ANTI         | 168   | 30    | 1   | 43   | 9.21 | 49.69  |
| 11 | SUFFRAGIST  | 527   | ANTI         | 168   | 23    | 1   | 33   | 9.06 | 41.11  |
| 12 | SUFFRAGE    | 1,386 | WOMAN        | 25    | 9     | 1   | 12   | 8.95 | 23.85  |
| 13 | SUFFRAGE    | 1,386 | OPPOSING     | 66    | 22    | 2   | 30   | 8.87 | 36.66  |
| 14 | SUFFRAGE    | 1,386 | MANHOOD      | 31    | 13    | 1   | 16   | 9.05 | 28.60  |
| 15 | SUFFRAGIST  | 527   | PUBLICATIONS | 35    | 4     | 1   | 6    | 8.86 | 16.33  |
| 16 | SUFFRAGIST  | 527   | OUTRAGE      | 128   | 19    | 1   | 21   | 8.80 | 29.86  |
| 17 | SUFFRAGISTS | 615   | SUSPECTED    | 43    | 8     | 2   | 8    | 8.75 | 18.15  |
| 18 | SUFFRAGIST  | 527   | DISORDER     | 92    | 9     | 3   | 13   | 8.58 | 21.72  |
| 19 | SUFFRAGIST  | 527   | CONSPIRACY   | 90    | 11    | 1   | 13   | 8.61 | 21.97  |
| 20 | SUFFRAGE    | 1,386 | SOCIETIES    | 195   | 48    | 1   | 72   | 8.57 | 50.93  |
| 21 | SUFFRAGE    | 1,386 | FEMALE       | 113   | 23    | 1   | 41   | 8.55 | 38.08  |
| 22 | SUFFRAGE    | 1,386 | FEDERATED    | 14    | 5     | 3   | 5    | 8.52 | 13.19  |
| 23 | SUFFRAGIST  | 527   | SUPPOSED     | 132   | 16    | 1   | 17   | 8.45 | 23.65  |
| 24 | SUFFRAGIST  | 527   | DISTURBANCE  | 82    | 9     | 1   | 11   | 8.51 | 19.44  |

frequency alphabetical statistics filenames mutual information notes

330 Type-in OUTRAGE

# 2つの異なるワークフロー

thailand

Advanced Search

GALE PRIMARY SOURCES

West of the Mekong

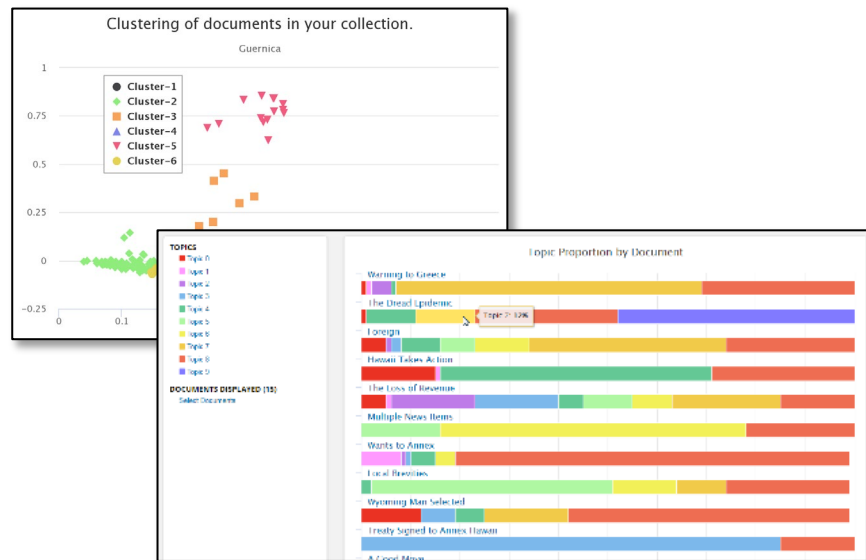
Publication: The Economist

GALE PRIMARY SOURCES

West of the Mekong

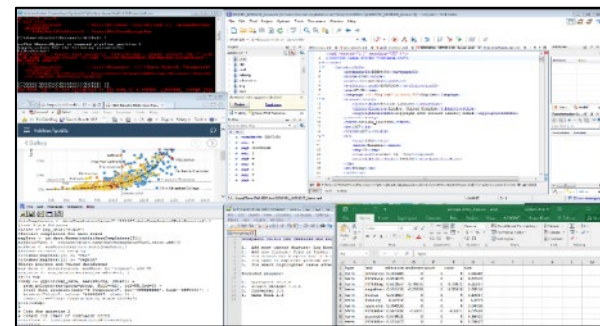
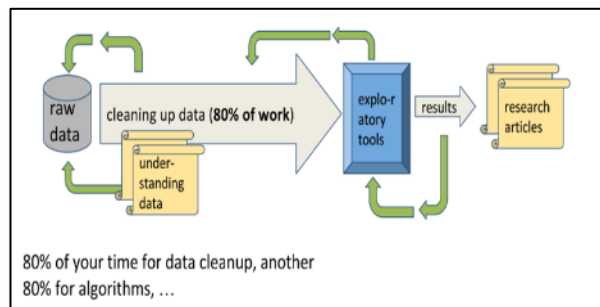
GALE PRIMARY SOURCES

検索 → 閲読

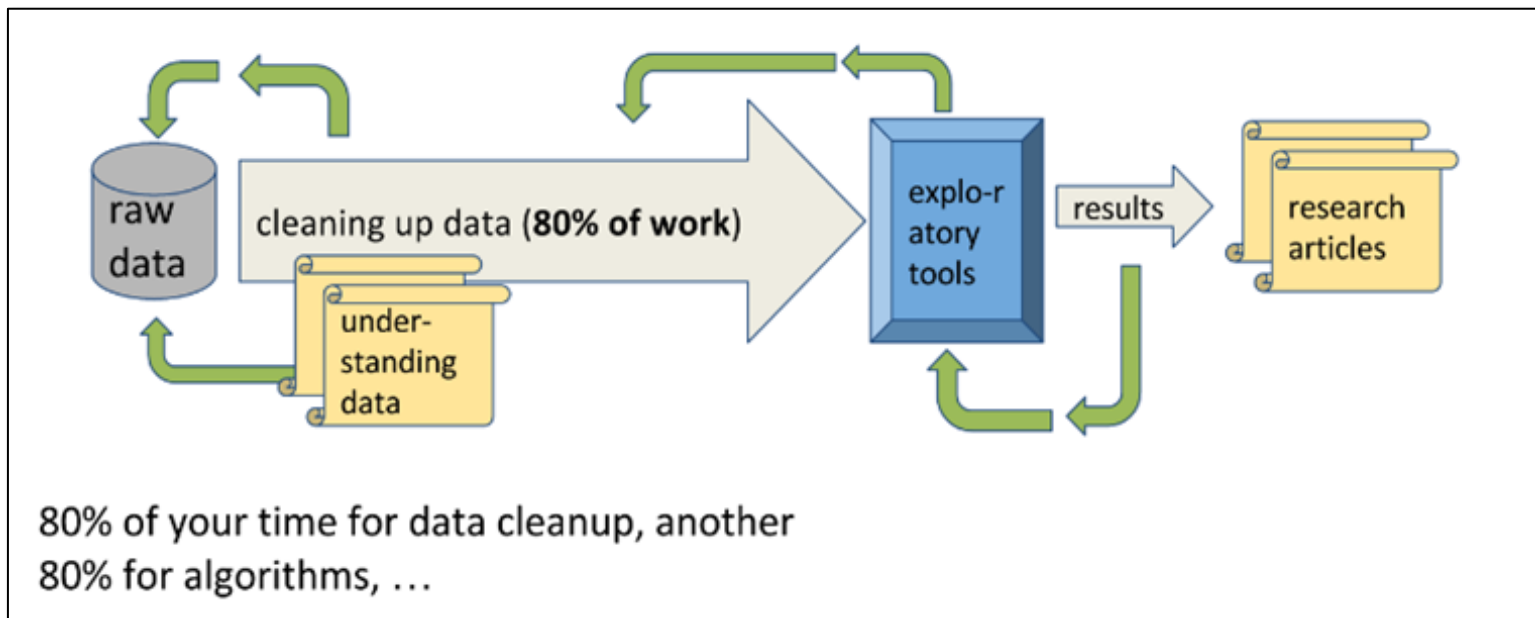


データ収集 → 分析

# デジタル人文学：3つの障害



# 1: 分析用に最適化された元データの入手



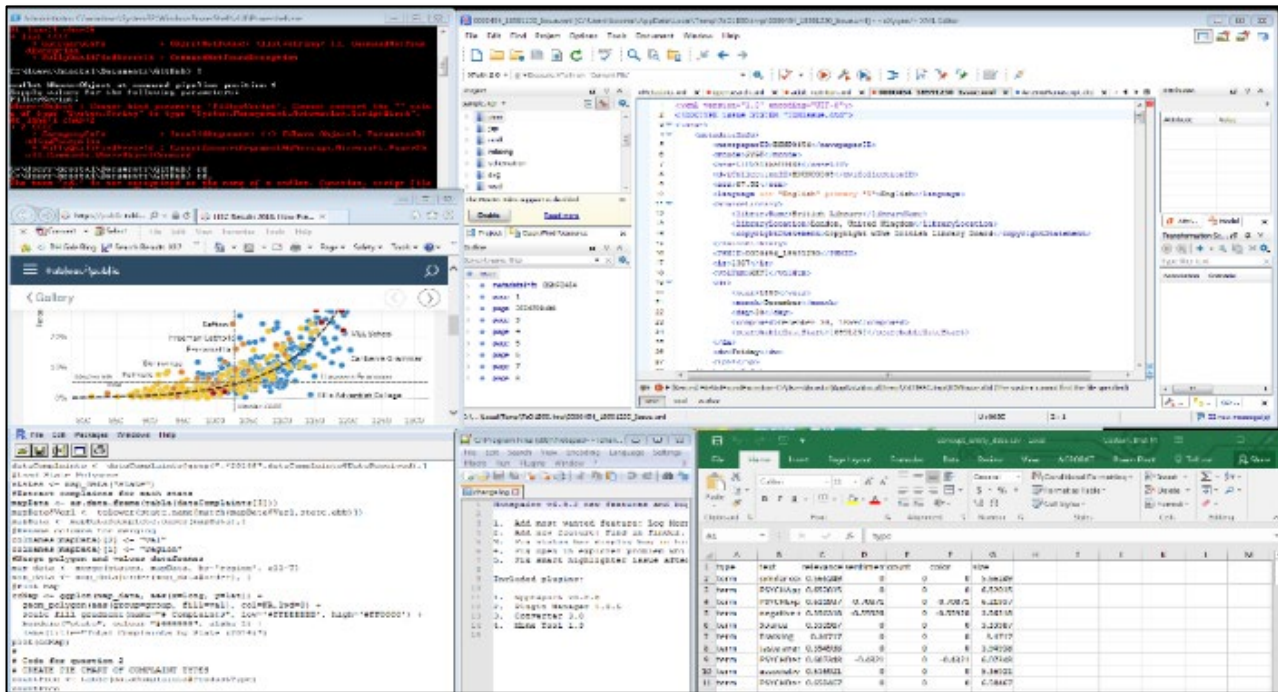
Slide courtesy of the COMHIS Collective: <https://comhis.github.io/> <http://j.mp/comhis-bsecs>



## 2: データの保存場所の確保と管理



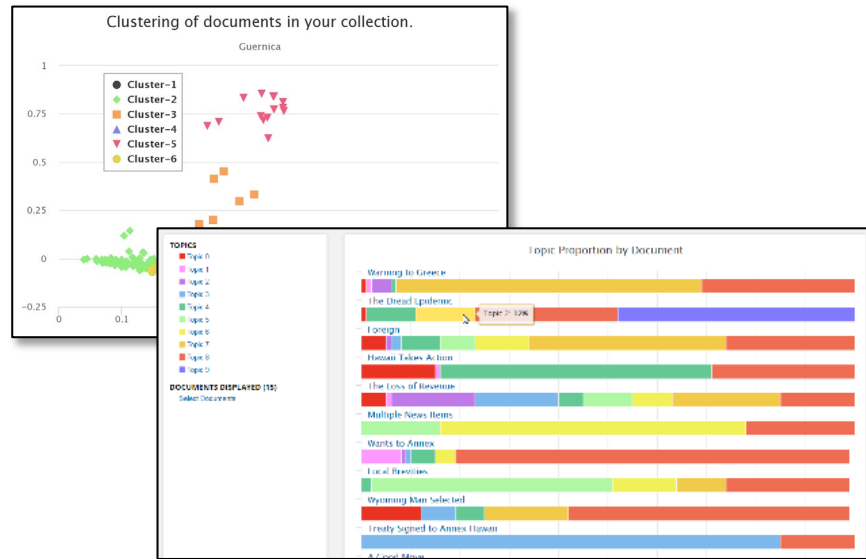
# 3: 分析ツールの難しさ



# 2つの異なるワークフロー

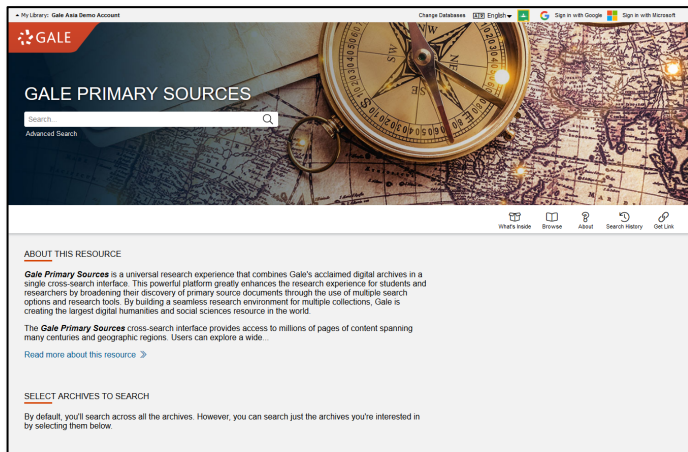
The screenshot shows the GALE PRIMARY SOURCES search interface. At the top, the text 'GALE PRIMARY SOURCES' is displayed. Below it, a search bar contains the word 'thailand'. The interface includes navigation options like 'Advanced Search' and 'My Library'. The main content area displays a search result for 'West of the Mekong' by The Economist. The article text is visible, discussing the region of Laos and Thailand. A map of Thailand and Laos is also shown on the right side of the article view.

検索 → 閲読

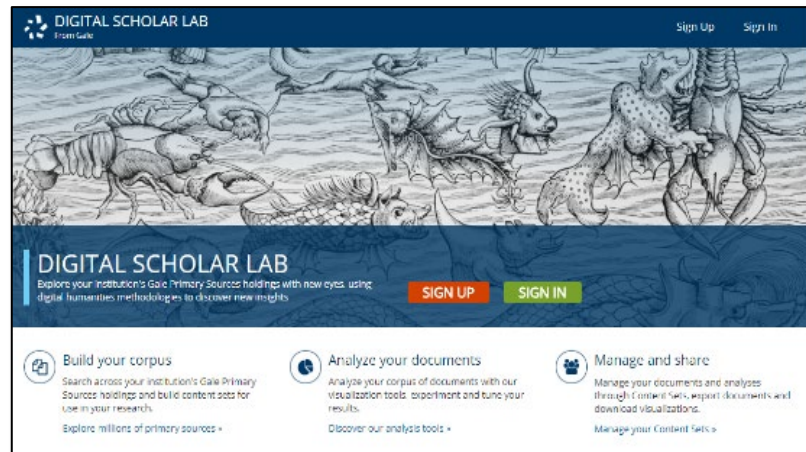


データ収集 → 分析

# それぞれに最適なプラットフォーム



**Gale Primary Sources**



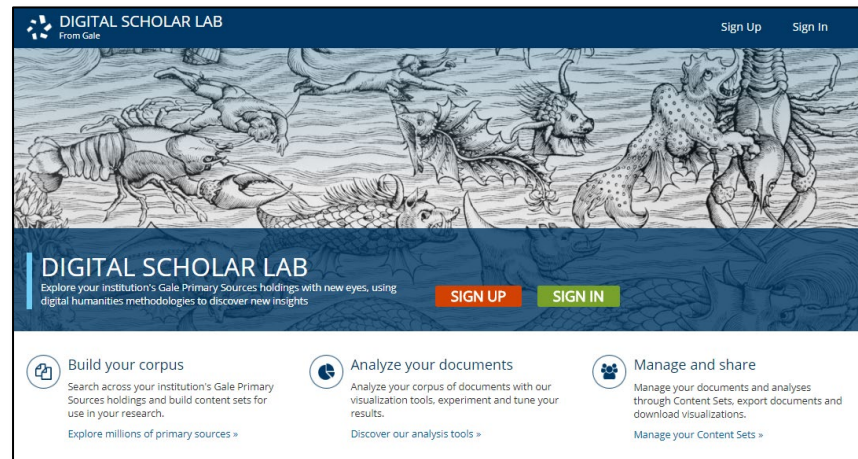
**Gale Digital Scholar Lab**



# Gale Digital Scholar Lab

## Galeの一次史料テキストマイニング専用プラットフォーム

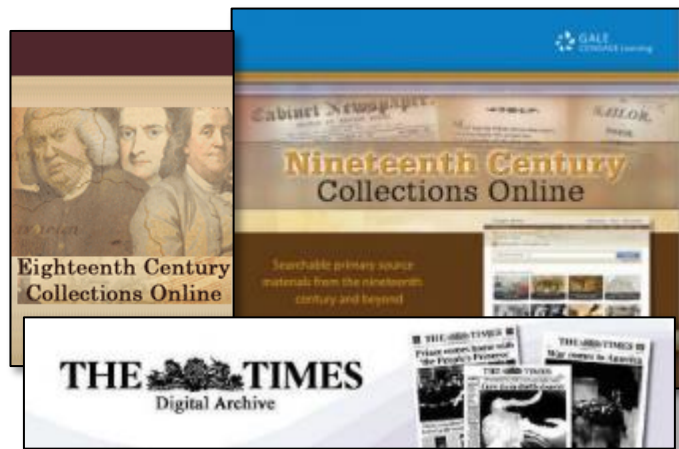
- Gale Primary Sources データベース群のOCRテキストを搭載
- 主要なテキストマイニング・ツール(分析ツール)を内蔵
- OCRクリーニング・ツールを内蔵
- 分析結果・OCRテキストのダウンロード機能
- ユーザー自身のテキストのアップロード機能(ベータ版)
- 充実したヘルプファイル、カリキュラムサポート
- プログラミング等の知識は不要!



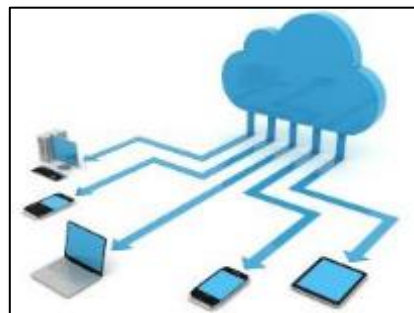
The screenshot shows the Gale Digital Scholar Lab website. At the top, there is a navigation bar with the logo 'DIGITAL SCHOLAR LAB From Gale' on the left and 'Sign Up' and 'Sign In' links on the right. Below the navigation bar is a large illustration of various animals, including a scorpion, a dragon, a goat, and a lobster. Underneath the illustration, the text reads 'DIGITAL SCHOLAR LAB' followed by a sub-headline: 'Explore your institution's Gale Primary Sources holdings with new eyes, using digital humanities methodologies to discover new insights.' There are two buttons: 'SIGN UP' in orange and 'SIGN IN' in green. Below this section are three main features, each with an icon and a brief description:

- Build your corpus**: Search across your institution's Gale Primary Sources holdings and build content sets for use in your research. Explore millions of primary sources >
- Analyze your documents**: Analyze your corpus of documents with our visualization tools, experiment and tune your results. Discover our analysis tools >
- Manage and share**: Manage your documents and analyses through Content Sets, export documents and download visualizations. Manage your Content Sets >

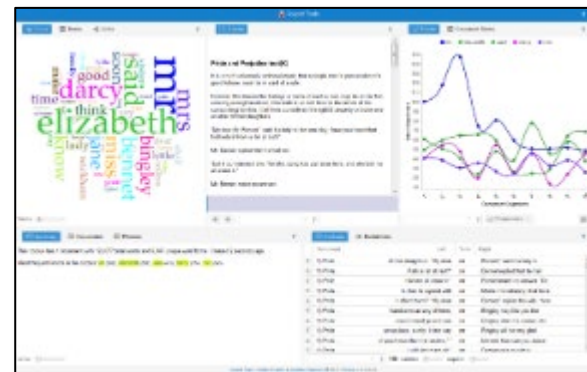
# 3つの障害を解決



既に分析用に最適化されたデータを格納



クラウド上で管理されている



主な分析ツールを搭載

# 現在 Lab に搭載可能なコンテンツセット

- Amateur Newspapers from the American Antiquarian Society
- American Fiction 1774-1920
- American Historical Periodicals from the American Antiquarian Society
- Archives Unbound
- Archives of Sexuality and Gender
- Associated Press Collections Online
- Brazilian and Portuguese History and Culture
- British Library Newspapers
- China and the Modern World
- Crime, Punishment, and Popular Culture, 1790-1920
- Daily Mail Historical Archive 1896-2016
- Declassified Documents Online: Twentieth Century British Intelligence
- The Economist Historical Archive, 1843-2015
- Eighteenth Century Collections Online (ECCO)
- The Illustrated London News Historical Archive 1842-2003
- The Independent Historical Archive
- Indigenous Peoples of North America
- International Herald Tribune Historical Archive, 1887-2013
- Liberty Magazine Historical Archive, 1924-1950
- The Listener Historical Archive, 1929-1991
- The Making of Modern Law: American Civil Liberties Union Papers
- The Making of Modern Law: Foreign Primary Sources, 1600–1970
- The Making of Modern Law: Foreign, Comparative and International Law, 1600–1926
- The Making of Modern Law: Landmark Records and Briefs of the U.S. Courts of Appeals, 1950-1980
- The Making of Modern Law: Legal Treatises, 1800–1926
- The Making of Modern Law: Primary Sources
- The Making of Modern Law: Trials, 1600–1926
- The Making of Modern Law: U.S. Supreme Court Records and Briefs, 1823–1978
- The Making of the Modern World (MOMW)
- Mirror Historical Archive 1903-2000
- Nineteenth Century Collections Online (NCCO)
- Nineteenth Century U.S. Newspapers
- Nineteenth Century UK Periodicals
- Picture Post Historical Archive 1938-1957
- Political Extremism and Radicalism
- Public Health Archives: Public Health in Modern America, 1890-1970
- Punch Historical Archive 1841-1992
- Refugees, Relief, and Resettlement
- Sabin Americana: History of the Americas, 1500-1926
- Seventeenth and Eighteenth Century Burney Newspapers Collection
- Seventeenth and Eighteenth Century Nichols Newspapers Collection
- Slavery and Anti-Slavery: A Transnational Archive
- Smithsonian Collections Online
- The Sunday Times Historical Archive 1822-2016
- The Telegraph Historical Archive 1855-2016
- The Times Digital Archive 1785-2014
- The Times Literary Supplement Historical Archive 1902-2019
- U.S. Declassified Documents Online
- Women's Studies Archive

# 現在 Lab に搭載できないコンテンツセット

- Financial Times Historical Archive
- National Geographic Magazine Archive
- Chatham House Online Archive
- World Scholar: Latin America & the Caribbean
- British Literary Manuscripts Online
- State Papers Online
- Early Arabic Printed Books from the British Library

## 主な理由:

- 権利関係の理由 (Financial Times、National Geographic)
- プラットフォームが未統合 (Chatham House、World Scholar)
- 手稿中心でOCRがない (British Literary Manuscripts、State Papers)
- 言語の障壁 (Early Arabic Printed Books)



# Gale Digital Scholar Lab デモンストレーション

# Gale Digital Scholar Lab のワークフロー



# 19世紀～20世紀前半の『タイムズ』紙における「社会主義」

- The Times Digital Archive 1785-2014
- タイトルが「socialis\*」を含む新聞記事
- 年代: 1820～1945

**SOCIALISM EXPOUNDED BY A FEMALE.**

The first of a course of three lectures on "social principles" was delivered last night at the "Lambeth Social Institution," as a room in the Westminster-road, recently occupied as a sing-song bazaar, is now called, by a feminine lecturer, who announces herself as "Mrs. Chappellsmith," a name which sounds very like a combination of Alice, indicating that the wearer of it has had some practical experience in the principles she propounds. She answered declarations herself to be the "late Miss Reynolds," one, if we are not mistaken, not unknown to fame—such fame as clings to the long-dormant, Botmda, and the almost forgotten self-titled David's Daughters, Judge, perhaps, that "Mrs. Reynolds-Chappellsmith" would not be a very conspicuous *non de guerre*, she has separated the names; but it would be no bad rule for the female socialist to imitate in some measure the lady in the Eastern tale, who numbered her lovers by rings on her fingers until she was completely encased; if they would, as often as they change partners, add each successive name to their former one, their appellations would become truly maguiloquent, and proclaim at the same time their triumph over the weakness and prejudices of mankind.

The room, which is of octagonal form, bodily ventilated, and capable of seating about 120 persons, was nearly filled by auditors of both sexes ranged before a platform, at one end of which was placed a large plaster bust of Milton, with a "most melancholic" expression, and at the other end, a table with a

**SOCIALISTS AND HITLERISM**

**INTERNATIONAL PROTEST URGED**

FROM OUR OWN CORRESPONDENT

PARIS, AUG. 27

The meeting of the Second (Socialist) International ended here early yesterday morning, after an all night sitting, with the adoption of a series of resolutions demanding an intensification of the struggle against capitalism, Fascism, and Hitlerism, and calling for an immediate resumption of the Disarmament Conference, together with the complete prohibition of the private manufacture of armaments. The resolution on disarmament also says that the general strike remains the supreme weapon of the working classes against war, adding that refusal to submit to arbitration should automatically be taken as proof of aggression.

The resolution dealing with Fascism and Hitlerism principally recommends the organization by the Socialists of all countries of

**SOCIALISM AND THE EMPIRE.**

**LORD OLIVIER'S ADDRESS.**

LORD OLIVIER gave an address on "Socialism and the Empire" at the conference of the University Labour Federation, at Manchester University on Saturday. He said that the complaint of Asiatic peoples against British Imperialism was in effect that relations which began as innocent trading relations had after several generations resulted in subservience and control to the nation with whom they traded. The Simon Commission was regarded with the greatest possible hostility in India. In his opinion the opposition on the part of Indians to giving any assistance to the Commission had been very bad political tactics and would increase the difficulties of a comfortable transition to Dominion status.

If the British Government had not taken control in the colonies the natives would have been eaten up by the independent action of capitalist Europeans and private adventurers. If the control of the British Government were taken away, exploitation of the natives would continue. It was the duty of the British Empire, whatever political party was in power, to protect the natives. What the Labour Party was standing for now was the inter-

**MR. CHURCHILL DENOUNCES SOCIALISTS**

**COUNTRY'S RECOVERY HAMPERED BY DOCTRINAIRES**

Mr. Churchill, in a vigorous speech to the Central Council of the Conservative Party yesterday, severely criticized the Socialist Government's conduct of affairs.

Blaming them for Britain's slow recovery, he said:—"We are harassed, harried, hampered, tied down, and stifled for the sake of vague, ill-though-out, and physically unattainable plans for a Socialist future."

**DANGERS AHEAD**

**LOWERED STANDARD OF LIVING**

The Central Council of the National Union of Conservative and Unionist Associations, which met at the Friends' House, Euston Road, yesterday, was addressed by Mr. Churchill during the morning session. The chairman, Mr. R. A. Butler, M.P., introduced Mr. Churchill as "The Master Fighter." When Mr. Churchill moved to the microphone, the delegates rose, cheered vociferously, and

**"COLLAPSE OF SOCIALISM"**

**THE MORAL OF EVENTS IN ENGLAND**

Addressing the 1913 Club last night Mr. VICTOR MOLLO said that never was it clearer than to-day that Socialism had ignominiously collapsed. A review of the dismal wreckage showed that State shipping, trading, railways, mines, whether it be in Australia, Canada, New Zealand, France, or the United States, had almost without exception resulted in huge losses. The deficit had had to be made up by taxing or borrowing from the resources of private enterprise. Collectivization, so essential to Socialism, so beloved by Fabians, was definitely and rapidly receding.

From the initial stage of Socialism they could pass on to its full maturity in Russia. All labour in Russia was compulsory. No matter what goods the Soviets exported, whether it be cereals, oil, dairy products, or finished articles, they invariably undersold

**Imperial and Foreign**

**SOCIALISM IN INDIA**

**CONGRESS'S TWO VOICES**

From Our Delhi Correspondent

Socialism, which will have a voice in the Congress gathering at Lucknow this weekend, is gradually emerging as a reality in Indian politics. While it exists loosely and is still unrepresented by any *ad hoc* organization, it is nevertheless gaining ground among the politically conscious people of this country. Socialists within the ranks

# Gale Digital Scholar Lab で現在利用可能な分析ツール

| ツール名                            | 主な用途              | ライブラリ・ツールキット | オープンソース |
|---------------------------------|-------------------|--------------|---------|
| Nグラム頻度 Ngram                    | 頻出単語・フレーズの抽出      | Lucene       | ×       |
| クラスタリング Clustering              | 文書を類縁性で分類         | SciKit Learn | ○       |
| トピック・モデリング Topic Modelling      | 複数の文書に共通するトピックの抽出 | Mallet       | ○       |
| 感情分析 Sentiment Analysis         | 文書内容が肯定的か否定的かを分析  | OpenNLP      | ○       |
| 固有表現抽出 Named Entity Recognition | 固有名詞・数字・日付等の抽出    | spaCy        | ○       |
| 品詞タグ付け Parts-of-Speech Tagger   | 著者ごとに品詞の使用頻度を比較   | spaCy        | ○       |



# Gale Digital Scholar Lab のワークフロー



- 基本検索／詳細検索
- メタデータの表示
- ページ画像とOCRテキストの対照表示
- 検索履歴(セッション毎)
- 検索結果の文書を一度に選択可能(ただし最大1万件)
- 自身のコンテンツのアップロード(ベータ版)

- 検索／アップロードした文書のコンテンツセットへの追加
- コンテンツセットの名称・説明文を指定
- 各コンテンツセットに最大1万件の文書を追加可能
- 複数のコンテンツセットを管理
- コンテンツセットの複製
- 検索履歴の表示

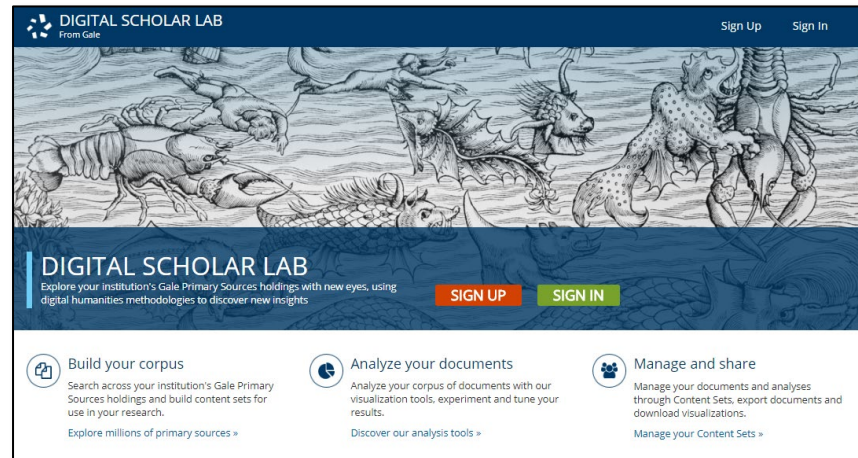
- デフォルトのクリーニングを適用する(特段の操作は不要)
- もしくは、任意のクリーニング設定を作成し、適用する

- 分析ツールの適用
- プログラミング技術は不要
- 視覚化／表形式の出力
- 簡単な設定項目
- ドキュメントへのリンク
- 実行履歴からの再現
- 1セッションあたり最大5千件の文書のOCRをダウンロード

# Gale Digital Scholar Lab

## Galeの一次史料テキストマイニング専用プラットフォーム

- Gale Primary Sources データベース群のOCRテキストを搭載
- 主要なテキストマイニング・ツール(分析ツール)を内蔵
- OCRクリーニング・ツールを内蔵
- 分析結果・OCRテキストのダウンロード機能
- ユーザー自身のテキストのアップロード機能(ベータ版)
- 充実したヘルプファイル、カリキュラムサポート
- プログラミング等の知識は不要！



The screenshot shows the homepage of the Digital Scholar Lab. At the top, there is a navigation bar with the logo 'DIGITAL SCHOLAR LAB From Gale' on the left and 'Sign Up' and 'Sign In' links on the right. Below the navigation bar is a large illustration of various animals, including a scorpion, a dragon, a goat, and a lobster, set against a dark blue background. Below the illustration, the text 'DIGITAL SCHOLAR LAB' is displayed in large, bold letters, followed by the tagline 'Explore your institution's Gale Primary Sources holdings with new eyes, using digital humanities methodologies to discover new insights.' To the right of this text are two buttons: 'SIGN UP' (orange) and 'SIGN IN' (green). Below the main content area, there are three columns of text, each with an icon and a title: 'Build your corpus' (with a magnifying glass icon), 'Analyze your documents' (with a globe icon), and 'Manage and share' (with a group of people icon). Each column contains a brief description of the service and a link to explore more.

# インターフェースの更新予定について

# Gale Digital Scholar Lab インターフェース更新予定

- 新インターフェースの特色
  - 分析ツールの改良
  - 使いやすさ・見やすさの向上
  - 他のGale商品と使い勝手を統一
  - アクセサビリティの向上 (WCAG 2.0準拠)
- 2021年6月：新インターフェースを利用可能に (新旧インターフェースを並行して運用)
- 2021年9月：新インターフェースに完全移行

※時期は変更になる場合がございます。

## 分析結果の移行について

- インターフェース移行後も、ユーザーのコンテンツセットおよびクリーニング設定は保持されます。
- ただし、分析ツールのデザイン仕様が変更のため、分析結果のアウトプット (画像・CSV・JSON形式) は新旧インターフェースでは異なります。
- 念のため、完全移行前に旧インターフェースによる分析結果をローカルにダウンロードし保存しておくことをお勧めします。
- 上記の点についてはデータベース内においても通知する予定です。

# コンテンツセット画面

- フォルダーによる整理
- 複数のコンテンツセットを結合する機能

## 現行インターフェース

The current interface shows a list of content sets under the heading "CONTENT SETS". Each set includes a checkbox, a name, and options for "Edit" and "Copy". Below each set, "QUICK STATS" are provided, including the number of documents, archives used, and the time frame.

| Content Set   | Documents | Archives Used | Timeframe   |
|---------------|-----------|---------------|-------------|
| Big Cheese    | 10000     | 31            | 1983 - 2009 |
| Cheese        | 201       | 11            | 1734 - 2016 |
| Content Set 1 | 0         | 0             |             |
| Horses        | 101       | 15            | 1759 - 9999 |
| Wine          | 100       | 10            | 1734 - 2016 |

## 新インターフェース

The new interface features a "My Content Sets" section with a video introduction. Below this is a table with columns for "Type Name", "Build", "Clean", and "Analyze". Each row represents a content set with a description and corresponding statistics.

| Type Name  | Build             | Clean             | Analyze  |          |
|--|-------------------|-------------------|----------|----------|
| Alcohol as Medicinal<br><i>Add content through Build by searching for or uploading documents.</i>  | Documents: 0      | Configurations: 1 | Tools: 0 | Runs: 0  |
| Deep Sea Exploration Reporting<br><i>Reports of deep sea exploration.</i>  | Documents: 5,321  | Configurations: 1 | Tools: 2 | Runs: 2  |
| Motherhood during the 18th Century<br><i>This content set includes articles specifically around writings directed to women providing guidance on parenting. It also includes references to motherhood.</i> | Documents: 10,000 | Configurations: 1 | Tools: 2 | Runs: 2  |
| Women in Science<br><i>Includes writings from women in science and looks for similarities.</i>   | Documents: 1,788  | Configurations: 1 | Tools: 4 | Runs: 17 |
| Unions + 19th Century<br><i>Content pulled about Unions in the 19th Century from a number of historical newspapers including the Financial Times, Economist, and Sunday Times</i>                          | Documents: 209    | Configurations: 1 | Tools: 5 | Runs: 14 |



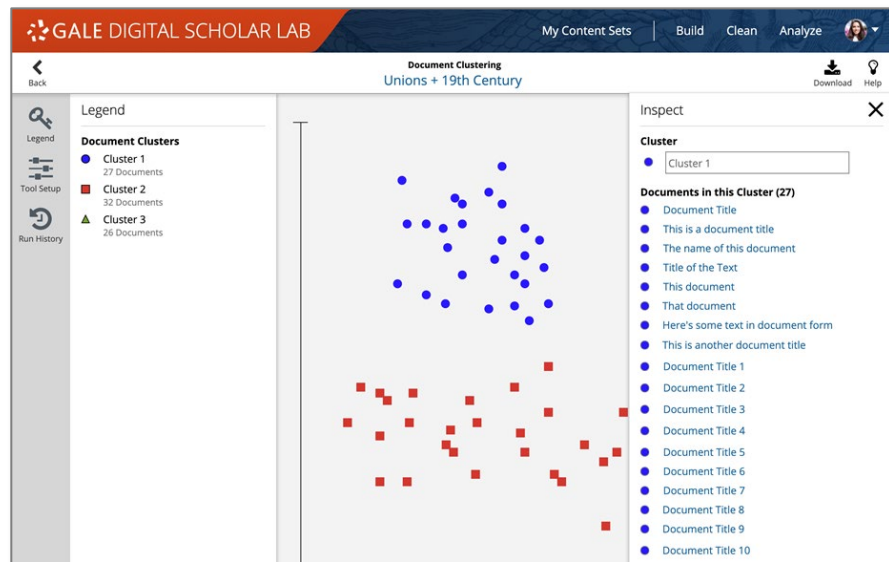
# クラスタリング・ツール

- 特定クラスタの文献の詳細情報を表示
- 特定クラスタをコンテンツセットへ追加する機能

## 現行インターフェース



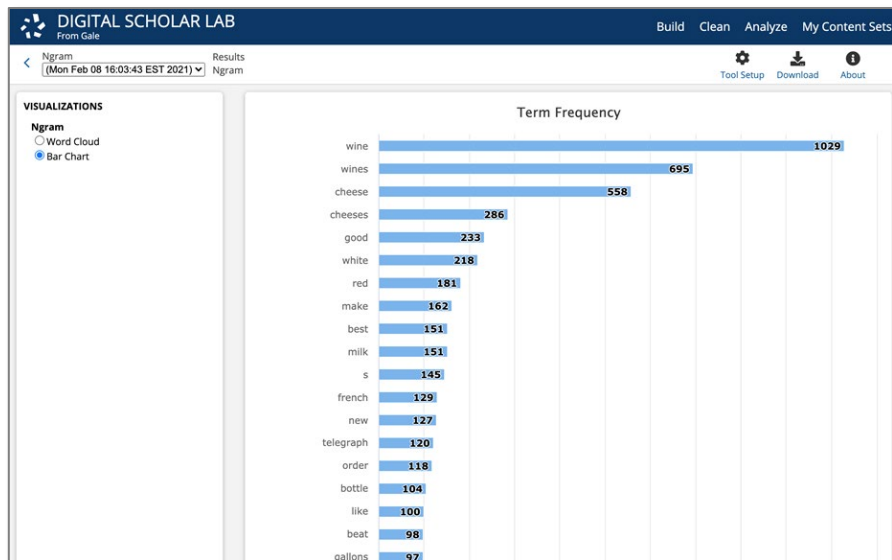
## 新インターフェース



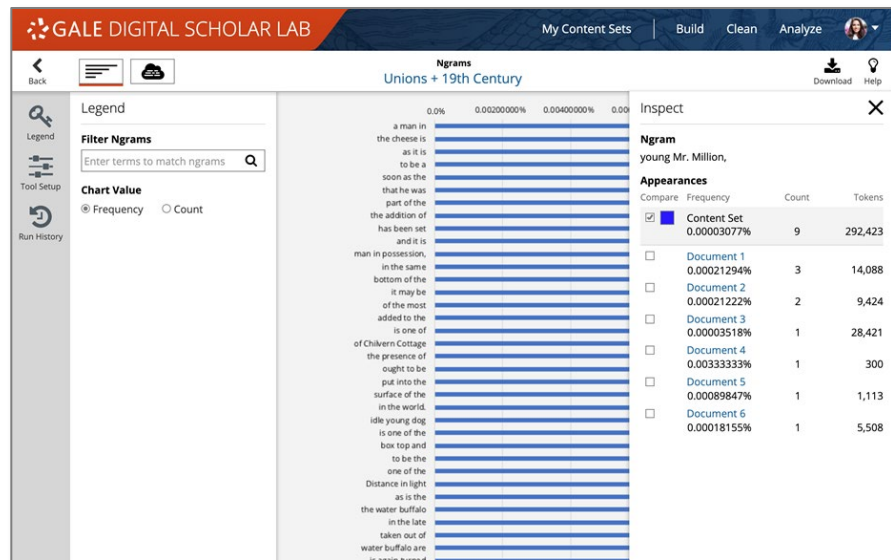
# Nグラム頻度ツール(グラフ表示)

- 個々のNグラムの詳細情報を表示
- 絶定数/頻度の切替表示

## 現行インターフェース



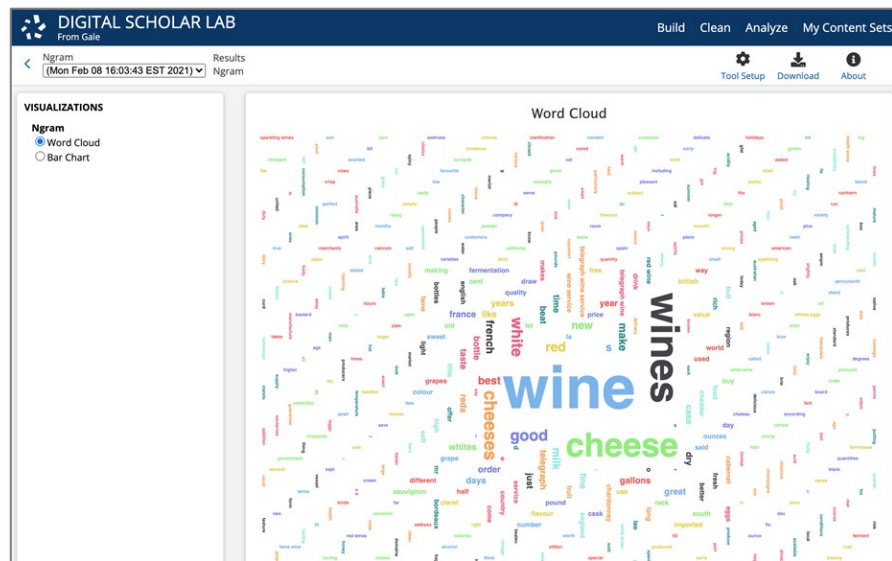
## 新インターフェース



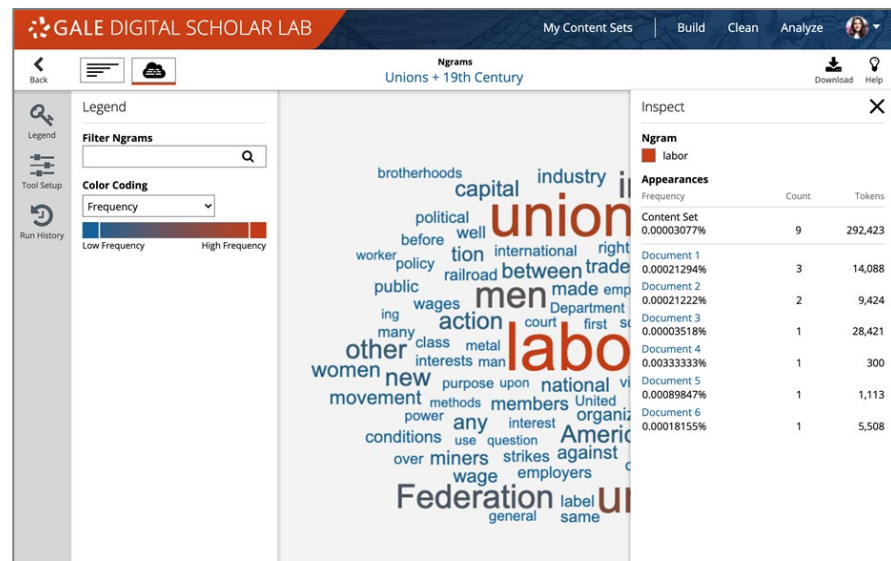
# Nグラム頻度ツール(ワードクラウド)

- より自然なクラウド型に
- 絶対数/頻度によるカラーコーディング

## 現行インターフェース



## 新インターフェース



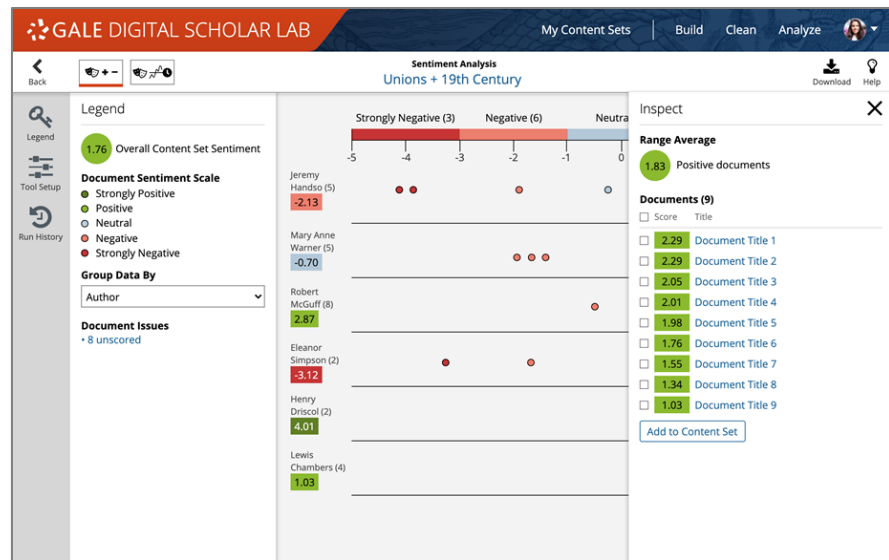
# 感情分析ツール

- メタデータによるグループ分け機能
- 感情スコアの視覚化

## 現行インターフェース



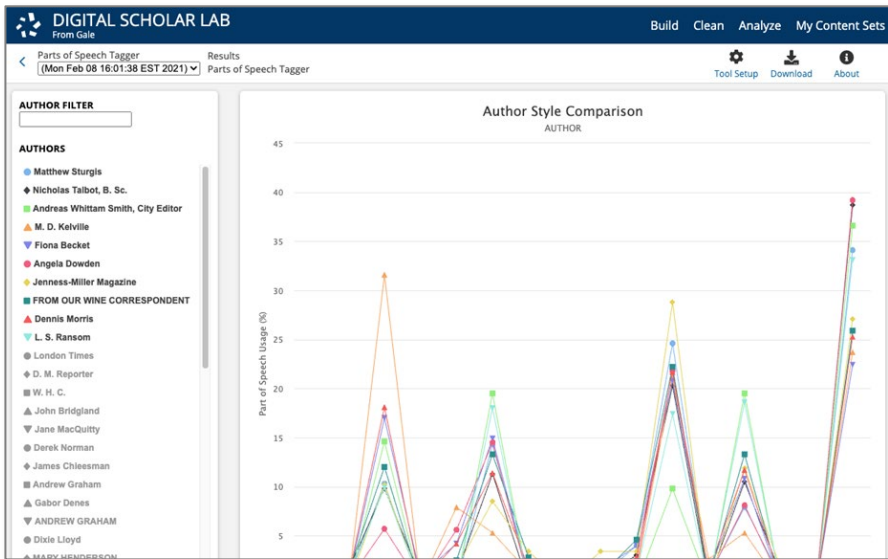
## 新インターフェース



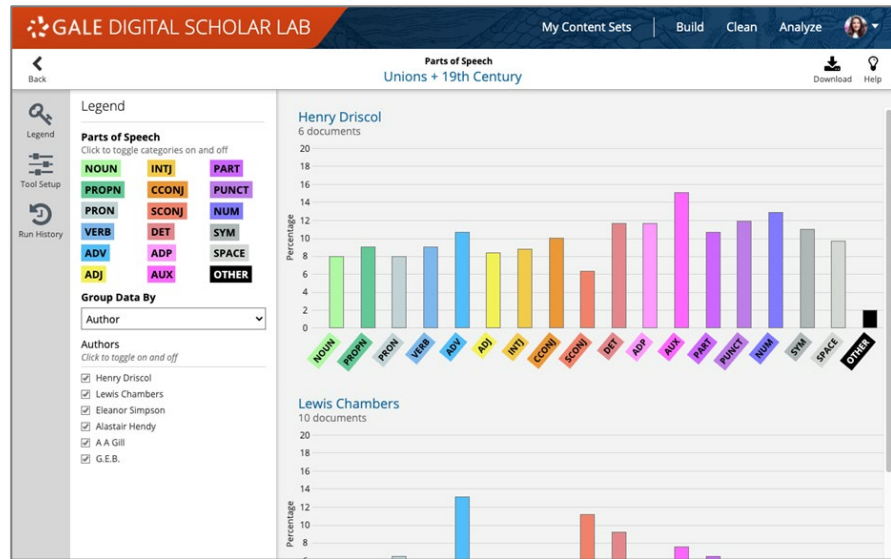
# 品詞タグ付けツール

- メタデータによるグループ分け機能
- 品詞分類の視覚化

## 現行インターフェース



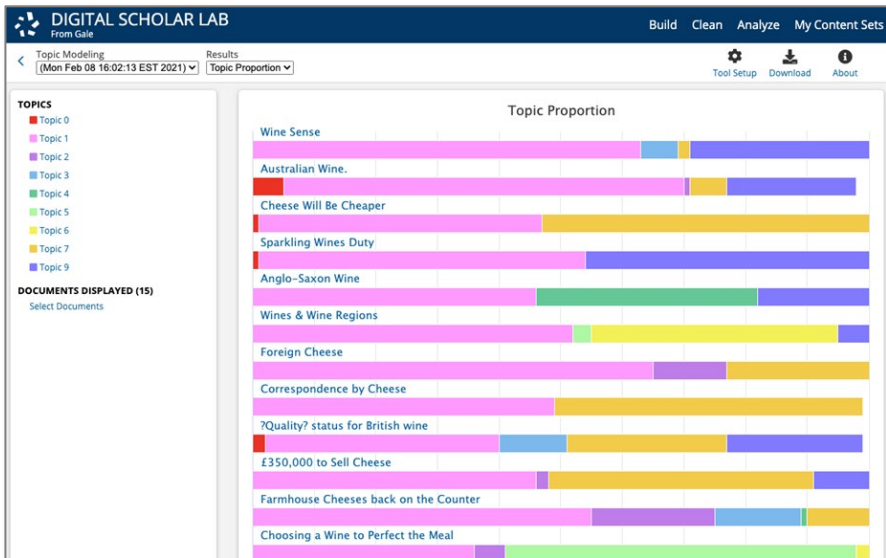
## 新インターフェース



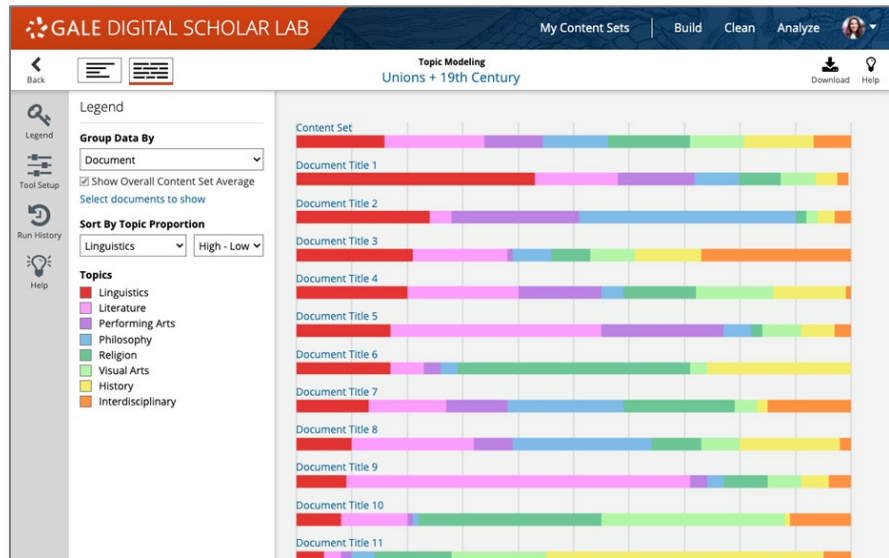
# トピックモデリング・ツール

- ・トピック含有比率による並び替え機能
- ・メタデータによるグループ分け機能

## 現行インターフェース



## 新インターフェース





# まとめ

- 一次資料のデジタル化が可能にする新たな研究
  - 〈検索⇒閲読〉 〈収集⇒分析〉の2つの異なるワークフロー
- デジタル人文学実践の3つの障害
  - 分析に適したデータの入手、その保存・管理、分析ツールの難しさ
- 一部の限られた研究者から、より幅広い層へ広がる可能性
  - デジタル時代の人文学教育に大きな示唆

**ご清聴ありがとうございました。**

**GaleJapan@cengage.com**