

Gale Primary Sources  
Start at the source.

# DIGITAL HUMANITIES

New Ways of Understanding and Exploring  
our History

30<sup>th</sup> May 2017

# Who is Gale?

- Gale is part of Cengage Learning; who with revenue of **\$1.6 Billion**, is the third largest educational publisher in the world
- Gale has been publishing books since **1954**
- Gale has been digitising major world newspapers and historical collections, beginning with the **Times Digital Archive** in 2002
- Gale Digital Archives now include over **250 Million** pages, sourced from the world's most prestigious libraries and archives

# Some of our Partner Libraries



As well as being source libraries for our archives, these institutions are also our customers. Currently **92 out of the top 100\* Universities in the world** have *at least one Gale Primary Source Archive*

\* Times Higher Education 2016 rankings

# Unique, World-class Content

THE  TIMES

The  
Economist



Smithsonian  
Institution



FINANCIAL  
TIMES

THE  INDEPENDENT



CHATHAM HOUSE

 NATIONAL  
GEOGRAPHIC

# Studying Primary Sources

There is opportunity in any situation...

Brussels **7** BELGIUM

Rome **5**  
The Vatican **6** ITALY/  
VATICAN CITY

Taormina **8**  
Sigonella **9**

**2** Tel Aviv  
**3** Jerusalem  
**4** Bethlehem

ISRAEL/  
WEST BANK

**1** Riyadh

SAUDI  
ARABIA





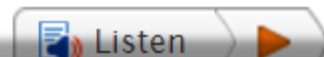


# Trump signs 'tremendous' deals with Saudi Arabia on his first day overseas

Philip Rucker and Karen DeYoung

*The Washington Post*. (May 20, 2017): News:

**Copyright:** COPYRIGHT 2017 The Washington Post  
<http://www.washingtonpost.com/>



Saturday, ushering in a new era in U.S.-Saudi Arabian relations by signing a joint "strategic vision" that includes \$110 billion in American arms sales and other new investments that the administration said would bring hundreds of thousands of jobs.

Byline: Philip Rucker and Karen DeYoung

RIYADH, **Saudi Arabia** - President **Donald Trump** made a splashy debut on the world stage here Saturday, ushering in a new era in U.S.-Saudi Arabian relations by signing a joint "strategic vision" that includes \$110 billion in American arms sales and other new investments that the administration said would bring hundreds of thousands of jobs.



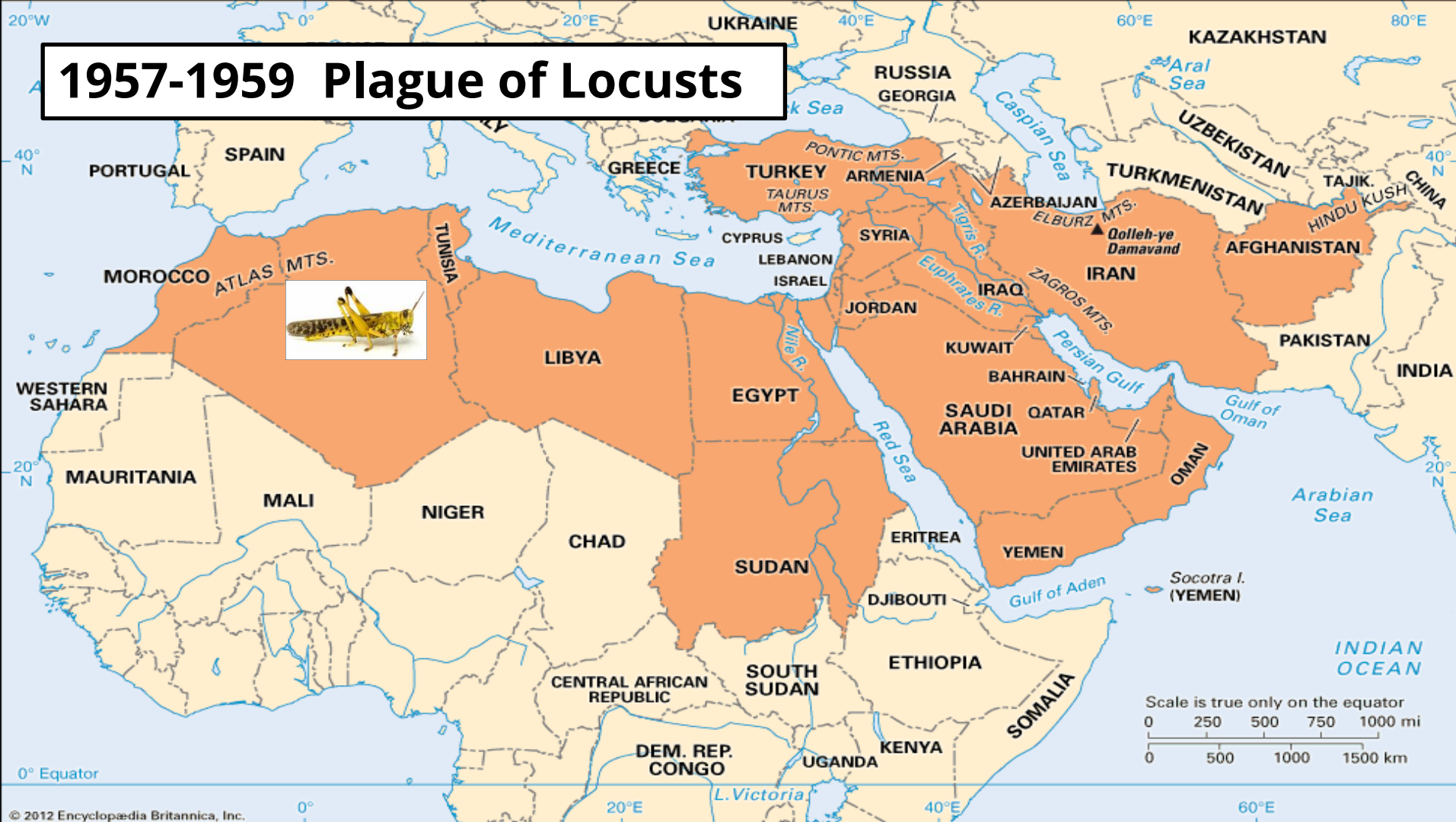
Why Saudi Arabia?







# 1957-1959 Plague of Locusts



# 1957-1959 Plague of Locusts

SUBJECT: Drive Against Locusts in Bahrain.

BPCIN UNCLASSIFIED.

During the first week of May an immature horde of locusts appeared in Bahrain. The Department of Agriculture and the Bahrain Petroleum Co. (B/P Co.) are working on the problem. The problem is a big job which will require US Government assistance presumably through CINCUSAFE. Believe this is a good opportunity if we move fast.

past the insects can be dealt with before reaching their final destructive of maturity. Still in the process of development, they appear to be the minor swarm which passed through Bahrain in March, an offshoot of the immense swarms which have plagued the Middle East since first on Tunis from the Algerian Sahara in December last year.

Documents from:

Saudi Arabia: Records of the U.S. Department of State, 1955-1959 (Archives Unbound)

INCOMING TELEGRAM

Department of State

38-W

OFFICIAL USE ONLY

Action

Control: 9997

NEA

Info

FROM: Jidda

Rec'd: Apr 16, 1958  
10:41 a.m.

TO: Secretary of State

RMR

NO: 1048, April 16, 1 p.m.

SS

C

PRIORITY

RUT 145, DHAHRAN 248

Chief, USMIM:

farm was passed  
agency am trans-

ADM

AIR

Al Hasa area will be destroyed by locust. This includes the Riyadh and Al Kharj area where His Majesty's farms are located. It is the heaviest concentration in at least 14 years. Locust Control Mission ineffective and without insecticides. Is it possible for the American Government to render assistance? We are studying means of helping with local resources but this looks like big job which will require US Government assistance presumably through CINCUSAFE. Believe this is a good opportunity if we move fast.

END QUOTE

Am advising Foreign Office foregoing. Recommend US not (repeat not) take action until requested by SAG, but would appreciate indication from Department what USG prepared to do.

Dhahran notify Logan through Clark that he should explain SAG must make request.

THIS LOCUST... 38 RETURNED  
TO THE  
CENTRAL FILES

886A.22/4-1658

HBS

OFFICIAL USE



Studying Primary Sources

An Opportunity for your Institution...

# Resources for Eastern Europe

## **ARCHIVES UNBOUND**

- Poland: Records of the U.S. Dept. of State 1945-1963
- Romania: Records of the U.S. Dept. of State 1945-1963
- Socialism and National Unity in Yugoslavia, 1945-63: Records of the U.S. State Department Classified Files
- Alexander III and the Policy of "Russification," 1883-1886
- Czechoslovakia from Liberation to Communist State, 1945-63
- Correspondence from German Concentration Camps and Prisons

# Gale Primary Sources



The Economic  
Historical Archive,  
1843-2011

19th Century UK  
Periodicals

Punch Historical  
Archive, 1841-1992

American Civil Liberties  
Union Papers, 1912-  
1990

Portuguese History and  
Culture

Indigenous Peoples:  
North America

(47)

# Digital Humanities

# What is Text and Data Mining?

- “Automated processing of large amounts of structured digital textual content, for purposes of information retrieval, extraction, interpretation, and analysis.”

Bernie Reilly, President of the Center for Research Libraries

- To Gale, text and data mining represents a new opportunity to lead the way in making content highly useful for scholarly research and to add value to our historical archives
- Gale aim to facilitate faculty and library connections, take part in new research across the humanities and social science disciplines, and position the library as the centre of research.



# Why is it Important to Researchers?

- An almost entirely new front for investigation and analysis
- Allows researchers to enrich content through improved indexing, perform a systematic review of literature, create new databases that themselves can be mined, and do research on mining itself for computational linguistics research

Increasingly, scholars want to do more than just search and view special collections. They want to analyse them in large numbers using computer programmes, to find undiscovered patterns, and to ask new questions

# Beyond 'Search and Retrieve'

## Adam Crymble Digital humanities lab

About  
Research  
Teaching  
Crymble Awards  
Contact

Digital History takes a twenty-first century approach to understanding and sharing the past. The Digital History Lab of Adam Crymble explores ways that we can analyse large collections of historical sources in ways that would not be possible using traditional techniques.

Find out more about my research, and how we can work together to explore history.



Dr Adam Crymble (University of Hertfordshire) is a pioneer of this. He is interested in issues around how immigrants have been treated in Britain historically

My Research looks at  
a pretty typical  
**Historical  
Question:**

How were Irish Immigrants  
to London, England  
treated at the Dawn of the

**Industrial  
Revolution?**



d Kingdom—Conservatives  
eral Unionists support (or side  
es, and oppose **Irish** Home  
(or **Irish** members strugglin  
) , and the "Little England  
Imperial[istic] idea, or the  
ire) are the enemies of th

James Joyce, in an address  
chief element to be surpris  
y **Londoner** must be ex  
with him the sensation does  
eedily becomes amazement  
of the comparative type, a

348 pp., 1 facsim. 4". (Chronicles a  
of Great Britain, no. 48.)

**Weird tales.** **Irish.** **London:** H  
[1890?]. 256 pp. 18".

Selections from Carleton, Lover, Lever, Cro

**Wilde** (*Lady* [Jane Francesca  
Ancient legends, mystic charms and  
of Ireland. With sketches of the **Irish**

by **Irish** writers.  
ditions are given  
are those referred

BANIM, JOHN.  
Series. **Lon**  
Nowlans" al



**DUBLIN.**  
1845— Camb.U.I.; Dub.R.I.A.; Dub.T.C.; Gl  
The **Irish** Naturalist: A Monthly Journal of C  
History. Dublin, Belfast, **London**.  
1899— R.M.; Camb.P.S.C.; Camb.U.; Dub.

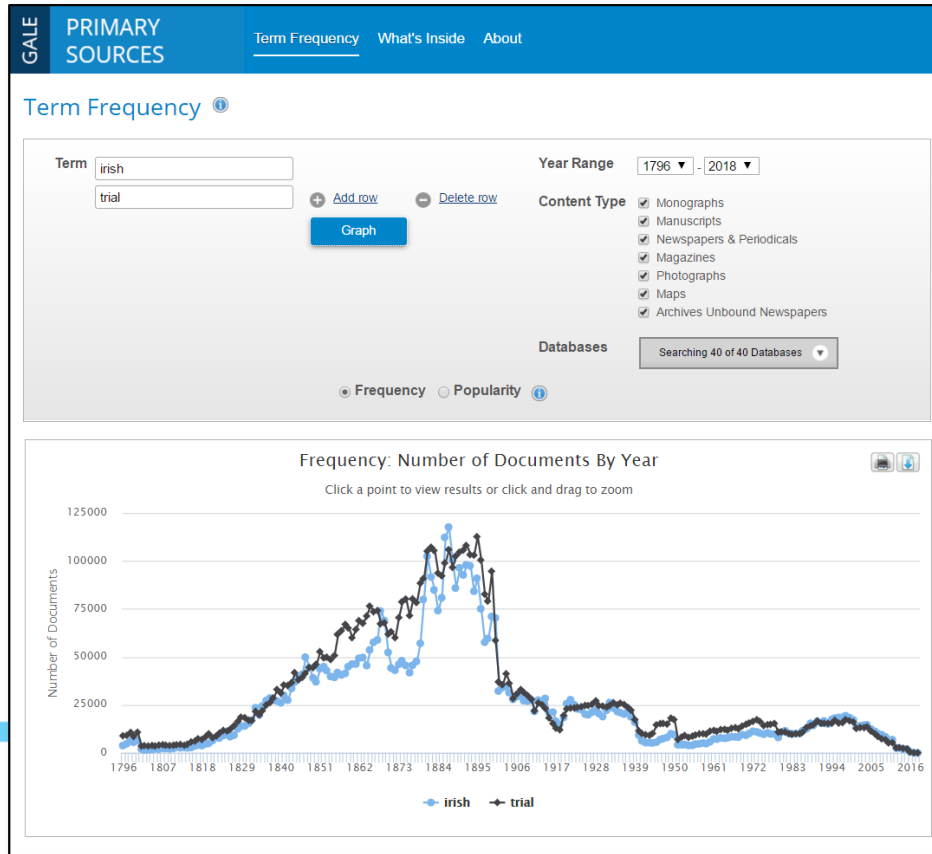




FIG 3. Irish persons were four times more likely to appear in court on trial for their lives.



# Using GPS to Analyse Texts



# Text Mining In Action

101	it again for himself: which he did at once. The	Jew	glanced round the room, as his companion tossed down the	Oliver Twist
102	anything to anybody here.' 'No more it is,' replied the	Jew;	'Charley will talk. Don't mind him, my dear; don't mind	Oliver Twist
103	the right shoulder; a piece of dumb show which the	Jew	appeared to understand perfectly. He then, in cant terms, with	Oliver Twist
104	what more could you have?" There was a red-eyed little	Jew	who came into the Close while I was loitering there,	Great Expect...
105	I was loitering there, in company with a second little	Jew	whom he sent upon an errand; and while the messenger	Great Expect...
106	errand; and while the messenger was gone, I remarked this	Jew,	who was of a highly excitable temperament, performing a jig	Great Expect...
107	ain't asleep. Catch a weasel at it, and catch a	Jew!	'Truly, sir, I fear I nodded,' said the old man.	Our Mutual F...
108	for?' 'Only to show you my meaning clearly,' said the	Jew,	raising his eyebrows. 'To be able to do that, you	Oliver Twist
109	a short pause. 'Yes, sir,' replied Oliver. 'Ah!' said the	Jew,	turning rather pale. 'They-- --they're mine, Oliver; my little property.	Oliver Twist
110				Twist
111				Twist
112				Twist
113				utual F...

Here we can see that Dickens used anti-Semitic language in his writing. This was typical for his day, but not something people generally notice when reading his novels



Dr Michaela  
Mahlberg,  
University of  
Nottingham

# What is being provided, and how?

- Gale will deliver a hard drive with the XML from each of the resources that the library requests (and for which the corresponding archive purchase has been made)
- The XML will differ depending on the resource (platform, date created, material type, etc.).
- Any resource that is cross-searchable on *Gale Primary Sources* will have the same data structure.
- The XML content is the same as the OCR text viewable in the interface as well as accompanying metadata
- The library will be responsible for getting their students/faculty access to the content after Gale has delivered the drive.

# How can the data be used, and by whom?

- The library will be responsible for managing access to the data
- The data can be used by any faculty, student, librarian, or scholar that would have access under the original license agreement signed
- The terms and conditions of use are covered in the licensing addendum
- The data can be used in myriad ways. It is all up to the researchers
- Researchers do not have to let Gale know when they are publishing research based on the data but we would love to know what work is being done!

**Gale Primary Sources**

Start at the source.

Questions?