

HERITAGE CONNECTOR: TRANSFORMING TEXT INTO DATA TO EXTRACT MEANING AND MAKE CONNECTIONS

**SCIENCE
MUSEUM
GROUP**



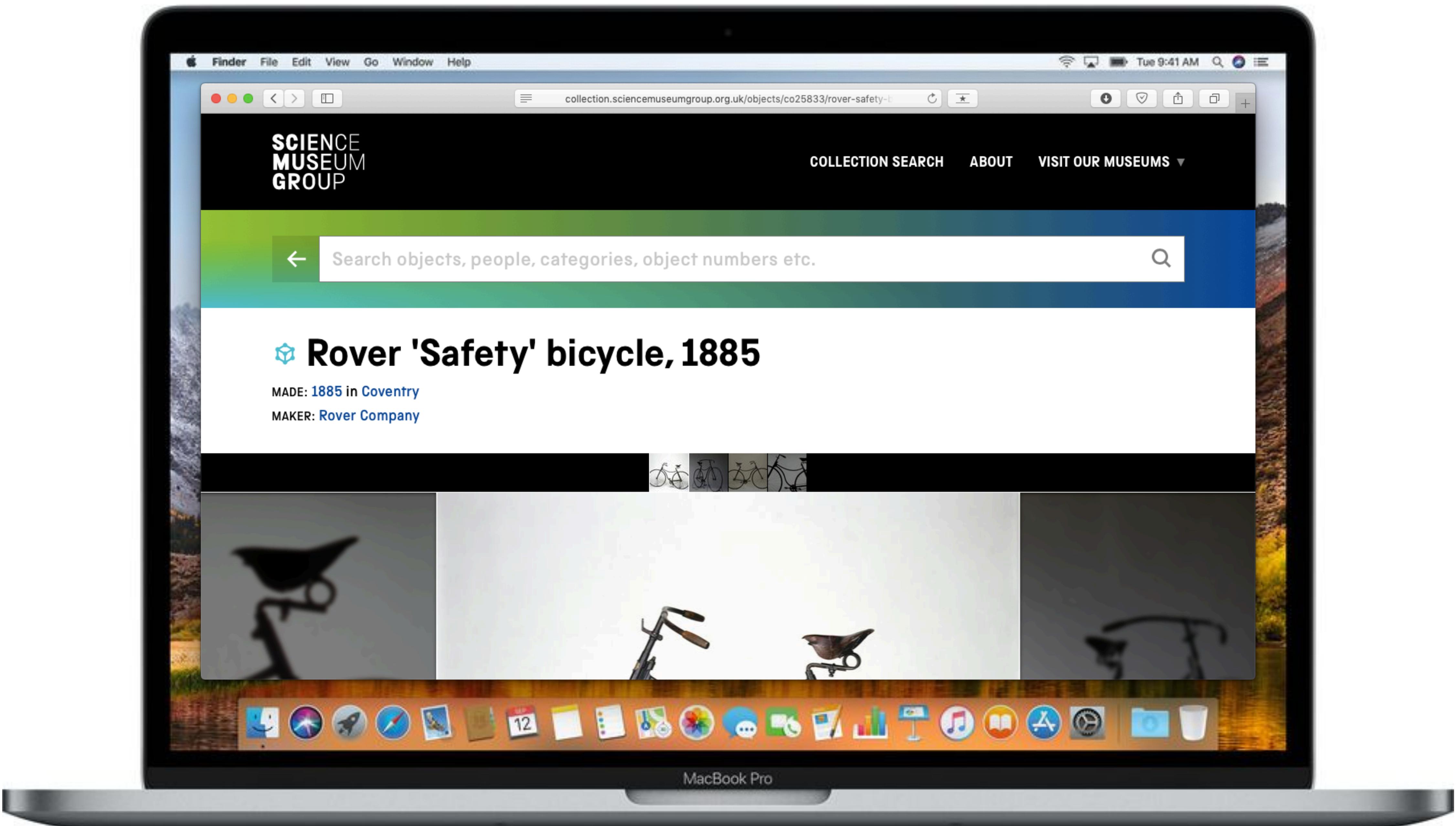
V&A

John Stack, Digital Director
Jamie Unwin, Technical Architect Collections
Rhiannon Lewis, Project Co-ordinator
TBC, Research Developer

Jane Winters, Professor of Digital Humanities & Pro-Dean for Libraries

Angela Wolff, Full Stack Developer, V&A





CORE RESEARCH QUESTION

How can existing digital tools and methods be used to build relationships at scale between poorly and inconsistently catalogued digitised collection objects and other content sources?



WIKIDATA

en.wikipedia.org/wiki/LNER_Class_A3_4472_Flying_Scotsman

Not logged in Talk Contributions Create account Log in

Article Talk Read Edit View history Search Wikipedia

LNER Class A3 4472 Flying Scotsman

From Wikipedia, the free encyclopedia

This article is about the locomotive. For the train service, see [Flying Scotsman \(train\)](#). For other uses, see [Flying Scotsman \(disambiguation\)](#).

LNER Class A3 4472 Flying Scotsman is a Pacific steam locomotive built in 1923 for the [London and North Eastern Railway](#) (LNER) at [Doncaster Works](#) to a design of [Nigel Gresley](#). It was employed on long-distance express [East Coast Main Line](#) trains by the LNER and its successors, [British Railways](#) Eastern and North-Eastern Regions, notably on the [London](#) to [Edinburgh](#) [Flying Scotsman](#) train service after which it was named.

The locomotive set two world records for steam traction, becoming the first steam locomotive to be officially authenticated at reaching 100 miles per hour (160.9 km/h) on 30 November 1934.^[1] and then setting a record for the longest non-stop run by a steam locomotive when it ran 422 miles (679 km) on 8 August 1989 while in Australia.^[2]

Retired from regular service in 1963 after covering 2.08 million miles,^{[1][3][4]} [Flying Scotsman](#) enjoyed considerable fame in preservation under the ownership of, successively, [Alan Pegler](#), [William McAlpine](#), [Tony Marchington](#), and finally the [National Railway Museum](#) (NRM).

As well as hauling enthusiast specials in the United Kingdom, the locomotive toured extensively in the United States and Canada from 1969 until 1973^[5] and Australia in 1988/89.^[6] [Flying Scotsman](#) has been described as the world's most famous steam locomotive.^{[7][8]} In a 2015 poll which questioned people from four continents it was again ranked the most famous locomotive.^[9]

Contents [hide]

- 1 History
- 2 Preservation
 - 2.1 Alan Pegler
 - 2.2 William McAlpine
 - 2.3 Tony Marchington
 - 2.4 National Railway Museum
 - 2.5 Overhaul 2006–2016

Type and origin [hide]

Power type Steam
Designer Sir Nigel Gresley
Builder Doncaster Works
Build date February 1923
Website www.flyingscotsman.org.uk

Specifications [hide]

Configuration:

- Whyte 4-6-2

Gauge 4 ft 8½ in (1,435 mm)
standard gauge
Driver dia. 80 in (2,032 mm) diameter



Search objects, people, categories, object numbers etc.

Marks Tey Station With Mr L Sydney Bracher

MADE: 1899-06-06 in Marks Tey Railway Station
MAKER: Edgar Tarry Adams



'Marks Tey Station With Mr. L. Sydney Bracher, 1899 A photograph of Mr L. Science Museum Group Collection © The Board of Trustees of the Science Museum

Use this image

A photograph entitled 'Marks Tey Station With Mr L. Sydney Bracher', taken by Edgar Tarry Adams in June 1899. Mr Bracher, probably a friend of Adams, is captured smoking his pipe and tipping his straw boater at the camera at the station in Essex.

A photograph entitled 'Marks Tey Station With Mr L Sydney Bracher', taken by Edgar Tarry Adams in June 1899.

Mr Bracher, probably a friend of Adams, is captured smoking his pipe and tipping his straw boater at the camera at the station in Essex.

DETAILS

CATEGORY:	Photographs
OBJECT NUMBER:	1997-5003/157
TYPE:	photograph
TAXONOMY:	visual and verbal communication

CITE THIS PAGE

Science Museum Group. Marks Tey Station

RIGHTS

We encourage the use and reuse of our collection data.



Data in the title, made, maker and details fields are released under Creative Commons

Search objects, people, categories, object numbers etc. 

_marks Tey Station With Mr L Sydney Bracher

MADE: 1899-06-06 in Marks Tey Railway Station
MAKER: Edgar Tarry Adams



'Marks Tey Station With Mr. L. Sydney Bracher, 1899 A photograph of Mr L. Science Museum Group Collection © The Board of Trustees of the Science Museum

< > Use this image

A photograph entitled 'Marks Tey Station With Mr L. Sydney Bracher', taken by Edgar Tarry Adams in June 1899. Mr Bracher, probably a friend of Adams, is captured smoking his pipe and tipping his straw boater at the camera at the station in Essex.

A photograph entitled 'Marks Tey Station With Mr L Sydney Bracher', taken by Edgar Tarry Adams in June 1899.

Mr Bracher, probably a friend of Adams, is captured smoking his pipe and tipping his straw boater at the camera at the station in Essex.

DETAILS

CATEGORY: [Photographs](#)

OBJECT NUMBER: 1997-5003/157

TYPE: photograph

TAXONOMY: visual and verbal communication

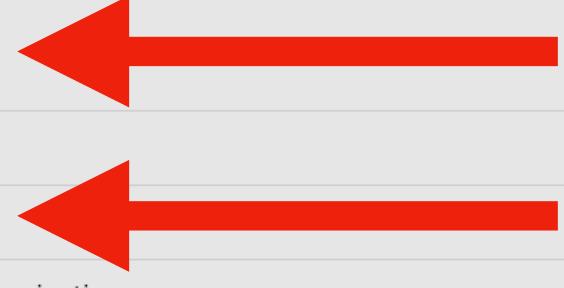
CITE THIS PAGE

Science Museum Group. Marks Tey Station

RIGHTS

We encourage the use and reuse of our collection data.

Data in the title, made, maker and details fields are released under Creative Commons Zero



Search objects, people, categories, object numbers etc. 

Marks Tey Station With Mr L Sydney Bracher

MADE: 1899-06-06 in Marks Tey Railway Station
MAKER: Edgar Tarry Adams



'Marks Tey Station With Mr. L. Sydney Bracher, 1899 A photograph of Mr L. Science Museum Group Collection
© The Board of Trustees of the Science Museum

A photograph entitled 'Marks Tey Station With Mr L. Sydney Bracher', taken by Edgar Tarry Adams in June 1899. Mr Bracher, probably a friend of Adams, is captured smoking his pipe and tipping his straw boater at the camera at the station in Essex.

A photograph entitled 'Marks Tey Station With Mr L Sydney Bracher', taken by Edgar Tarry Adams in June 1899.

Mr Bracher, probably a friend of Adams, is captured smoking his pipe and tipping his straw boater at the camera at the station in Essex.

DETAILS

CATEGORY: [Photographs](#)

OBJECT NUMBER: 1997-5003/157

TYPE: photograph

TAXONOMY: visual and verbal communication

CITE THIS PAGE

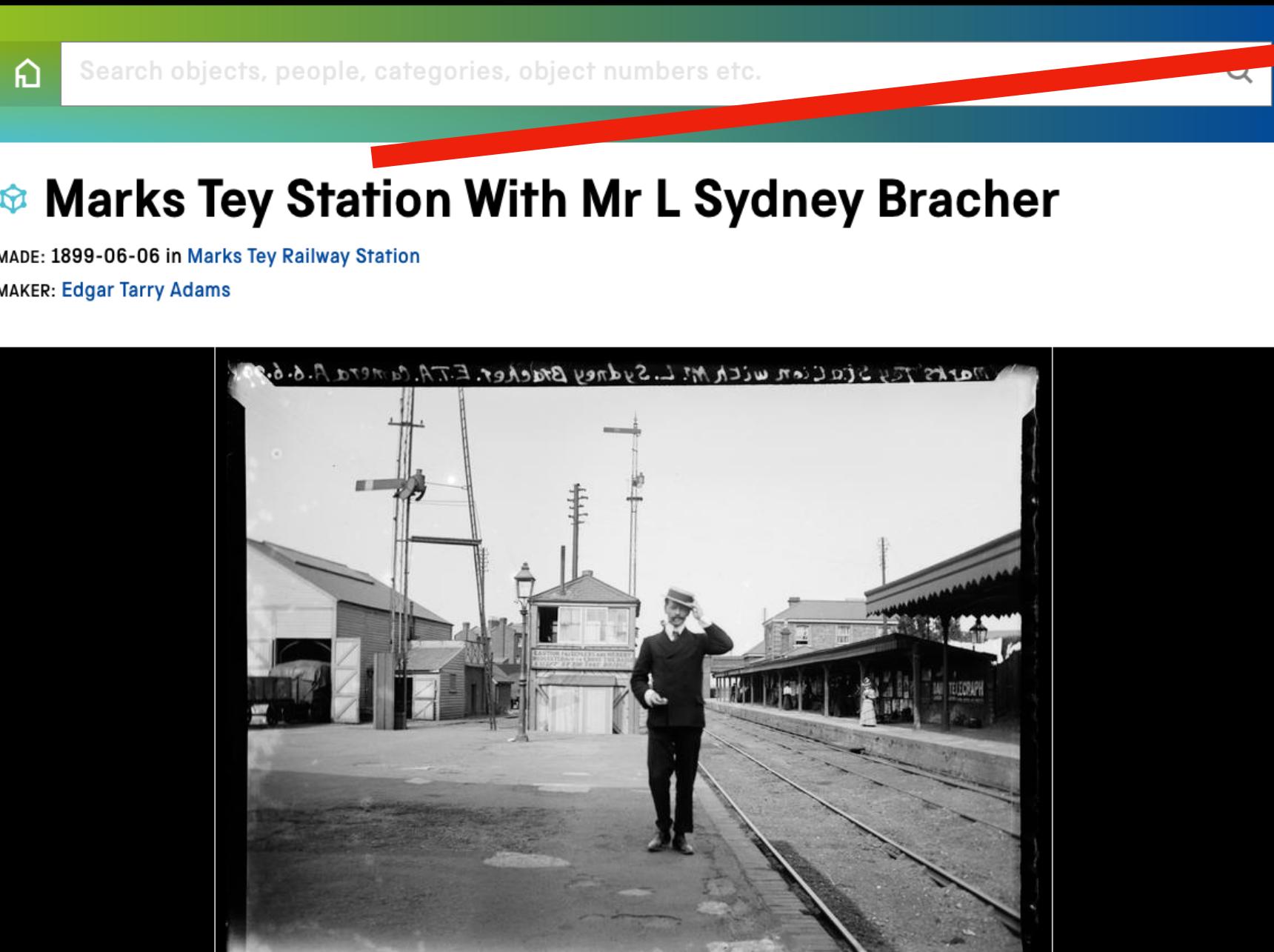
Science Museum Group. Marks Tey Station 

RIGHTS

We encourage the use and reuse of our collection data.

 CC BY-NC-ND

The title, made, maker and details fields are released under Creative Commons



'Marks Tey Station With Mr. L. Sydney Bracher, 1899 A photograph of Mr L. Science Museum Group Collection
© The Board of Trustees of the Science Museum

A photograph entitled 'Marks Tey Station With Mr L. Sydney Bracher', taken by Edgar Tarry Adams in June 1899. Mr Bracher, probably a friend of Adams, is captured smoking his pipe and tipping his straw boater at the camera at the station in Essex.

A photograph entitled 'Marks Tey Station With Mr L Sydney Bracher', taken by Edgar Tarry Adams in June 1899.

Mr Bracher, probably a friend of Adams, is captured smoking his pipe and tipping his straw boater at the camera at the station in Essex.

DETAILS

CATEGORY: Photographs

OBJECT NUMBER: 1997-5003/157

TYPE: photograph

TAXONOMY: visual and verbal communication

CITE THIS PAGE

Science Museum Group. Marks Tey Station

RIGHTS

We encourage the use and reuse of our collection data.

Data in the title, made, maker and details fields are released under Creative Commons



Item Discussion

English Not logged in Talk Contributions Create account Log in

Read View history Search Wikidata

Marks Tey railway station (Q2025932)

No description defined

In more languages

Statements

instance of	railway station	edit
	▼ 0 references	+ add reference
		+ add value

image		edit
	Markstey1.jpg	648 × 429; 98 KB
	► 1 reference	+ add value

named after	Marks Tey	edit
	▼ 0 references	+ add reference
		+ add value

country	United Kingdom	edit
	▼ 0 references	+ add reference
		+ add value

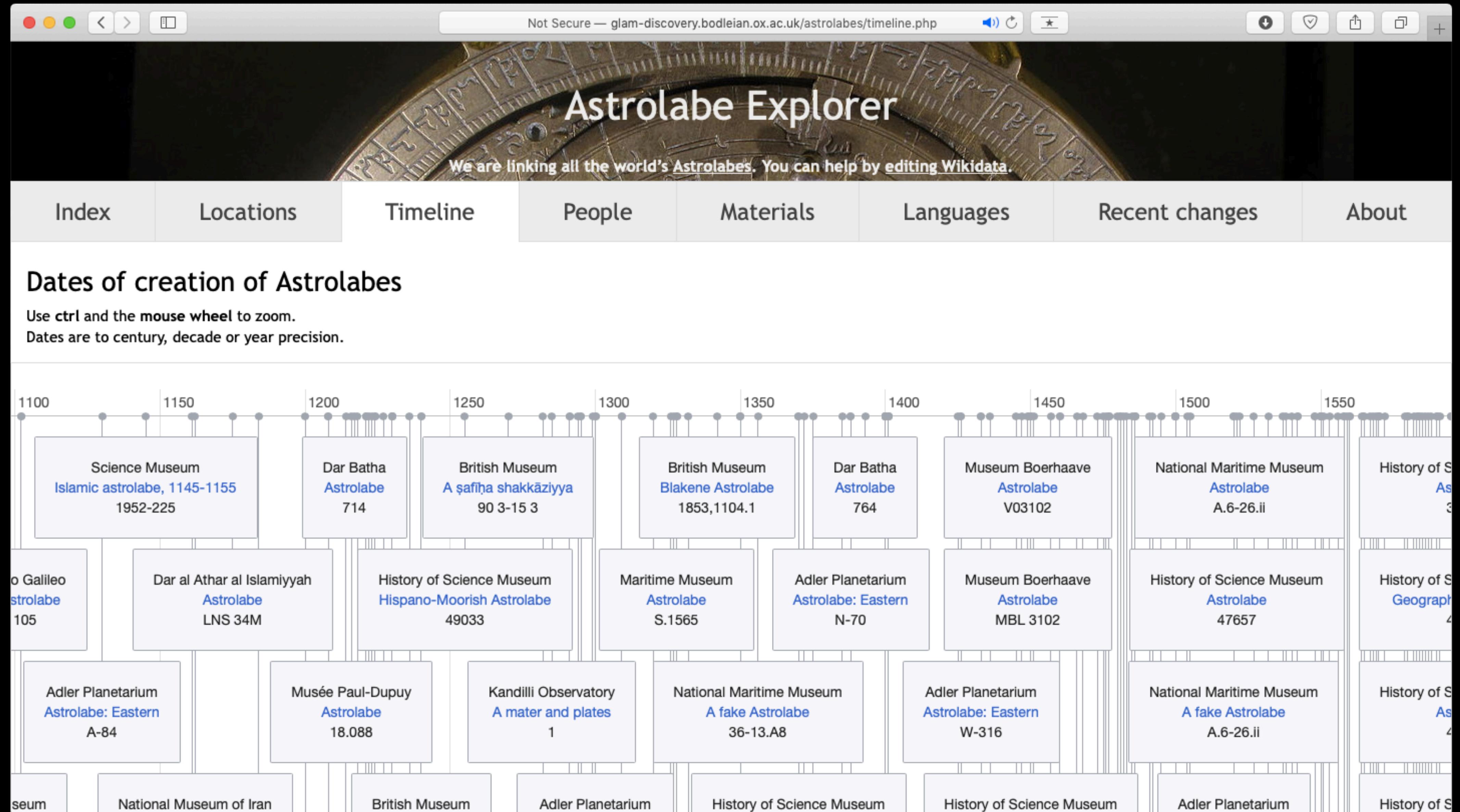
located in the administrative territorial entity	Marks Tey	edit
	► 1 reference	+ add value

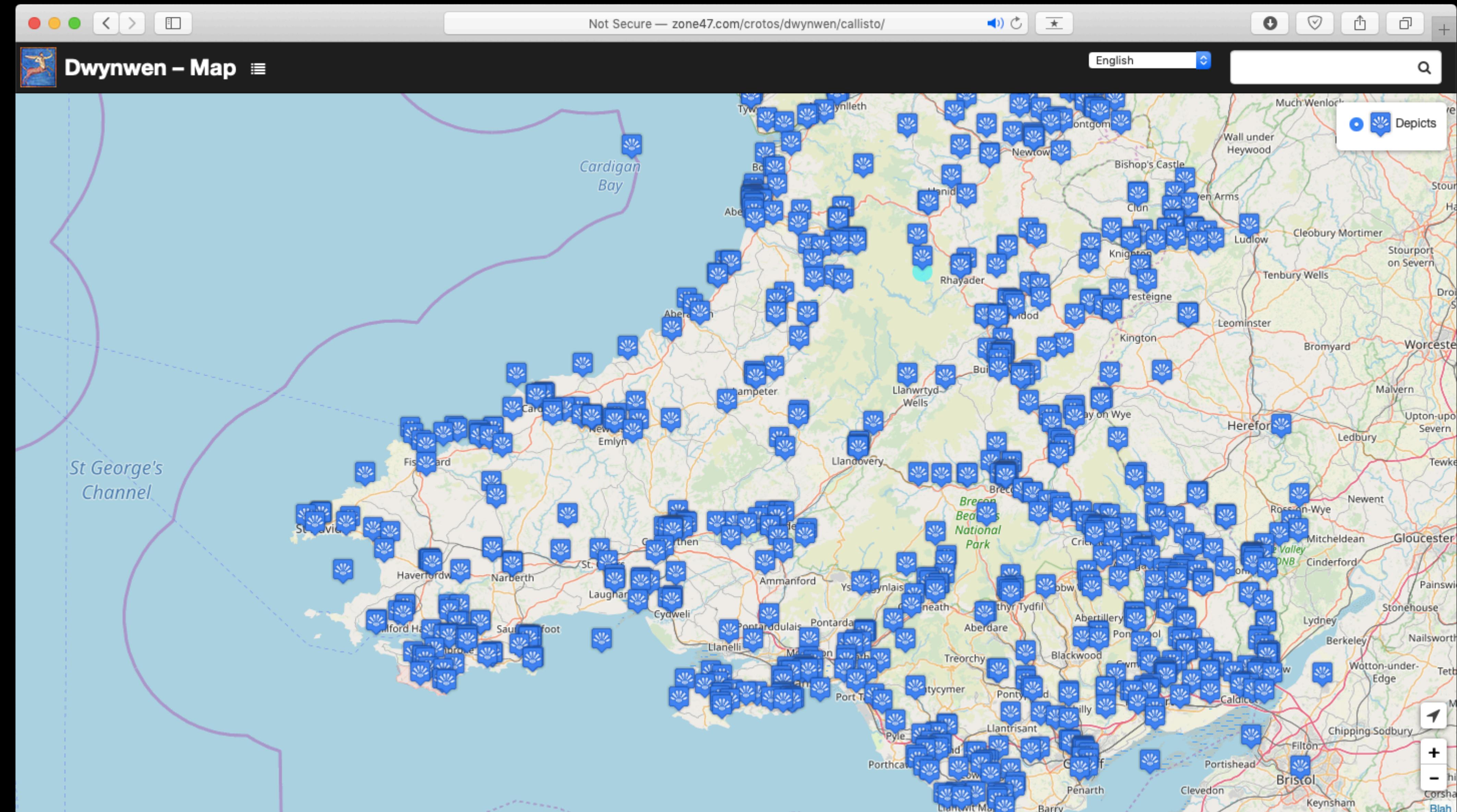
coordinate location		edit
	51°52'51"N, 0°46'58"E	Wikimedia maps Map data © OpenStreet...
	► 1 reference	+ add value

operator	Greater Anglia	edit



VIAF ID		French Sculpture	Photographers' Identities Catalog ID	Great Russian Encyclopedia Online ID	GTAA ID	Japan Search name	Semantic Scholar
15873		Gran Enclopèdia Catalana ID	Census artist ID picasso-pablo	906	3138704	artist-info artist ID 26937	author ID 69470173
GND ID	0050796	National Gallery of Victoria artist ID	elFilm person ID 312623	Finnish National Gallery artist ID 9BF45A03-	US National Archives Identifier 10580567	Open Media Database person ID 209065	パブロ・ピカソ Rijksmuseum
ISNI	0000 0001 2275 5381	OpenPlaques subject ID	1167	NYT topic ID	F8AF-49C5-	BHCL ID 275410	Research Library authority ID 90063
Library of Congress authority ID	nm0681444	Thyssen-Bornemisza artist ID	person/pablo-picasso 455	A3EE-66318C675CF0	Luminous-Lint ID Pablo_Picasso	Artyclopedia artist ID picasso_pablo	Comic Vine ID 4005-17471 NCL ID 000093750
n78086005		Munzinger person ID	FAST ID 28123	NE.se ID	Boijmans artist ID 14334	CONOR ID 8036195	Edvard Munch's correspondance person ID 911
RKDartists ID	00000001291	Perlentaucher ID	Nationalmuseum Sweden artist ID	Theatricalia person ID	Internet Broadway Database person ID 25216	Runeberg author ID picasbab	Art Gallery of South Australia creator ID 3504
63276	pablo-picasso	4044	425f	PictoRight ID code M-006359	497091	DACS ID 35fb9c7-d2bb-4cc1-afdd-26b1c342498d	abART person ID 7762
IdRef ID	027071790	NORAF ID	Encyclopædia Britannica Online ID biography/Pablo-Picasso	Auckland Art Gallery artist ID	PM20 folder ID pe/013593	Benetiz ID B00141033	Rotterdam City Archives actor ID 2519257
00452768	90058585	DBNL author ID	Les Archives du Spectacle Person ID	1027	Cooper-Hewitt Person ID	Angelicum author ID 82807	National Diet Library Auth ID 000105563
NKCR AUT ID	jn19990006541	NNDB people ID	373/000024301	National Library of Greece ID	Enciclopédia Itaú Cultural ID	Libris-URI wt79b7cf3r8x8lx	Brockhaus Enzyklopädie online ID picasso-pablo
SELIBR ID	211453	British Museum person-institution	22735	63112	18047293	BanQ author ID 0000007686	CALIS ID n2004366423
Biblioteca Nacional de España ID	41931	Filmportal ID	ad3d6e96b6c74cfa9df316fd239187dd	MNCARS artist ID	ISFDB author ID 26270	MNAV artist ID 354	National Library of Korea Identifier KAC199621664
XX1637941	National Portrait Gallery (London) person ID	SNAC Ark ID	w6g1618s	picaso-pablo	CiNii author ID (books)	Project Gutenberg author ID 49945	Store norske leksikon ID Pablo_Picasso
Libraries Australia ID	mp62055	Tate artist ID	1767	IPI base code I-001068130-6	DA00376731	BNC ID 000037089	354
35422161		BiblioNet author ID	11685	Isidore ID	Unz Review author ID	MusicBrainz artist ID 9d5f2c22-e2d2-40be-9919-534437143bb3	RISM ID pe41008105
ULAN ID	500009666	SBN author ID	IT\ICCU\CFIV\051221	picaso_pablo	Christie's creator ID	MusicBrainz artist ID 9d5f2c22-e2d2-40be-9919-534437143bb3	SFMOMA artist ID pablo_picasso
Bibliothèque nationale de France ID	87305a20-73cd-45e9-8d1e-022b0b0c29e2	WikiTree person ID	Picasso-1	54716	IPI name number 00119750569	NLP ID (unique) A11808937	Europeana entity agent/base/60206
11919660r	KulturNav-ID	Kinopoisk person ID	99747	Spanish Biographical Dictionary ID 9357/pablo-ruiz-picasso	ARTIC artist ID	36198	NLP ID (PLWABN unique record) 62374
Artsy artist ID	pablo-picasso	openMLOL author ID	77098	RA Collections ID	Web umenia creator ID	9810616547105606	PUSC author ID 9810616547105606
NTA ID	068447809	AlloCiné person ID	898	8219	7821	Saint Louis Art Museum person ID	WorldCat Identities ID lccn-n78086005
Open Library ID	OL44790A	Architectuurgids architect ID	2163	Babelio author ID	Invaluable.com person ID nvddduiinb	2325	RERO ID 02-A000130420
Portuguese National Library ID	9658	AllMovie person ID	p56661	16890	Krugosvet article kultura_i_obrazovani_e/izobrazitelnoe_iskussvo/_PIKASSO_PABLO.htm	NSK ID 000003465	DBC author ID 87097968824419
Museum of Modern Art artist ID	4609	ČSFD person ID	65102	Encyclopædia Universalis ID	Minneapolis Institute of Art constituent ID 443	a11549853	ELNET ID a11549853
Freebase ID /m/060_7	National Gallery of Art artist ID	Swedish Film Database person ID	115119	pablo-picasso	Flanders Arts Institute person ID	V. B. Mirimanov	NLR ID (Russia) 778208
Kunstindeks Danmark Artist ID	1783	Elonet person ID	36439	picasso-erotique	1878168	BNB person ID PicassoPablo1881-1973	RSL ID (person) 000033155
1312	Musée d'Orsay artist ID	SHARE Catalogue author ID	36439	Städels Museum artist ID	Nelson-Atkins Museum of Art person ID	73	LNB ID 000029604
CANTIC ID a10443824	12184	PORT person ID	52093	Online Books Page author ID	6846	WikiArt ID pablo-picasso	National Library of Brazil ID 000590262
EGAXA ID	947012	NUKAT ID	1801	Picasso, Pablo, 1881-1973	BVMC person ID	Klexikon article ID Pablo_Picasso	Adv10229580
002657273	AGORHA person/institution ID	NLR ID (Romania)	112292	Artnet artist ID	3448	YCBA agent ID	ADV11654647
		000174361		pablo-picasso	55607	Pablo_Picasso	ADV11654648
				Le Delarge ID	5856	MutualArt artist ID	ADV12539089
				11390_artiste_PICAS	68EFD50CBA356F91	National Gallery of Canada artist ID	
				SO_Pablo		pablo-picasso	





Not Secure — lc-photo-expt.s3-website-us-east-1.amazonaws.com/#/

armyphoto division of the United States Army | View Wiki

1830	(549)
1820	(528)
1840	(514)
1810	(484)
1800	(383)
1930	(371)
1790	(297)
1770	(188)
1750	(187)
1780	(183)
1940	(177)
1740	(174)

Show More

P19 - Place of Birth

New York City	(482)
London	(286)
Paris	(233)
Philadelphia	(182)
Brooklyn	(127)
Boston	(122)
Chicago	(118)
Washington, D.C.	(77)
Vienna	(65)
San Francisco	(63)
Dublin	(55)
Manhattan	(50)
Baltimore	(50)
St. Louis	(48)
Rome	(46)
Pittsburgh	(46)
Los Angeles	(45)
Cincinnati	(45)
Berlin	(39)
Detroit	(38)

Show More

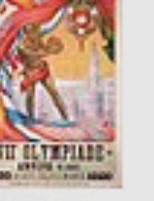
Open "<https://www.loc.gov/item/2011647755>" in a new tab

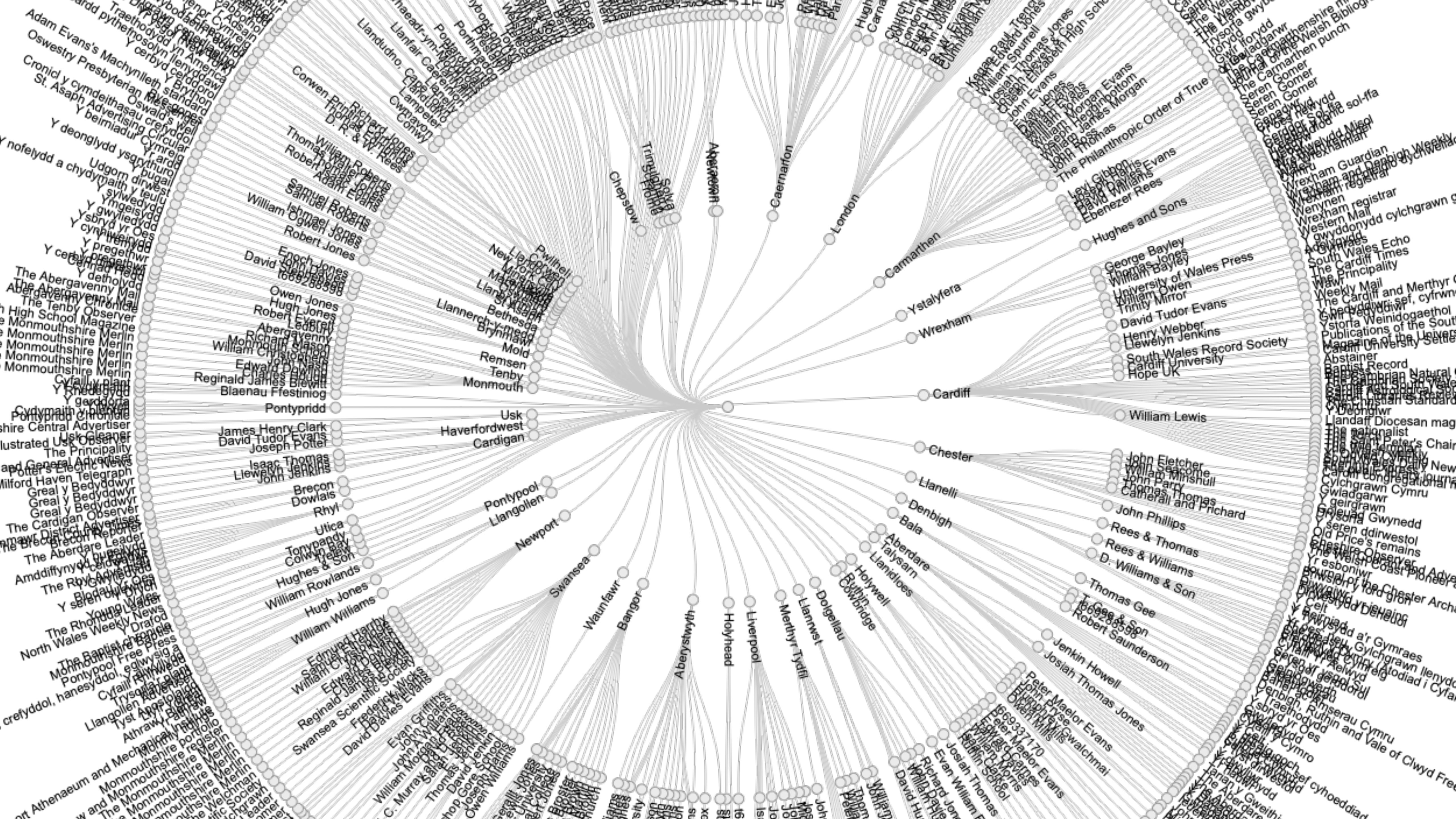


Roswell K. Bishop, Co. I, 123
Reg. N.Y.S.V. Killed at the Battle
of Chancellorsville, May 3d 1863.



headquarters of Her Majesty's Government | View Wiki


1920 Summer Olympics
games of the VII Olympiad, celebrated in Antwerp, Belgium, in 1920 | View Wiki



ACTIVITIES

1. Test and evaluate software tools for link building.
2. Hold ~~convening~~ webinar to explore potential.
3. Apply methods to V&A collection and build links.
4. Apply tools to non-collection resources an attempt link building.
5. Apply methods to handful of other collections.
6. Hold hackathon to explore working with resulting dataset of links.

THANKS

ALL RESEARCH QUESTIONS

- ★ How can existing digital tools and methods be used to build relationships at scale between poorly and inconsistently catalogued digitised collection objects and other content sources?
- ★ Which software tools and approaches (machine learning, natural language processing, topic modelling, entity extraction, fuzzy matching, named entity linking, etc.), in combination, provide the most effective approaches?
- ★ How might the relationships built using these techniques complement and amplify the benefits of the use of persistent identifiers (eg museum inventory numbers, QR codes, DOIs, etc.)?
- ★ What gaps and biases emerge when these relationships are created, and which hitherto unexpected connections are made?
- ★ How might confidence in these (machine-generated) relationships impact on their usefulness in research and discovery? Can 'degrees of confidence' become a useful concept in searching across collections as it is in statistics?
- ★ Is such an approach scalable to larger volumes of content and different types of collections?
- ★ Where is the best use of human input in supporting such an approach? What expertise and skills are required for this input?
- ★ How might these new automatically established relationships help museums to gain greater understanding of how their collections are being and could be used and interpreted?
- ★ What benefits arise from using such tools to link to scholarship such as theses and journal article abstracts/keywords? What new forms of research are afforded?
- ★ What ethical issues are raised by such an approach?
- ★ What are the most effective methods for organising interdisciplinary and cross-sectoral research projects concerned with digital cultural heritage, and how can these methods usefully be shared across the IROs?