

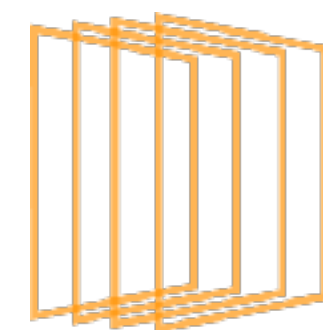
HERITAGE CONNECTOR

TRANSFORMING TEXT INTO DATA TO EXTRACT
MEANING AND MAKE CONNECTIONS

**SCIENCE
MUSEUM
GROUP**



V&A



**TOWARDS
A NATIONAL
COLLECTION**



**Arts and
Humanities
Research Council**

Kalyan Dutia, Research Developer

Rhiannon Lewis, Project Co-ordinator

John Stack, Digital Director

Jamie Unwin, Technical Architect Collections

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

Angela Wolff, Full Stack Developer, V&A

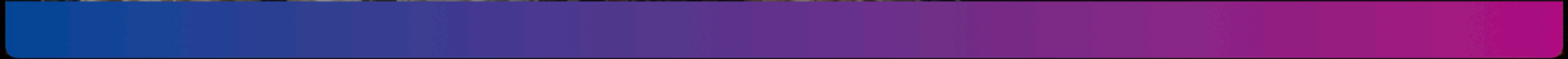
How can existing digital tools and methods be used to build relationships at scale between poorly and inconsistently catalogued digitised collection objects and other content sources?



SEARCH OUR COLLECTION

Search objects, people, categories, object numbers etc. 🔍

Explore over 350,000 objects and archives from the Science Museum, Science and Industry Museum, National Science and Media Museum, National Railway Museum and Locomotion.



All 50

People 0

Objects 50

Documents 0

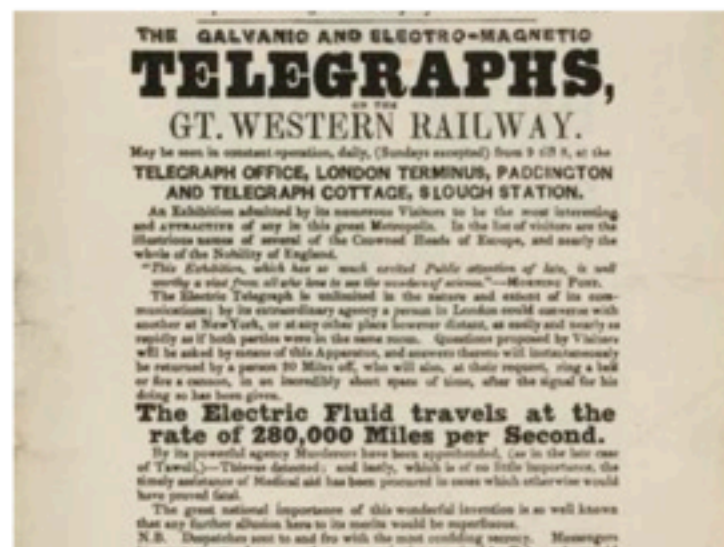
View:



Filter search



Geoffrey Tippet Collection - Swindon Works & GWR Locomotives [supplied title] Photographic Collections (Railway) c. 1930



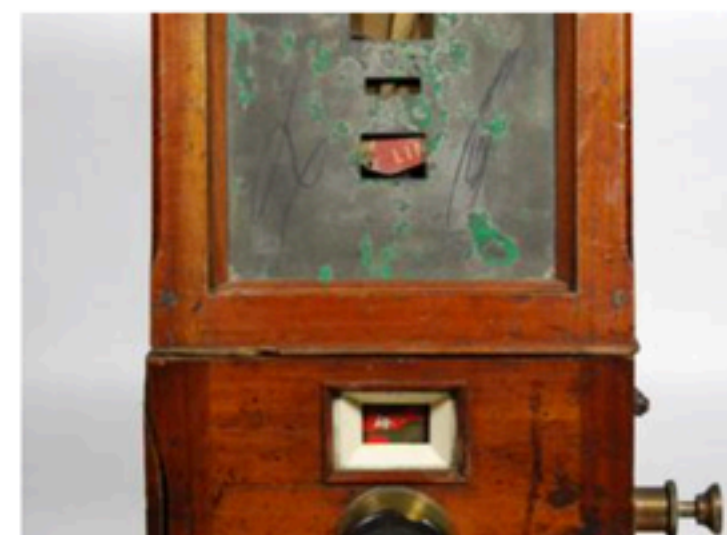
Advertisement by GWR for the exhibition of the Galvanic and Electro-Magnetic Telegraphs, GT. Western Railway Art 1845



Milk van, Great Western Railway Locomotives and Rolling Stock 1936



Diesel Railcar No 4, 1934, Great Western Railway Locomotives and Rolling Stock 1934



FinderFileEditViewGoWindowHelp

collection.sciencemuseumgroup.org.uk/search/images?q=gwr

A black and white photograph showing a steam locomotive in a workshop or factory setting, with various mechanical parts and structures visible.

Geoffrey Tippet Collection - Swindon Works & GWR Locomotives [supplied title]
Photographic Collections (Railway)
c. 1930

A historical advertisement featuring a large, ornate plate with a central emblem and text describing the exhibition of the Gal.

Advertisement by GWR for the exhibition of the Gal
Art
1845

A black and white photograph of a milk van, a specialized railway vehicle used for transporting milk, with a distinctive design and multiple compartments.

Milk van, Great Western Railway
Locomotives and Rolling Stock
1936

A color photograph of a diesel railcar, a motorized passenger vehicle, shown in a workshop or storage area.

Diesel Railcar No 4, 1934, Great Western Railway
Locomotives and Rolling Stock
1934

A white ceramic plate with a gold rim and a central emblem, displayed on a stand.

Ceramic plate, Great Western Railway - Refreshment Department
Passenger Comforts

A close-up photograph of a pocket watch with a white face, black Roman numerals, and a small seconds sub-dial.

Pocket watch, Great Western Railway
Railway Timepieces

A photograph of a telegraph instrument, a mechanical device used for sending messages over a telegraph line.

Telegraph instrument
Telecommunications

A black and white photograph of a steam locomotive, a large engine used for pulling trains, shown in a railway yard.

Steam locomotive
Locomotives and Rolling Stock

A photograph of a telegraph block instrument, a component used in telegraph systems for sending and receiving messages.

Telegraph block instrument
Telecommunications

A black and white photograph of the interior of Nottingham Victoria Joint Station, showing the platform, tracks, and the large glass and steel roof structure.

Nottingham Victoria Joint Station
Railway Infrastructure

Category

☐ Signalling & Telecommunications 11

☐ Locomotives and Rolling Stock 6

☐ Railway Posters, Notices & Handbills 5

☐ Art 4

☐ Photographic Collections (Railway) 3

☐ Pictorial Collection (Railway) 3

☐ Railway Uniform & Costume 3

☐ Locomotives and Rolling Stock Components 2

☐ Miscellanea & Curiosities 2

☐ Railway Infrastructure 2

On Display

☐ Railway Museum 10

☐ Warehouse 6

☐ Search Engine Exhibition Area 3

☐ Great Hall 1

☐ Locomotion 1

☐ The National Railway Museum at Shildon 1

Object type

☐ poster 5

☐ block instrument 4

☐ steam locomotive 4

☐ telegraph instrument 4

☐ print 3

☐ telegraph block instrument 3

DISCOVERY AND EXPLORATION

1. knowing about the collection
2. understanding that it may contain relevant contents
3. knowing that content is available online



4. navigating keyword-based search
5. grasp the relevance of the content to the search
6. navigate from that content to related content

Zeiss Compound Microscope

MADE: 1907



DETAILS


CATEGORY:	Scientific Instruments & Research
OBJECT NUMBER:	Y1991.49.2/1
TYPE:	compound microscope
TAXONOMY:	<div>furnishing and equipment<ul style="list-style-type: none">tools & equipment<ul style="list-style-type: none">optical instrument<ul style="list-style-type: none">microscopefurnishing and equipment<ul style="list-style-type: none">tools & equipment<ul style="list-style-type: none">optical instrument<ul style="list-style-type: none">microscope</div>
CREDIT:	Gift of Central Manchester Health Authority


CITE THIS PAGE

Science Museum Group. Zeiss Compound

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](#)


 Descriptions and all other text content are licensed under a [Creative Commons Attribution 4.0 licence](#)

[Using our data](#)

DOWNLOAD

Download catalogue entry as [JSON](#)

View  manifest in [IIIF viewer](#)

Add  to [Animal Crossing Art Generator](#)

Download  manifest [IIIF](#)

Our records are constantly being enhanced and improved, but please note that we cannot guarantee the accuracy of any information shown on this website.

BETA

From the Collections

From ancient Chinese ceramics to Alexander McQueen evening dresses, take an incredible journey through 5000 years of human creativity with our online collections.



EXPLORE OUR COLLECTIONS

Search more than 1.2 million objects

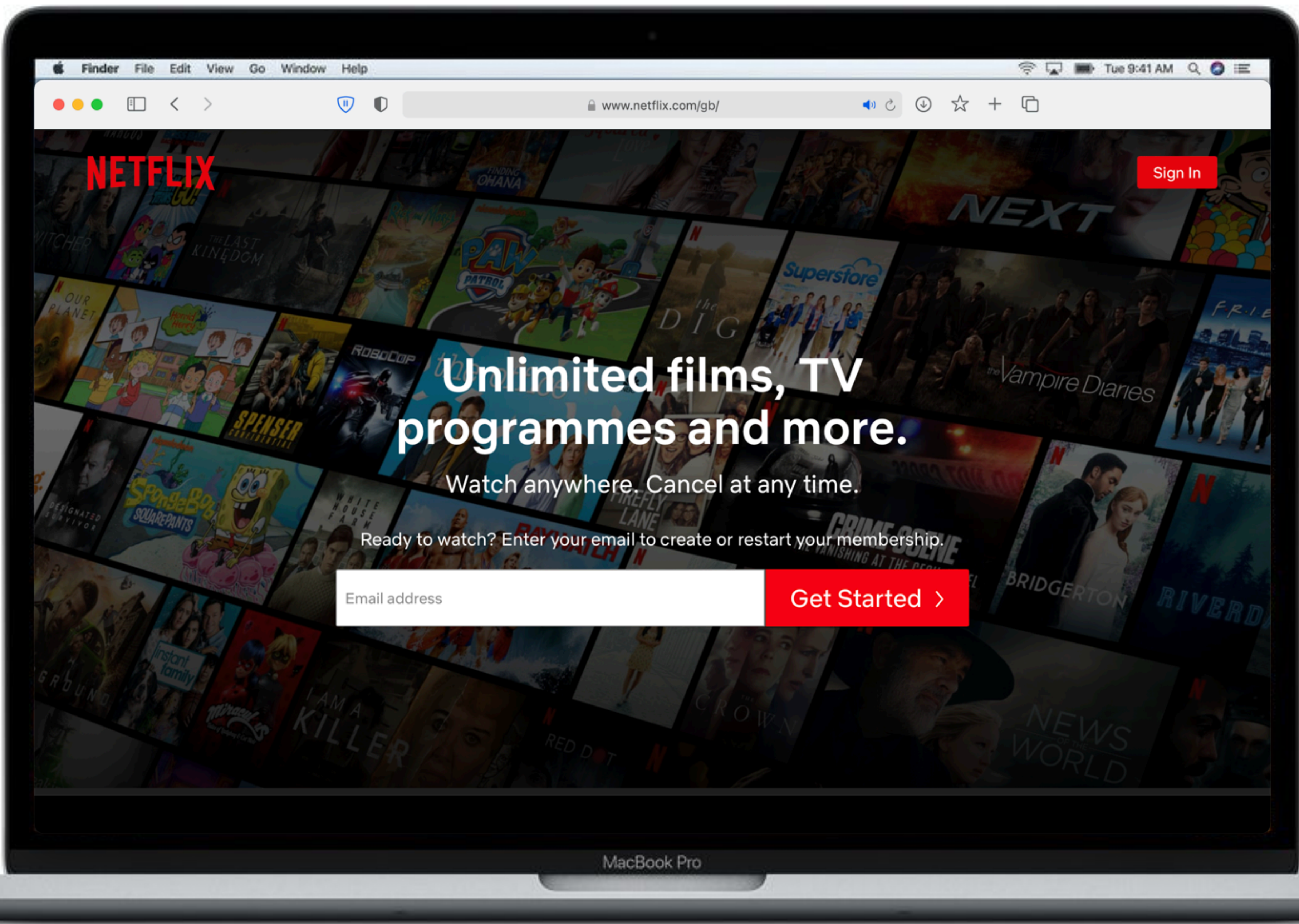
Search by object, artist, maker...

SEARCH ►

[Add dates](#) +

HERITAGE CONNECTOR PROJECT

1. Knowledge graphs
2. Linked data
3. Artificial intelligence



NETFLIX

Sign In

Unlimited films, TV programmes and more.

Watch anywhere. Cancel at any time.

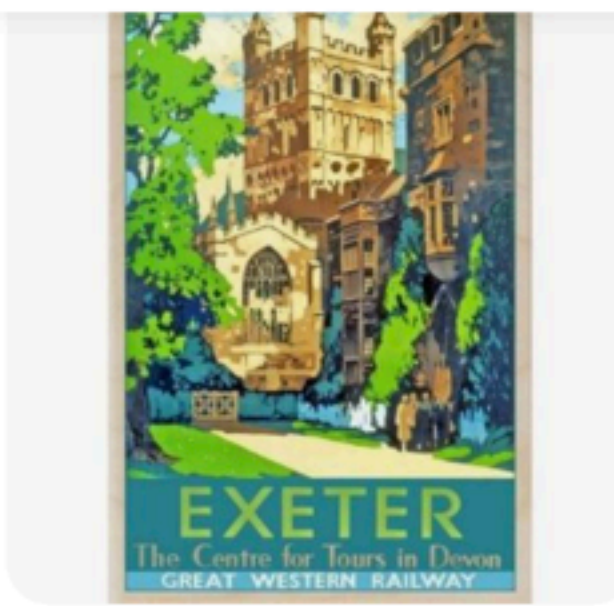
Ready to watch? Enter your email to create or restart your membership.

Email address

Get Started >

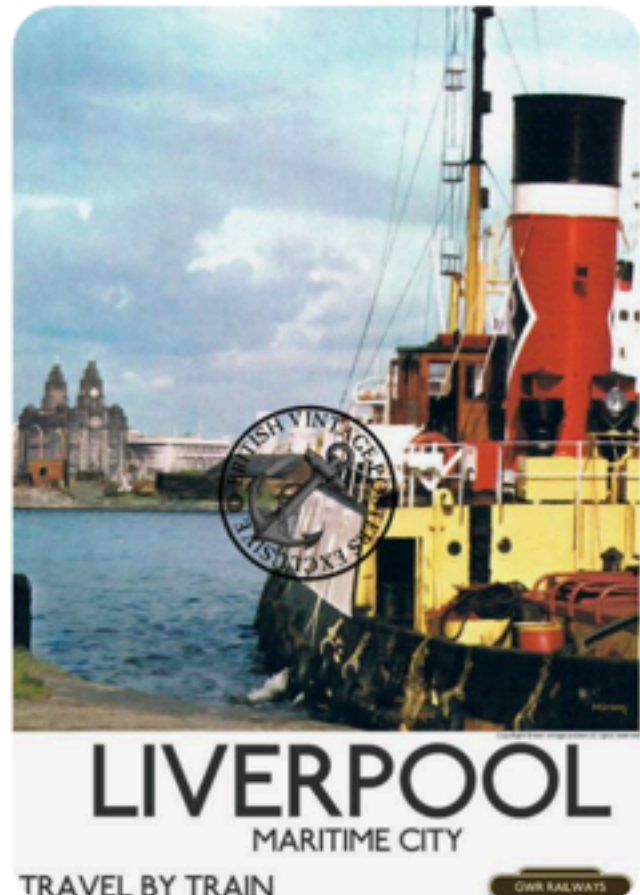
MacBook Pro

Explore Shop



The Wooden Postcard Company-Exeter Cathedral Wooden Fridge Magnet

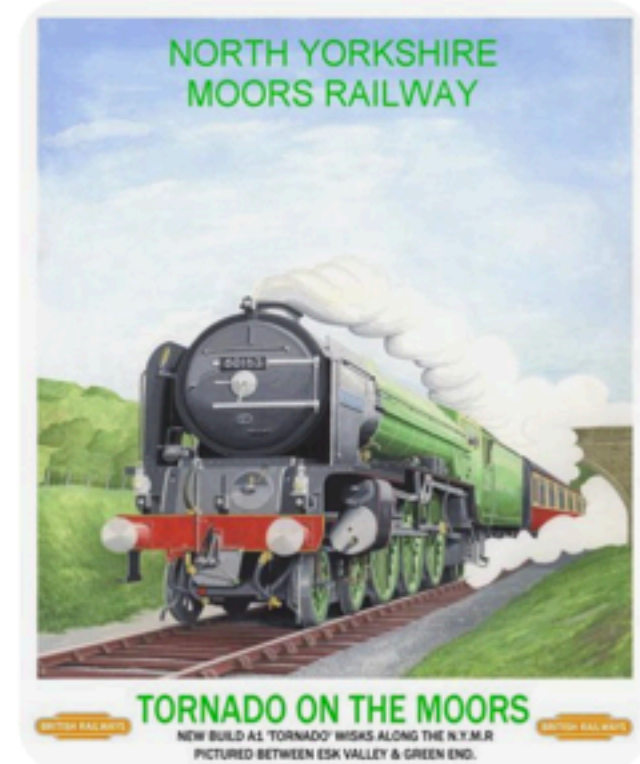
Promoted by Trouva



Details about Vintage Poster... eBay



Details about TX341 Vintage... eBay



Tornado North Yorkshire Moors... Etsy

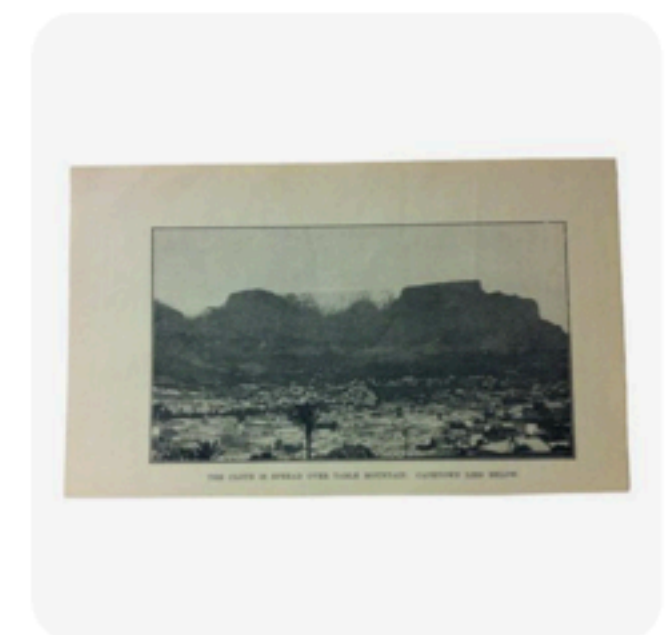
YORK THE SHAMBLES TRAVEL BY TRAIN



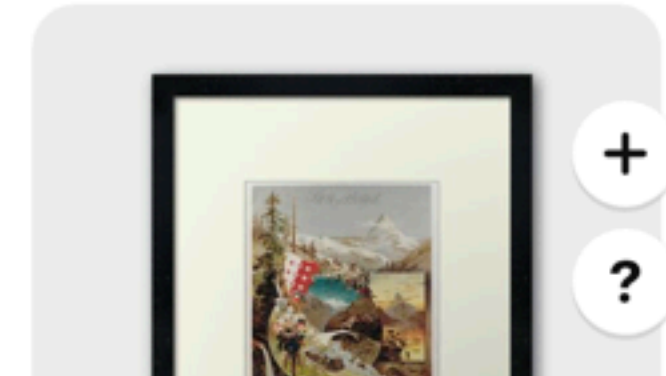
Details about VINTAGE RAILWAY... eBay



VINTAGE RAILWAY POSTER York The Shambles Old TRAIN...



1929 "The Cloth Is Spread Over Table Mountain" South African...





Start your search



Become a host



Newquay
4-hour drive



Lyme Regis
2-hour drive



Plymouth
3.5-hour drive



Windermere
5.5-hour drive

HERITAGE CONNECTOR PROJECT

1. Knowledge graphs

2. Linked data

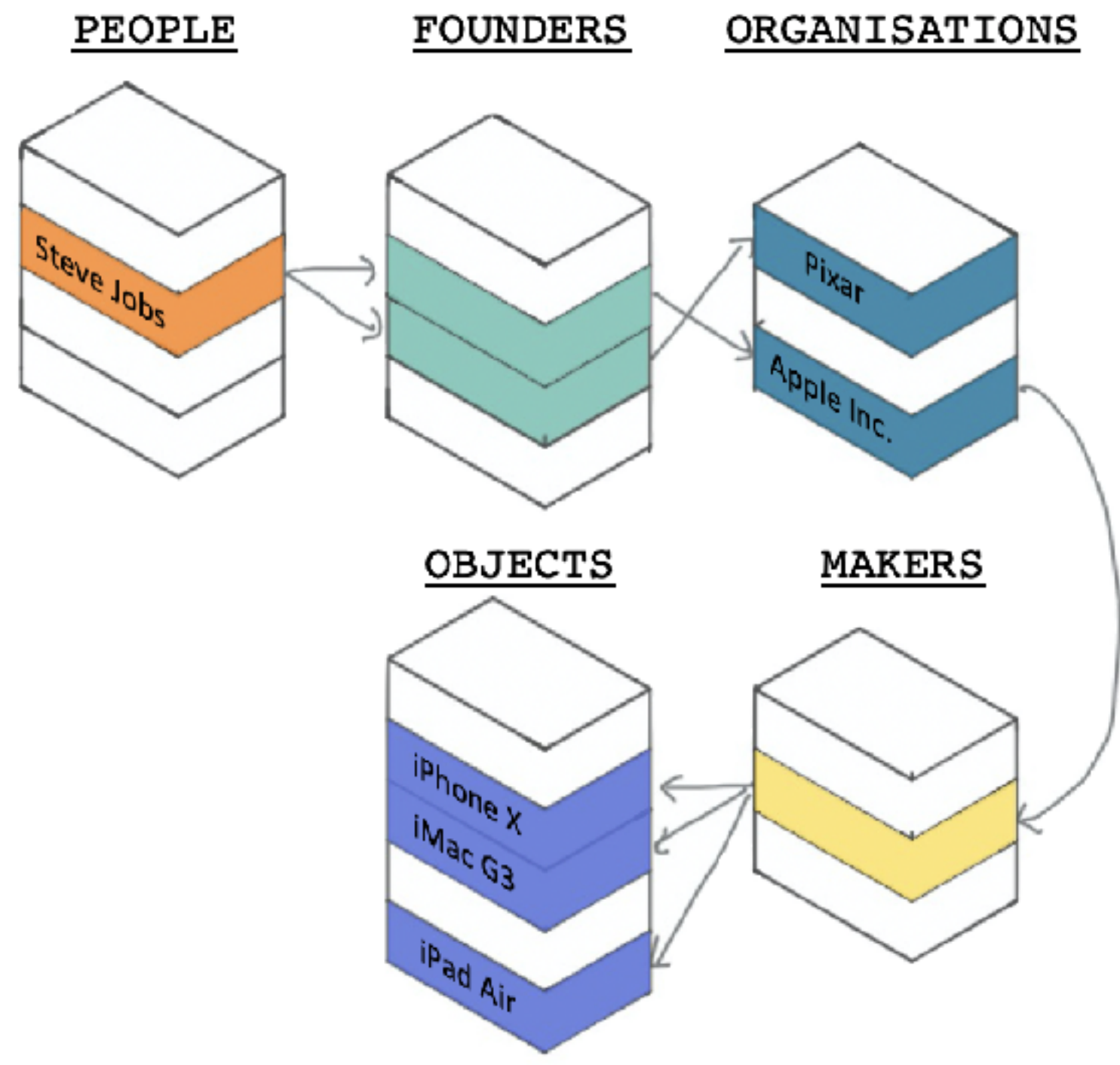
3. Artificial intelligence

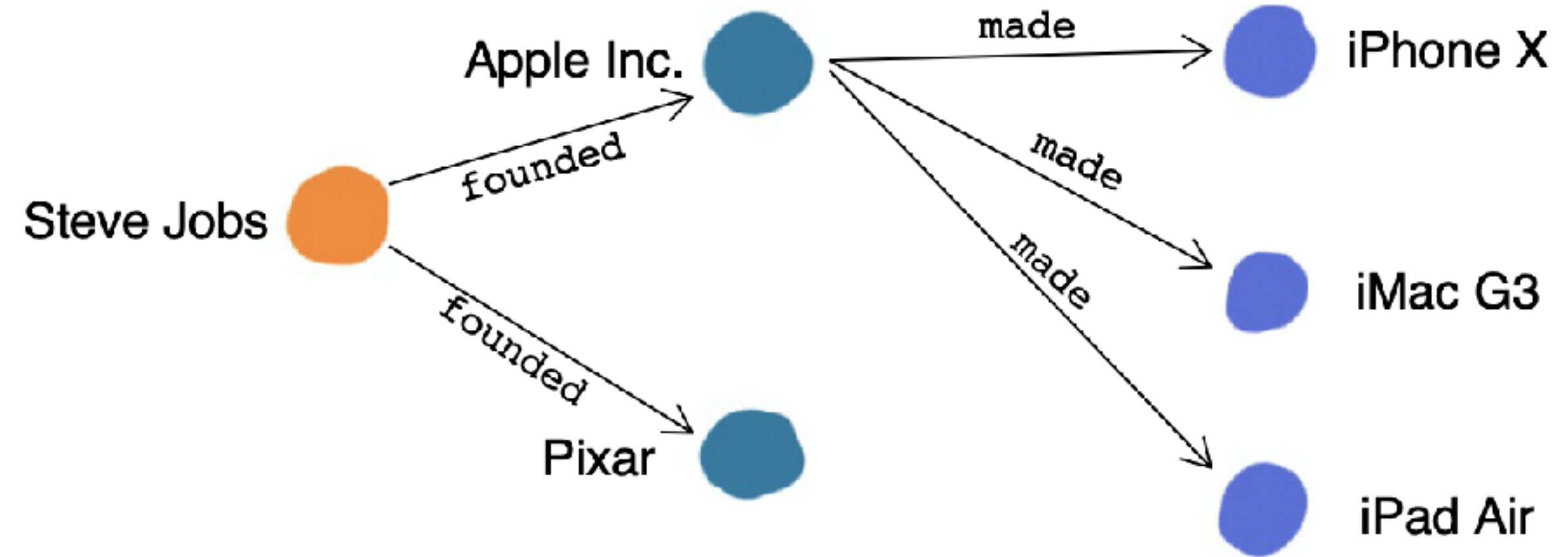
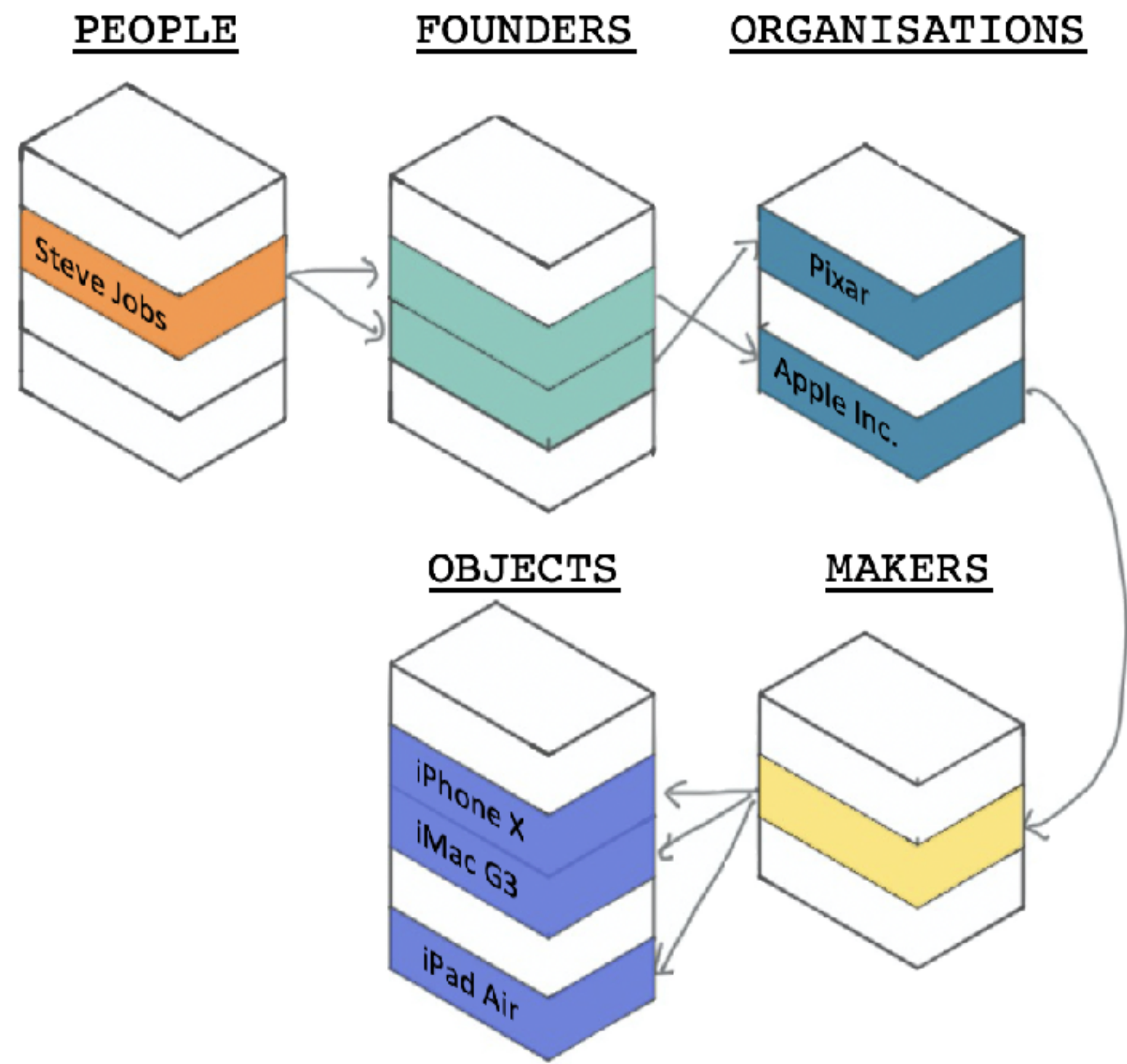
1. Improve collection interfaces

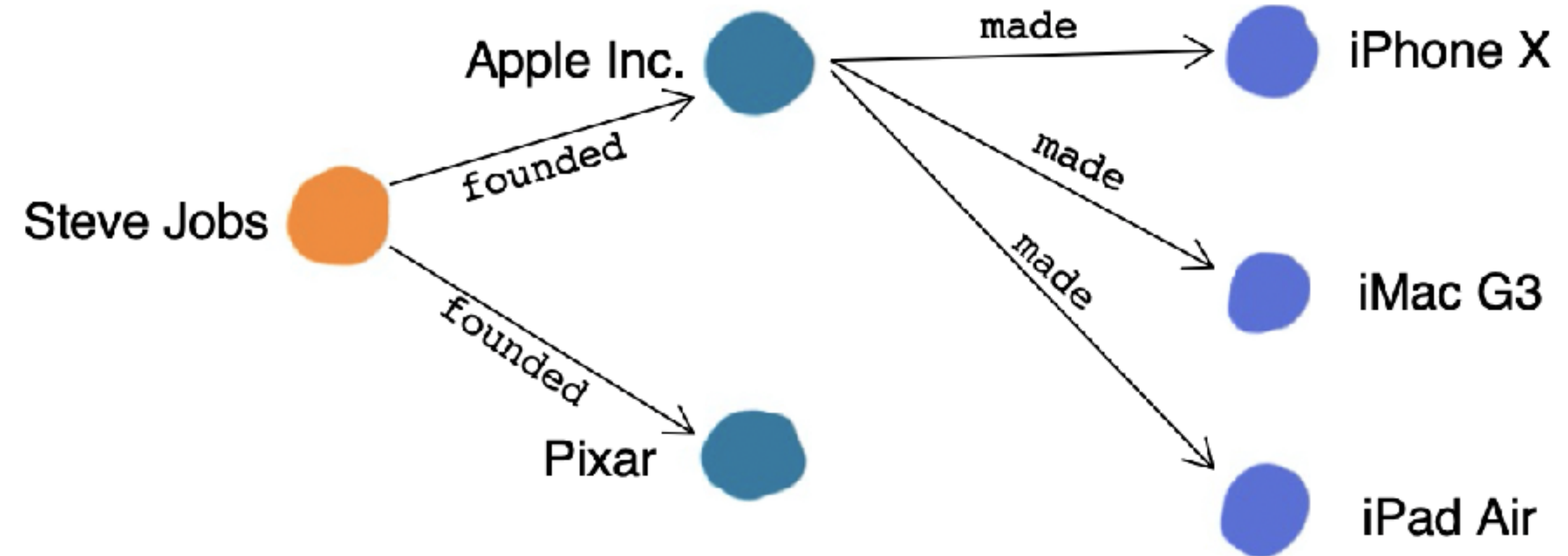
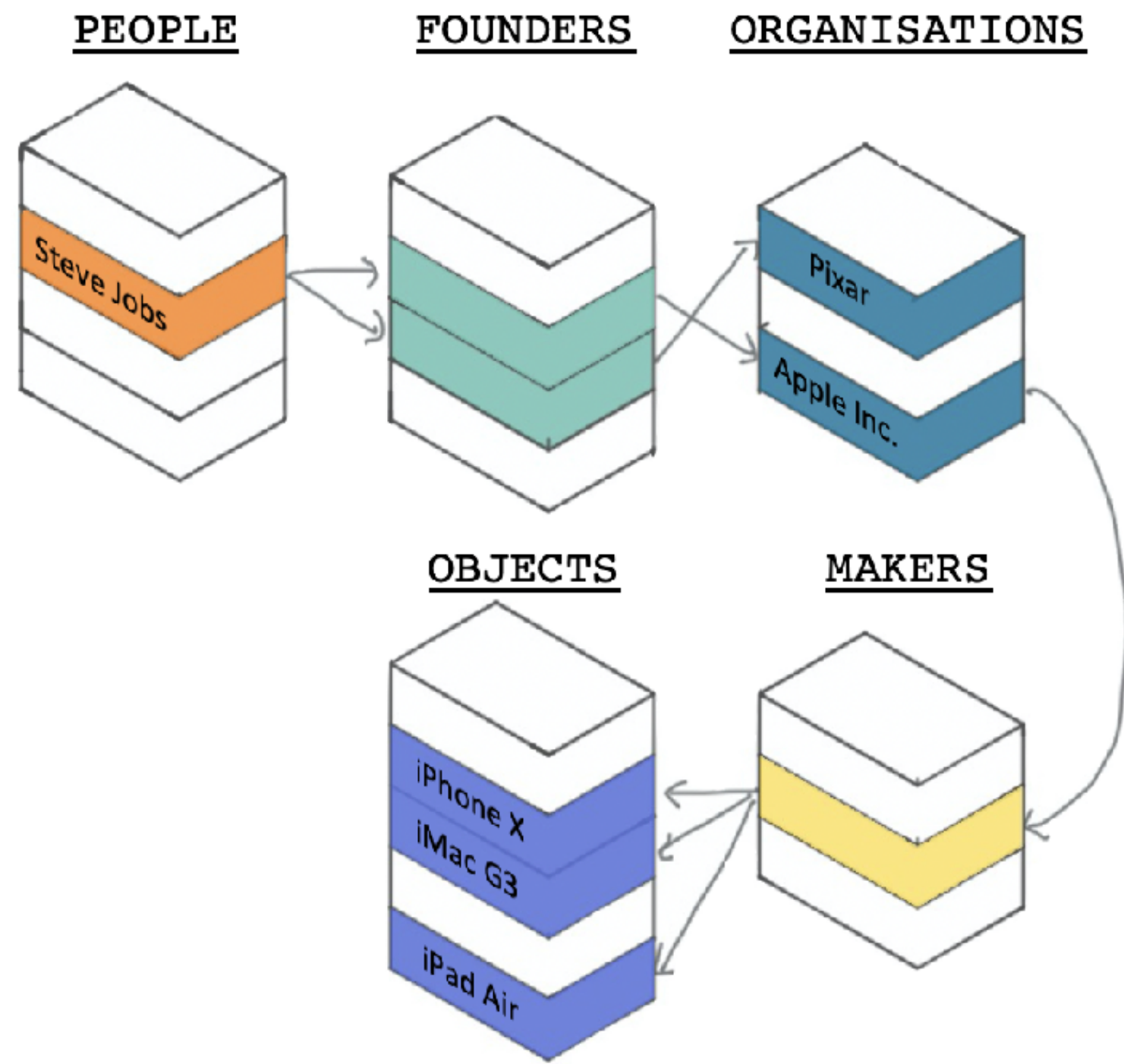
2. Improve discovery

3. Improve links to other data sources

KNOWLEDGE GRAPHS AND LINKED DATA







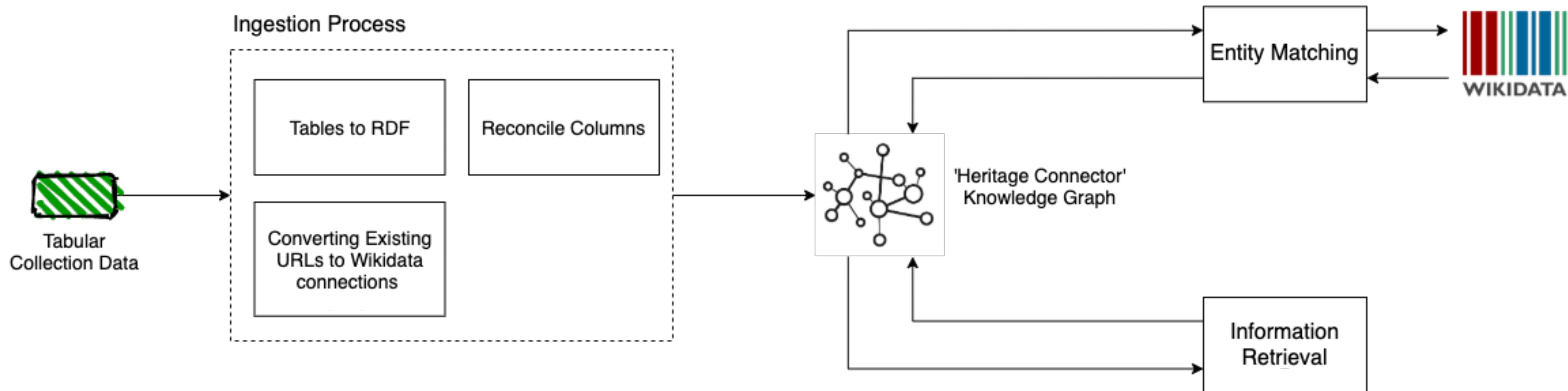
text **Anna Atkins** was born in Tonbridge, Kent

triples **anna_atkins**, birth_place, tonbridge_kent

linked data <https://collection.sciencemuseumgroup.org.uk/people/cp113991/>
<https://www.wikidata.org/wiki/Property:P19>,
<https://www.wikidata.org/wiki/Q936183>

ARTIFICIAL INTELLIGENCE

1. Processing IDs and URLs (links)
2. Adding new links to Wikidata with machine learning
3. Adding new links from texts with named entity recognition



Get all IDs and URLs
from the notes field
using regex (pattern
matching)



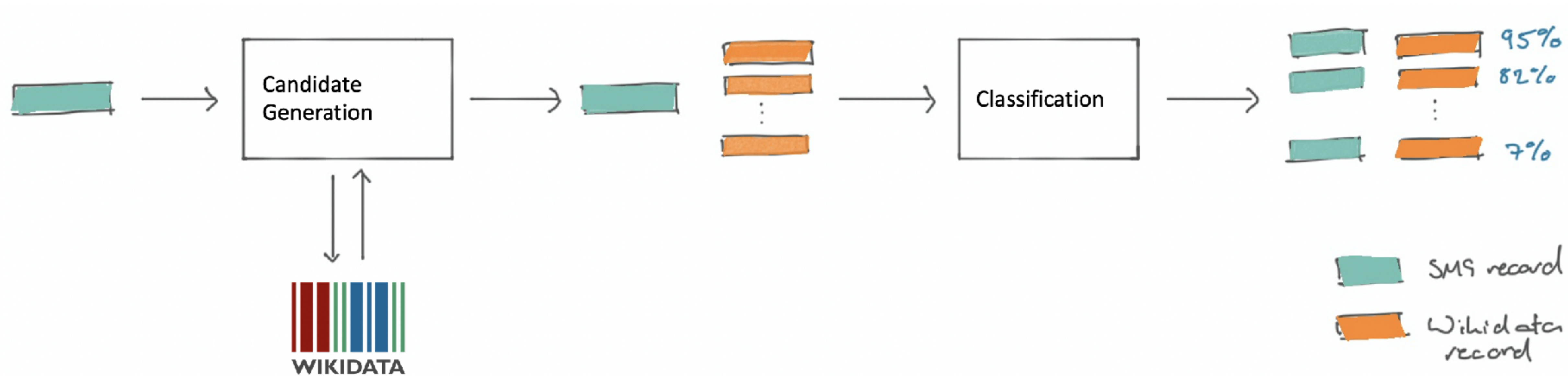
Resolve these URLs
to IDs of external
databases, where
possible



Use a SPARQL
query to exchange
external IDs for
Wikidata IDs



Filter found IDs to
ones that match
the record using
labels and dates



Sokol space suit

MADE: 1991 in Russia



SOKOL space suit worn by Helen Sharman
Science Museum Group Collection
© The Board of Trustees of the Science Museum

↗ > Use this image

British astronaut, Helen Sharman's Sokol spacesuit made by Zvezda. Sharman wore this rescue suit during the space flight on board the SOYUZ-TM-12 and MIR spacecraft in May 1991. Space suit model number KV-2 No. 167.

Sokol-KV-2 rescue suit worn by Helen Sharman during the Juno mission to the Mir space station, 1991

СПАСАТЕЛЬНЫЙ СКАФАНДР

Helen Sharman was the first British person in space. Sharman wore this suit for two hours on the ground to check its fit. Lying back, she tried to read but her arms ached from holding the book for so long. Despite the suit's cooling systems she sweated 2 litres during the mission launch. Once she could remove the suit, she dried it thoroughly to ensure it would not go mouldy.

The Sokol suit was developed after three unsuited cosmonauts asphyxiated on the Soyuz 11 mission in 1971 when their descent module depressurised during the return to Earth. Every cosmonaut now wears one during launch and return from space. It will keep the wearer alive for a number of hours in the event of a cabin depressurisation. Each suit is tailor made to the individual cosmonaut and comprises an inner, airtight 'bladder' of rubberised plastic and an outer layer of nylon canvas. There are connecting rings on the lower abdomen for air (cooling) and oxygen supplies and a centrally positioned pressure adjustment valve control on the chest; the pressure gauge is on the left wrist. The helmet and boots are integral with the rest of the suit; the gloves are attached with anodized aluminium bayonet fixings. Today's Sokol design is little changed from the original.

ON DISPLAY

[Science Museum: Exploring Space Gallery](#)

If you are visiting to see this object, [please contact us](#) in advance to make sure that it will be on display.

RELATED PEOPLE

[Helen Sharman](#)

RELATED ARTICLES

National Science and Media Museum

[Bring the National Science and Media Museum collection home in Animal Crossing](#)
Science Museum

[Highlights on display](#)

[Science Museum announces National Lottery ticket sales trial as Helen Sharman spacesuit goes back on display](#)

[UK tour of Tim Peake's spacecraft attracts 1.3 million visitors as Science Museum marks Apollo anniversaries with Summer of Space](#)

British **NORP** astronaut, Helen Sharman's **PERSON** Sokol **OBJECT** spacesuit made by Zvezda **ORG** . Sharman **PERSON** wore this rescue suit during the space flight on board the SOYUZ-TM-12 and MIR spacecraft in May 1991 **DATE** . Space suit model number KV-2 No. 167 **CARDINAL** .

Sokol-KV-2 **OBJECT** rescue suit worn by Helen Sharman **PERSON** during the Juno **OBJECT** mission to the Mir **OBJECT** space station, 1991 **DATE**

СПАСАТЕЛЬНЫЙ СКАФАНДР

Helen Sharman **PERSON** was the first British **NORP** person in space. Sharman **PERSON** wore this suit for two hours on the ground to check its fit. Lying back, she tried to read but her arms ached from holding the book for so long. Despite the suit’s cooling systems she sweated 2 litres during the mission launch. Once she could remove the suit, she dried it thoroughly to ensure it would not go mouldy.

The Sokol **OBJECT** suit was developed after three unsuited cosmonauts asphyxiated on the Soyuz 11 **OBJECT** mission in 1971 **DATE** when their descent module depressurised during the return to Earth **LOC** . Every cosmonaut now wears one during launch and return from space. It will keep the wearer alive for a number of hours in the event of a cabin depressurisation. Each suit is tailor made to the individual cosmonaut and comprises an inner, airtight 'bladder' of rubberised plastic and an outer layer of nylon canvas. There are connecting rings on the lower abdomen for air (cooling) and oxygen supplies and a centrally positioned pressure adjustment valve control on the chest; the pressure gauge is on the left wrist. The helmet and boots are integral with the rest of the suit; the gloves are attached with anodized aluminium bayonet fixings. Today **DATE** 's Sokol **ORG** design is little changed from the original.

Helen Sharman 1963

OCCUPATION: [Astronaut](#), [Broadcaster](#), [Chemist](#), [Engineer](#), [Lecturer](#)

NATIONALITY: [British](#)

BORN IN: [Sheffield, South Yorkshire, England, United Kingdom](#)

NPP Zvezda (Q541905)

company in [Moscow, Russia](#)

 [edit](#)

[K-36DM I Zvezda \(Russia\)](#) | [Research-and-production enterprise "Zvezda" to them.](#)

[GI Severin I Zvezda Research and Production Enterprise](#)

[British](#) **NORP** astronaut, [Helen Sharman's](#) **PERSON** [Sokol](#) **OBJECT** spacesuit made by [Zvezda](#) **ORG** . [Sharman](#) **PERSON** wore this rescue suit during the space flight on board the SOYUZ-TM-12 and MIR spacecraft in [May 1991](#) **DATE** . Space suit model number KV-2 No. [167](#) **CARDINAL** .

[Sokol-KV-2](#) **OBJECT** rescue suit worn by [Helen Sharman](#) **PERSON** during the [Juno](#) **OBJECT** mission to the [Mir](#) **OBJECT** space station, [1991](#) **DATE**

СПАСАТЕЛЬНЫЙ СКАФАНДР

[Helen Sharman](#) **PERSON** was the first [British](#) **NORP** person in space. [Sharman](#) **PERSON** wore this suit for two hours on the ground to check its fit. Lying back, she tried to read but her arms ached from holding the book for so long. Despite the suit's cooling systems she sweated 2 litres during the mission launch. Once she could remove the suit, she dried it thoroughly to ensure it would not go mouldy.

The [Sokol](#) **OBJECT** suit was developed after three unsuited cosmonauts asphyxiated on the [Soyuz 11](#) **OBJECT** mission in [1971](#) **DATE** when their descent module depressurised during the return to [Earth](#) **LOC** .

Every cosmonaut now wears one during launch and return from space. It will keep the wearer alive for a number of hours in the event of a cabin depressurisation. Each suit is tailor made to the individual cosmonaut and comprises an inner, airtight 'bladder' of rubberised plastic and an outer layer of nylon canvas. There are connecting rings on the lower abdomen for air (cooling) and oxygen supplies and a centrally positioned pressure adjustment valve control on the chest; the pressure gauge is on the left wrist. The helmet and boots are integral with the rest of the suit; the gloves are attached with anodized aluminium bayonet fixings. [Today](#) **DATE**

's [Sokol](#) **ORG** design is little changed from the original.

Sokol space suit (Q1197668)

Russian spacesuit used on Soyuz

[Sokol IVA](#) | [Sokol](#)

Soyuz 11 (Q648581)

Manned Soviet space mission to the Salyut 1 Space Station

```

DATE_PATTERNS = [
    {"label": "DATE", "pattern": [{"SHAPE": "dddd"}, {"ORTH": "-"}, {"SHAPE": "dddd"}]}, # 1984 - 1990 | 1984-1990
    {"label": "DATE", "pattern": [{"ORTH": "c."}, {"SHAPE": "dddd"}]}, # c. 1200
    {"label": "DATE", "pattern": [{"TEXT": {"REGEX": r"c\.\d{3,4}"}}]}, # c.1200
    {"label": "DATE", "pattern": [{"TEXT": {"REGEX": r"c\.\d{3,4}"}, {"ORTH": "-"}, {"SHAPE": "dddd"}]}, # c.1200 - 1220 | c.1200-1220
    {"label": "DATE", "pattern": [{"TEXT": {"REGEX": r"\d{1,2}/\d{1,2}/(\d{4}|\d{2})"}]}]}, # 03/12/2000
    {"label": "DATE", "pattern": [{"TEXT": {"REGEX": r"\d{1,2}\.\d{1,2}\.(\d{4}|\d{2})"}]}]}, # 03.12.2000
    {"label": "DATE", "pattern": [{"SHAPE": "dd"}, {"ORTH": "-"}, {"SHAPE": "dd"}, {"ORTH": "-"}, {"SHAPE": "dddd"}]}, # 03-12-2000
    {"label": "DATE", "pattern": [{"SHAPE": "d"}, {"ORTH": "-"}, {"SHAPE": "dd"}, {"ORTH": "-"}, {"SHAPE": "dddd"}]}, # 3-12-2000
    {"label": "DATE", "pattern": [{"SHAPE": "dd"}, {"ORTH": "-"}, {"SHAPE": "d"}, {"ORTH": "-"}, {"SHAPE": "dddd"}]}, # 03-1-2000
    {"label": "DATE", "pattern": [{"SHAPE": "d"}, {"ORTH": "-"}, {"SHAPE": "d"}, {"ORTH": "-"}, {"SHAPE": "dddd"}]}, # 3-1-2000
    {"label": "DATE", "pattern": [{"SHAPE": "dddd"}, {"ORTH": "to"}, {"SHAPE": "dddd"}]}, # 1805 to 1860
]

```

```

COLLECTION_NAME_PATTERNS = [
    # TODO: use 'POS': 'PROPN' here instead of IS_TITLE: True for better detection of proper nouns
    {"label": "ORG", "pattern": [{"IS_TITLE": True, 'OP': '+'}, {'LOWER': 'collection'}]}, # Sforza collection
    {"label": "ORG", "pattern": [{"IS_TITLE": True, 'OP': '+'}, {'LOWER': 'archive'}]}, # Charles Urban archive
]

```


{"label": "ORG", "pattern": "Thames Archway Company", "id": "<https://collection.sciencemuseumgroup.org.uk/people/cp15926>"}

{"label": "ORG", "pattern": "Hodbarrow Mining Company", "id": "<https://collection.sciencemuseumgroup.org.uk/people/cp16807>"}

{"label": "ORG", "pattern": "HMS Vanguard (1815)", "id": "<https://collection.sciencemuseumgroup.org.uk/people/cp17108>"}

{"label": "ORG", "pattern": "Wind Energy Group", "id": "<https://collection.sciencemuseumgroup.org.uk/people/cp17473>"}

{"label": "ORG", "pattern": "E R and F Turner Limited", "id": "<https://collection.sciencemuseumgroup.org.uk/people/cp17945>"}

{"label": "ORG", "pattern": "Baird Television Limited", "id": "<https://collection.sciencemuseumgroup.org.uk/people/cp17663>"}

{"label": "ORG", "pattern": "Alliance Box Company", "id": "<https://collection.sciencemuseumgroup.org.uk/people/cp24886>"}

{"label": "ORG", "pattern": "Hell", "id": "<https://collection.sciencemuseumgroup.org.uk/people/cp21022>"}

{"label": "ORG", "pattern": "Paradigm Models Limited", "id": "<https://collection.sciencemuseumgroup.org.uk/people/cp22440>"}

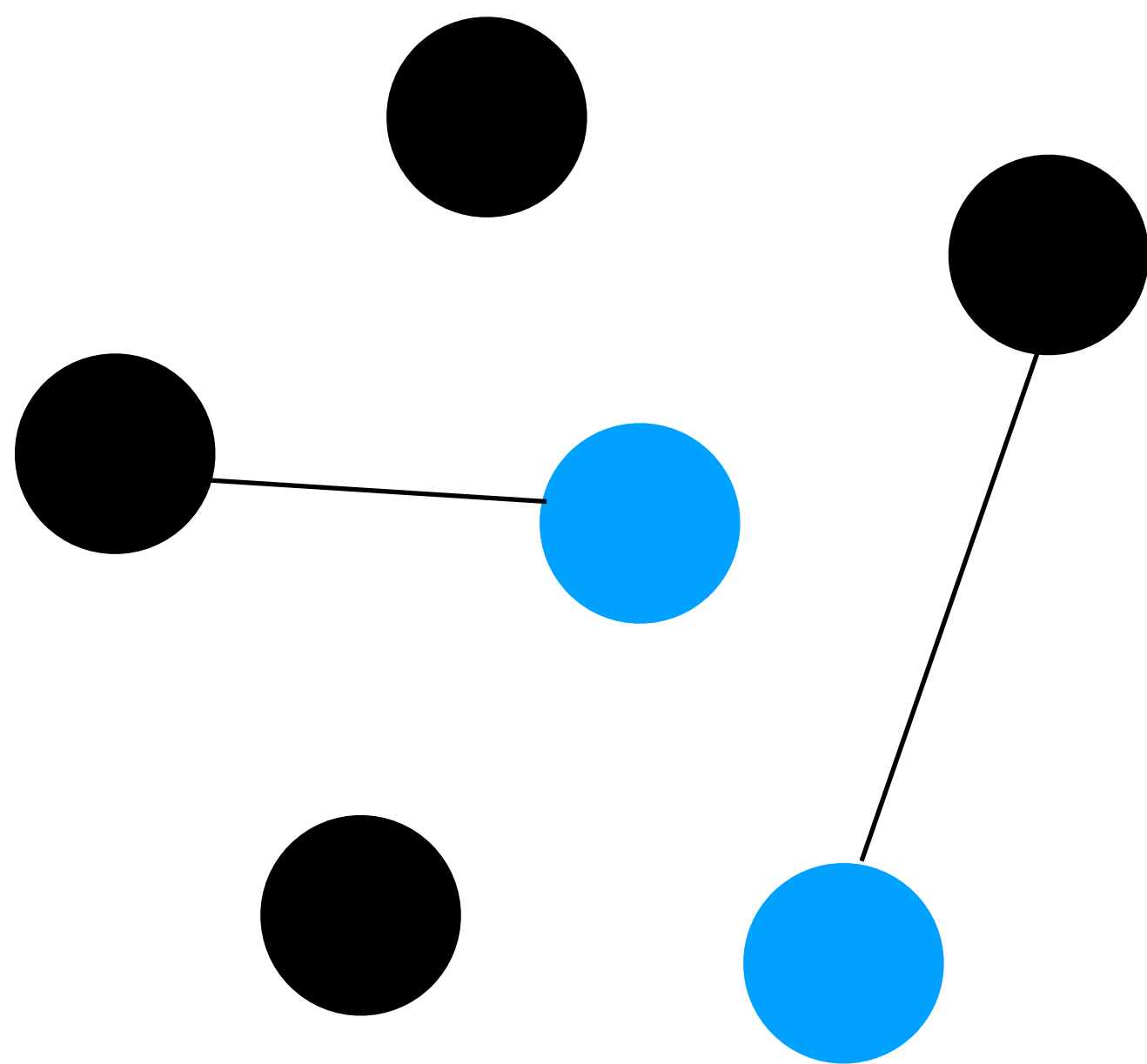
{"label": "ORG", "pattern": "City of York Council", "id": "<https://collection.sciencemuseumgroup.org.uk/people/cp19207>"}

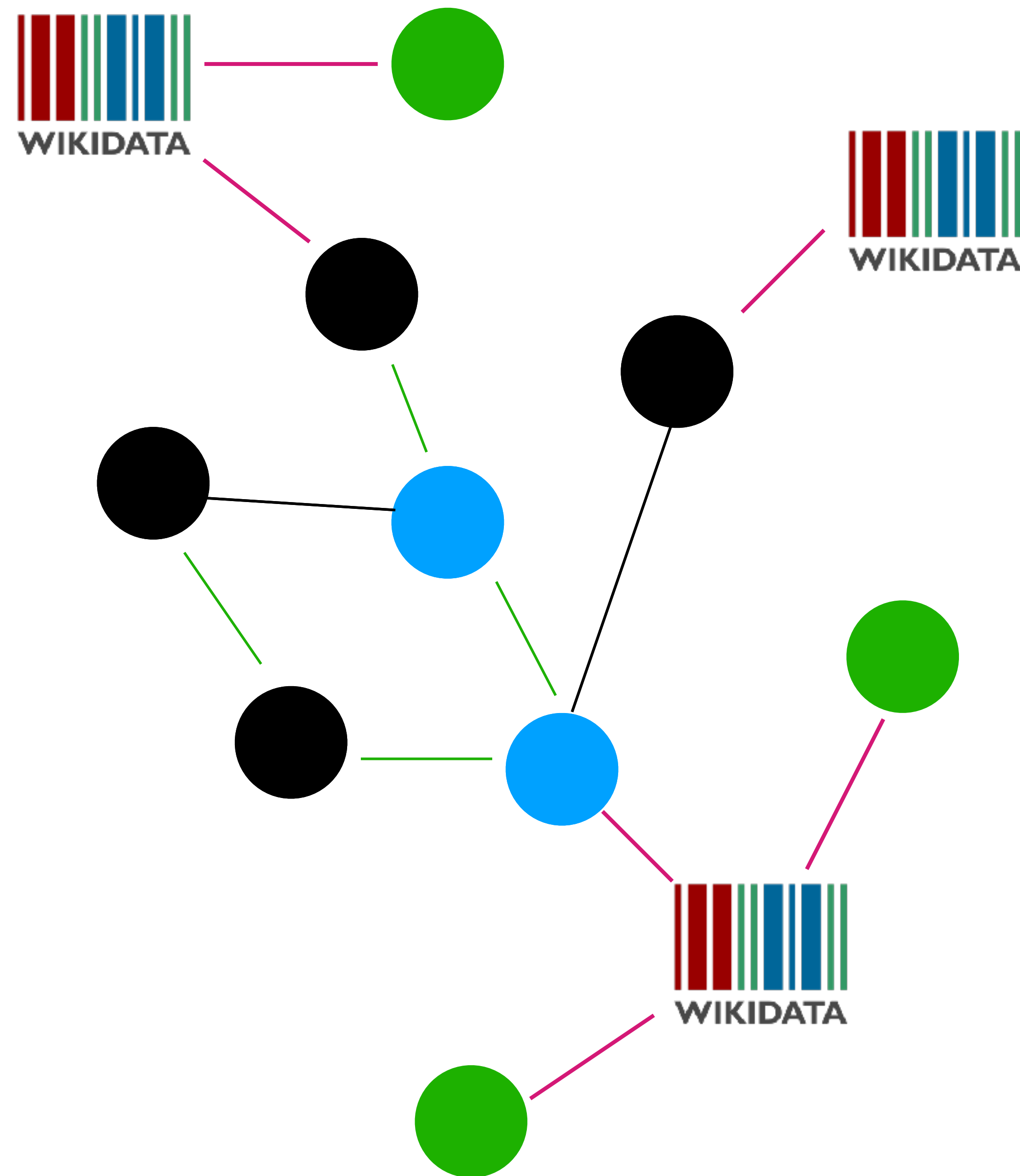
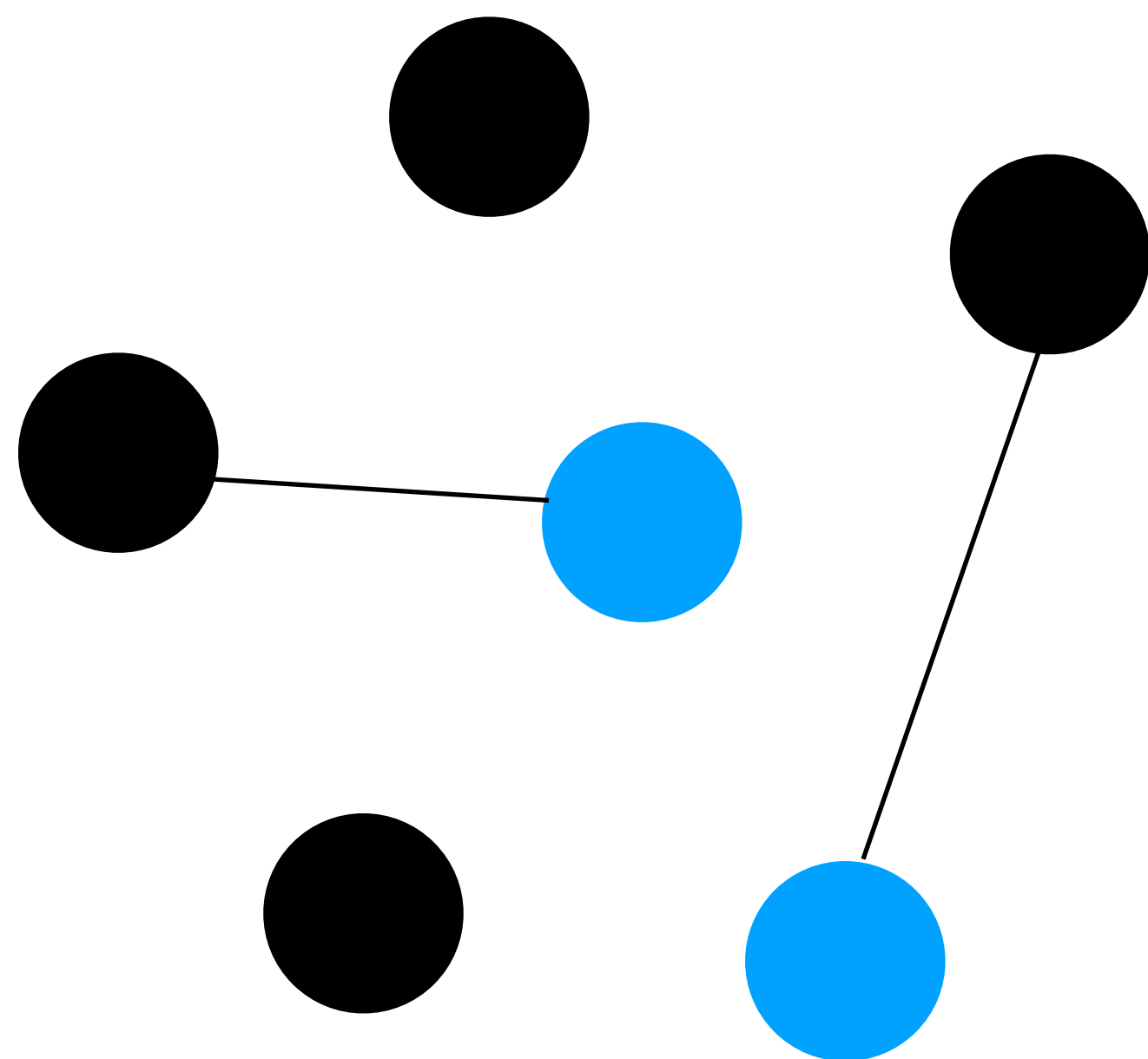
{"label": "ORG", "pattern": "Kvaerner Masa-Yards", "id": "<https://collection.sciencemuseumgroup.org.uk/people/cp24946>"}

{"label": "ORG", "pattern": "Frederick Bateman and Company Limited", "id": "<https://collection.sciencemuseumgroup.org.uk/people/cp20289>"}

{"label": "ORG", "pattern": "Normal School of Science, Astronomy Laboratory", "id": "<https://collection.sciencemuseumgroup.org.uk/people/cp20442>"}

{"label": "ORG", "pattern": "T Green & Son Ltd", "id": "<https://collection.sciencemuseumgroup.org.uk/people/cp20553>"}





PRELIMINARY FINDINGS

- Cultural heritage databases:
 - rich
 - large
 - complex
 - limited standardisation
- Standardisation challenging

- Motivations for working with Linked Open Data include:
 - more visible
 - exposing 'hidden' / 'hidden' aspects
 - enrichment
 - data reuse in new contexts
 - better user experience

- Working with linked data at any kind of scale:
 - time consuming
 - resource intensive

- Aligning free-text fields to entities can take a significant time using existing tools
- More robust methods therefore exist in the Heritage Connector approach.

- Creating external and internal links work better when used iteratively
- As named entity recognition creates more entities and relations, the effectiveness of the disambiguator increases.

- You can expect varying success disambiguating records with Wikidata depending on their type

- It is not a question of *if* human intervention and curation is needed, but at what point it should be used and how it may be most usefully focused

NEXT STEPS

- Ingest V&A Collection data

- Extend data beyond collections and Wikidata to include text content such as articles

- Build and test a robust internal link creation method for heritage collections data

- Explore how knowledge graphs enable new forms of interaction and discovery in practice.

THANKS

<https://www.sciencemuseumgroup.org.uk/project/heritage-connector/>

<https://thesciencemuseum.github.io/heritageconnector/>

https://www.zotero.org/groups/2439363/heritage_connector

<https://www.youtube.com/channel/UCz06jrolvj-JbFuiQ9BpZdQ>

<https://github.com/TheScienceMuseum/heritage-connector/>