

Tattoos in the Digital Panopticon Database, 1793-1925

Description

This dataset contains information about tattoos on 58,002 criminal convicts from records created between 1793 and 1925. These datasets were created by extracting data from datasets in the Digital Panopticon (www.digitalpanopticon.org), a compilation of 50 datasets containing records pertaining to men and women convicted of crimes at the Old Bailey court in London in the eighteenth and nineteenth centuries. The eight datasets from which tattoos data have been extracted include substantial physical descriptions of the convicts' bodies, including evidence of tattoos. Other information about the convicts was extracted from the 'life archives' of these convicts, which contain evidence from all the records pertaining to that convict in the Digital Panopticon. For more information about this data, see: https://www.digitalpanopticon.org/Tattoos,_1793-1925.

The information given includes details of the names and demographic details (age, gender, occupation, religion) of the convicts, their tattoos (descriptions, subjects, body location), and other physical marks on their bodies (such as scars and boils). For further details, see below.

There are two tables in the dataset:

- convict_descriptions: each row consists of the physical description of one convict (241,207 rows). Many convicts have separate entries from records created at different times.
- convict_description_segments: each row consists of a 'segment' of a physical description, which describes the marks on a specific part of a body, for example the right arm (565,280 rows).

Background

This dataset was created as part of a British Academy funded project, 'Analysing Criminal Tattoos through Data Mining and Visualisation'. Using data from the AHRC funded 'Digital Panopticon' project, the project developed new methods of data extraction and visualisation in order to better understand the meanings embedded within this and other rich bodies of textual evidence.

This dataset is one of a number to be released as open data by the Digital Panopticon project. It can be used for standalone research but additionally, the project will release record linkage data (known as 'life archives' data) which will enable researchers to reconstruct information about individuals from multiple datasets in ways that go beyond what is possible using the search facilities at www.digitalpanopticon.org.

The data comes from two Digital Panopticon datasets, with separate codes:

convict_descriptions (TTD)
convict_description_segments (TTC)

Original records

The tattoos dataset is extracted from datasets included in the Digital Panopticon project. For more information, see https://www.digitalpanopticon.org/Tattoos,_1793-1925. Links on that page link to information about each of these constituent datasets.

Data creation and limitations

The source datasets contain physical descriptions of convicts' bodies. Extracting and analysing information about tattoos from these descriptions is not straightforward, because they contain a wide range of other information: physical characteristics (eye colour, shape of face, etc), bodily infirmities (lame leg, broken nose), and personal details (such as occupation or religion), in a variety of formats. The challenge in this project was to extract the information about tattoos from all the other information in these descriptions. The methodology adopted, using a combination of segmentation, pattern matching, dictionaries, iterative manual checking, and historical interpretation, is described under 'Methodology' on the webpage https://www.digitalpanopticon.org/Tattoos,_1793-1925.

As far as we are aware, this data deposit contains the largest amount of information about tattoos available for the nineteenth century. The information about convict tattoos provides otherwise unavailable access to the sentiments of men and women who left few if any other records in their own voice. Nonetheless, there are some significant limitations to this data:

- First, these are only written descriptions, and often they provide very limited evidence of what the tattoos actually looked like. The clerks who recorded this information used their own words, and sometimes censored obscene language. They may also have failed to record all tattoos. Accounts vary as to whether convicts were stripped to the waist or totally naked while the tattoos were recorded, and it is likely that the relatively small number of tattoos on the lower half of the body is due to the fact such tattoos were often not visible.
- Second, the computational methodologies used to extract this information (described above), are not infallible, and it is likely that some tattoos have been omitted or misidentified.
- Finally, the historical judgement involved in identifying designs, subjects, and words written on bodies is subjective and fallible; it is likely that other historians would have made somewhat different decisions.

The data

The dataset contains two tables: `convict_descriptions` and `convict_description_segments`.

convict_descriptions

columns

- subrecordid = The unique ID of the Digital Panopticon subrecord from which this description is taken.
- recordid = The unique ID of the Digital Panopticon record from which this description is taken.
- life_id = The unique ID of the Digital Panopticon life archive from which this description is taken.
- dataset = dataset from the Digital Panopticon from which this evidence is taken (see list below)
- hastattoo (n/y/u) = does this record indicate the convict has a tattoo: no, yes, unknown (the physical description of the convict is blank)
- haspunishmentmark (n/y/u) = does convict have a mark from physical punishment: no, yes, unknown (the physical description of the convict is blank)
- designs = segmented description of each separate tattoo design on convict's body, divided by pipe characters (|)
- writtenwords = words written on the convict's body
- writtenyears = years written on the convict's body
- punishmentmarks = language used to describe any punishment marks on the convict's body
- body = words used to describe the body, such as 'right forearm', with body parts separated by pipe character (|)
- digest = location on convict's body of each tattoo together with description of tattoo, each tattoo separated by pipe character (|)
- subjects = subjects of any tattoos, determined manually by project staff by selecting from a defined list (provided below)
- url = link to the convict's life archive in the Digital Panopticon
- fulldescription = full transcription of the description of marks on the convict's body in the original data
- given = name of convict
- surname = name of convict
- gender (f/m/u) = female, male, unknown
- transported (no/yes) = was the convicted transported to Australia.
- ticket_of_leave (no/yes) = if transported, was the convict eventually awarded a 'ticket of leave', a form of probation.
- penalservitude (no/yes) = if imprisoned in Britain, was the convict sentenced to penal servitude, a sentence which replaced transportation in 1857.
- granted_prison_license (no/yes) = a prison licence was a form of probation.
- married (unknown/yes) = indicates whether there is evidence in the Digital Panopticon that the convict was married. This is of limited use since the Digital Panopticon records have very limited evidence about marriages.
- insane (no/yes) = indicates whether there is any evidence in the Digital Panopticon records that the judicial authorities deemed the convict to be insane.
- in_hulks (unknown/yes) = indicates whether there is evidence in the Digital Panopticon that the convict served time in the hulks.
- religion

- religion_category = the descriptions of convicts' religions have been consolidated into a small number of categories to enable statistical analysis; see list below.
- occupation
- occupation_top_50 = standardised name of the occupation if it is in the top fifty occupations in the database (list below)
- hisco = a number indicating the social class category of the occupation, as indicated by HISCO, a a historical international classification of occupations (list below)
- place_of_birth
- ship = name of the ship if the convict was transported to Australia
- earliest_trial_offence_category = the category of offence the convict was first charged with in the records of the Digital Panopticon (explanation below)
- earliest_trial_sentence_category = the category of punishment sentence for the convict's first conviction in the records of the Digital Panopticon(explanation below)
- latest_trial_offence_category = the category of offence the convict was last charged with in the records of the Digital Panopticon(explanation below)
- latest_trial_sentence_category = the category of punishment sentence for the convict's last conviction in the records of the Digital Panopticon(explanation below)
- descyear = the year in which the description was written
- desc_count = The number of descriptions of this person in the database.
- born = year of the convict's birth. This is often calculated from their age as recorded in a later record, so may be imprecise.
- trials = number of trials of this convict recorded in the Digital Panopticon
- earliest_trial
- latest_trial

convict_description_segments

columns

Many of these are the same as in the convict_descriptions database, and are explained above. The following columns, however, are only used in this database, which divides the full description of a convict's marks into separate 'segments'. A 'segment' is a part of the description which refers to a specific body part, for example their right arm. The additional columns are:

- life_id = The Digital Panopticon life ID of the person for whose description the segment is derived.
- tattoo_in_segment (n/y) = is there a tattoo in this specific segment, no or yes
- pm_in_segment (n/y) = is there a punishment mark in this specific segment, no or yes
- tattoo_in_description (n/y) = is there a tattoo in the full description of the convict's marks, no or yes
- pm_in_description (n/y) = is there a punishment mark in the full description of the convict's marks, no or yes
- segment_type = The tattoo identification process characterised each segment with respect to the features it thought it contained, e.g. if it contained a body part and a tattoo, it was assigned the segment_type 'body tattoo'.

- `segment_parts` = text of the segment, divided into parts with each mark and body part separated by a pipe character (|)
- `segment_part_classifications` = How each part of the segment has been classified. For example, as a part of the body, or as a mark.
- `other_descriptive_terms` = Descriptive terms which could not be assigned to a specific body part, such as 'pockpitted'.
- `adjectives` = any adjectives used in the description, including colours, numbers, and sizes
- `marks` = words used to describe any non-tattoo marks on the convict's bodies, such as scars or boils.
- `injuries` = words used to describe any injuries recorded on the convict's body.
- `unknown` = any words in the description which our methods failed to categorise.
- `segment_year` = year in which the specific segment was recorded.

Explanations of entries in specific columns

datasets

For information about these source datasets, see
https://www.digitalpanopticon.org/Tattoos,_1793-1925

dataset	description
cin	Western Australia Convict Indents
fas	Founders and Survivors
fas_pgo	Founders and Survivors data from the Police Gazette
hcr	Criminal Registers
mpr	Millbank Prison Registers
pld	Female Prison Licences
rhc	Metropolitan Police Register of Historical Criminals
tlm	Male Prison Licences

subjects

`america` = Phrases or symbols relating to America or American history, culture, identity.

`astronomy` = Phrases or symbols relating to astronomy. For example, sun and moon, stars.

`australia` = Phrases or symbols relating to Australia or Australian history, culture, identity.

`britain` = Phrases or symbols relating to Britain or British history, culture, identity.

`death` = Phrases or symbols relating to death. For example, skull, cross-bones.

`invention` = Phrases or symbols relating to nineteenth and twentieth-century inventions e.g. camera.

ireland = Phrases or symbols relating to Ireland or Irish history, culture, identity.

jewellery = Tattoos that represent jewellery. For example, ring, bracelet.

justicepunishment = Phrases or symbols relating to justice and punishment. For example, scales, whip.

love = Phrases or symbols relating to love. For example, heart, 'I love you'

military = Phrases or symbols relating to the military or warfare. For example, sword, Lord Kitchener, pistol, soldier.

namesinitials = Any names or sequences of letters that can be interpreted as initials.

nationalidentity = Term refers to tattoos that represent an aspect of national identity. For example, flag, cross flags. Most have also been given subject categories for specific countries.

nature = Phrases and symbols related to animals, flowers and plants. For example, elephant, flower pot

naval = Phrases or symbols relating to the Navy or maritime and seafaring cultures. For example, anchor, mermaid, sailor.

pleasure = Phrases or symbols relating to leisure practices including drinking, gambling, recreation. For example, cards, tobacco, glass, dancing.

religion = Phrases or symbols relating to the Christian religion. For example, seven candlesticks, cross, crucifix, clasped hands

sex = Phrases or symbols relating to sex or obscenity. For example, naked, obscene.

year = A tattoo with the four digit number of a year.

religion_category

<u>religion_category</u>	<u>description</u>
catholic	
church_of_england	
jewish	
miscellaneous	Includes Greek Church.
mixed	This category is used when multiple records in a convict's life archive indicate religions in more than one category.
muslim	
other_protestant	Includes baptist, calvinist, lutheran, methodist, presbyterian
unspecified_protestant	Religion is given as 'protestant' only.

occupation_top_50

List of the top fifty occupations in the whole Digital Panopticon dataset, for which the names have been standardized:

agent
baker
barman
blacksmith
bootmaker
boy
bricklayer
butcher
cabinet
carman
carpenter
charwoman
clerk
cook
costermonger
dealer
dressmaker
engineer
errand
farm
french
gardener
groom
hawker
jeweller
labourer
laundress
maker
married
merchant
other
painter
plasterer
polisher
porter
postman
printer
sailor
salesman
seaman
servant
shoemaker

shopman
smith
soldier
solicitor
spinster
stoker
tailor
traveller
waiter
weaver
wife

hisco

<u>hisco_code</u>	<u>description</u>
0.5	aristocracy
1	higher managers
2	higher professionals
3	lower managers
4	lower professionals, higher clerical and sales personnel
5	lower clerical and sales personnel
6	foremen
6.5	small manufacturers
7	medium-skilled workers
8	farmers and fishermen
9	lower-skilled workers
10	lower-skilled farm workers
11	unskilled workers
12	unskilled farm workers
13	workers (no further information)
14	children

earliest/latest _trial_offence_category

Offence categories are those used for the Old Bailey Online. For further information, see: <https://www.oldbaileyonline.org/static/Crimes.jsp>

Each offence is categorized both as a specific offence, and within a general category, with each entry providing the specific offence first, followed by a space and then the general offence category. So in the first entry below, the specific offence is 'animal theft' and the general category is 'theft'. The possibilities are listed below:

animal_theft theft
assault breaking_peace
breaking_peace libel
breaking_peace riot
breaking_peace vitriolage
breaking_peace wounding
burglary theft
concealing_a_birth miscellaneous
damage_to_property
damage_to_property arson
damage_to_property_other damage_to_property
deception
deception bankruptcy
deception deception_other
deception fraud
deception perjury
embezzlement theft
forgery deception
game_law_offences theft
grand_larceny theft
housebreaking theft
infanticide killing
keeping_a_brothel sexual_offences
kidnapping miscellaneous
killing
killing manslaughter
mail_theft theft
miscellaneous
miscellaneous conspiracy
miscellaneous disorderly
miscellaneous drunkenness
miscellaneous felony
miscellaneous habitual_criminal
miscellaneous illegal_abortion
miscellaneous miscellaneous_other
miscellaneous perverting_justice

miscellaneous piracy
miscellaneous prostitution
miscellaneous returning_from_transportation
murder killing
petty_larceny theft
pocketpicking theft
rape sexual_offences
receiving theft
royal_offences coining
royal_offences uttering
sexual_offences
sexual_offences bigamy
sexual_offences sodomy
sexual_offences_other sexual_offences
shopbreaking theft
shoplifting theft
simple_larceny theft
stealing_from_master theft
theft
theft extortion
theft_from_a_specified_place theft
threatening_behaviour breaking_peace
treason royal_offences
vagabonding breaking_peace
violent_theft highway_robbery
violent_theft robbery

earliest/latest _trial_sentence_category

Sentence categories are those used for the Old Bailey Online. For further information, see: <https://www.oldbaileyonline.org/static/Punishment.jsp>

Please note that these were the sentences meted out by the judges; they do not indicate that the convict actually received that punishment. For a variety of reasons, often owing to pardons, the actual punishment received was less harsh than that sentenced.

Each sentence is categorized both as a specific sentence, and within a general sentence category, with each entry providing the specific sentence first followed by a space and then the general category. So in the third entry below, the specific sentence is to be drawn and quartered, and the general category is 'death'. Where no specific subcategory was indicated, only the general category is provided, as in the first entry below. The possibilities are listed below:

death
death_respited death
drawn_and_quartered death
imprisonment

imprisonment hard_labour
imprisonment house_of_correction
imprisonment imprisonment_insanity
imprisonment imprisonment_newgate
imprisonment_other_institution imprisonment
miscellaneous fine
miscellaneous military_naval_duty
miscellaneous sureties
no_punishment
penal_servitude imprisonment
transportation
whipping corporal

Further information

Digital Panopticon descriptions of the source datasets:

https://www.digitalpanopticon.org/Tattoos,_1793-1925

Digital Panopticon research page outlining some of the project's findings concerning the practice of tattooing: https://www.digitalpanopticon.org/Convict_tattoos

Article in *The Conversation* about the research findings:

<https://theconversation.com/how-tattoos-became-fashionable-in-victorian-england-122487>

Acknowledgments

The British Academy funded this research project under its 'Digital Research in the Humanities 2018' funding scheme.

The Arts and Humanities Research Council funded the Digital Panopticon project (2013-2017).

Professor Robert Shoemaker (University of Sheffield) was the principal investigator of the Tattoos project and Dr Zoe Alker (University of Liverpool) was co-investigator.

Jamie McLaughlin, Research Software Engineer at the Digital Humanities Institute of the University of Sheffield, performed the data extraction and categorisation, and created the databases.

Sharon Howard, early modernist and digital historian, audited and analysed the data, and created visualisations which allowed us to better understand the data and its limitations. She and Jamie McLaughlin were primarily responsible for the creation of this data deposit.

Licence

The dataset is released under a Creative Commons Attribution Non-Commercial 4.0 (CC BY-NC) licence: <https://creativecommons.org/licenses/by-nc/4.0>

Please contact the Digital Panopticon project if you wish to use the data for purposes that fall outside the terms of the licence, or if you're interested in future research collaborations:
https://www.digitalpanopticon.org/Contact_Us

Citation

Digital Panopticon, 'Tattoos, 1793-1725', version 1.2.1 (February 2020),
www.digitalpanopticon.org.