

The AHDS Taxonomy of Computational Methods

This short paper presents the AHDS Taxonomy of Computational Methods which has been developed as part of the AHDS Projects and Methods Database project. It briefly outlines the taxonomy's rationale, explains in some more detail its framework, and includes the most recent version of the taxonomy.

What are the aims of AHDS Projects and Method Database?

The principal aim of the Projects and Method Database project is to establish a register containing detailed information on current and recent arts and humanities research projects using ICT, and on the computational methods employed by such projects.

What are the aims of the AHDS Taxonomy of Computational Methods?

The methods taxonomy serves as a controlled vocabulary for the method section of the database. Scholars, invited to provide input for the database will be asked to choose these terms from the taxonomy that denote most closely the computational methods they have employed in the course of their research. By gathering of information on the employment of methods in a controlled fashion, the database's capability for structured searching is greatly enhanced.

What kind of computational methods are represented in the taxonomy?

For the purpose of the method taxonomy, computational methods are defined as follows:

- a) The term method broadly denotes all the techniques and tools that are used to gain new knowledge in the various academic fields which constitute the arts and humanities.
- b) A method is a computational one if it is either based on ICT (i.e. database technology), or critically dependent on it (i.e. statistical analysis).
- c) As the initial version of the database focuses on projects involved in the creation of digital resources, the taxonomy focuses on computational methods used for the creation, analysis and dissemination of such resources.

How is the taxonomy structured?

Terms in the methods taxonomy are primarily classified according to the two dimensions "content type" and "function type".

a) Content types

The content types are used to denote the nature or genre of the content of the digital resource created. The taxonomy is constructed under the assumption that a resource's content remains stable during all stages of a resource creation project, regardless of possible changes in its digital manifestation. This should help scholars asked to provide input for the database. Indeed, once they established that they create a resource of a particular content type, they can locate all the terms needed to describe their methods under this particular content type heading.

E.g. Scanning textual material produces a resource of the data type "image" which can be transformed into a resource of the data type "text" by means of OCR software. The content type of the resource, however, always remains "text".

Content type terms

<i>Narrative text</i>	A resource that is primarily words for reading. For example books, letters, dissertations, poems, newspapers, articles, archives of mailing lists.
<i>Dataset/structured data and text</i>	A resource that primarily consists of structured information encoded in lists, tables, databases, etc., which will normally be in a format available for direct machine processing. Examples of datasets/structured data are: series of measurements, survey data, enumerator's data, census data, indexes and catalogues.
<i>Still Image/Graphics</i>	A resource that is primarily a static, 2-dimensional, visual representation. Examples are paintings, drawings or photographs of physical objects. Use the type "Spatial Resource" for plans, maps etc. if their spatial attributes are the main focus of interest.
<i>Moving Image</i>	A resource consisting of a series of visual representations that, when shown in succession, impart an impression of motion. Examples of moving images are: animations, movies, television programs, videos, zoetropes, or visual output from a simulation.
<i>3D object</i>	A resource that primarily describes the topology, surface and shape of an object three dimensionally. Examples include 3-d models, or related stereo images, of sculptures, buildings and other artefacts.
<i>Spatial</i>	A resource that primarily consists of geometrical primitives (i.e. points, polygons, lines) which are tied to a coordinate system or other positional information, allowing exact spatial measurements. Examples include maps, charts, and blueprints.
<i>Sound</i>	A resource that is primarily intended to be rendered as audio. Examples include recorded speech and music, and live performances of music.

The content type terms are located in first row of the taxonomy.

b) Function types

These denote the broad functions commonly undertaken in digital resource creation processes. Note that these functions are not completely discrete entities but can significantly overlap. Therefore, certain computational methods may need to be classified under more than one function type heading (i.e. the method "Record linkages" is classified as both a structuring/enhancing and as an analytical method).

Function type terms

<i>Capture</i>	Denotes the conversion of analogue information into (raw) digital data ("digitisation").
<i>Structuring and Enhancement</i>	Denotes the organisation and integration of the data captured from one or various sources into a uniform conceptual framework. This would typically include the design of the resource's information structure/data model as well as the normalisation, standardisation and enhancement of its data.
<i>Analysis</i>	Denotes the extraction of information/knowledge/meaning from the resource.
<i>Dissemination and presentation</i>	Denotes the presentation and dissemination/communication of the results of the research project.

The function terms are located in the first column of the method taxonomy.

Version 1.3.1 of the AHDS Taxonomy of Computational Methods

Resource type “Narrative text”

	Method Term	Definition
Capture	Manual transcription	Data capture by means of a manual transcription/inputting process.
	Scanning/photography - 2D	Data capture by means of a 2-dimensional scanner (i.e. flatbed scanner, film/slide scanner, drum scanner), or a digital camera.
	Digital sound recording	Data capture by means of sound recording equipment.
	Text recognition	i.e. OCR; refers to the conversion of scanned images of a text into text documents.
	Speech recognition	Refers to the conversion of spoken words and phrases into text.
	Usage of existing digital data	Refers to the usage of data that already exists in digital form
Structure	Markup/text encoding - descriptive - conceptual	Refers to the addition of character and symbols at certain places in a text in order to convey information about concrete and abstract concepts (i.e. genres, topical subjects).
	Markup/text encoding - descriptive - document structure	Refers to the addition of character and symbols at certain places in a text in order to convey information about its logical structure (i.e. identification of headings, paragraphs).
	Markup/text encoding - descriptive - linguistic structure	Refers to the addition of character and symbols at certain places in a text in order to convey information about its linguistic components (i.e. PoS-tagging, phonological and morphological markup).
	Markup/text encoding - descriptive - nominal	Refers to the addition of characters and symbols at certain places in a text in order to convey information about concrete and abstract named entities (i.e. identification of personal names, geographic names).
	Markup/text encoding - presentational	Refers to the addition of character and symbols at certain places in a text in order to convey information about its visual appearance.
	Markup/text encoding - referential	Refers to the addition of character and symbols at certain places in a text in order to convey information about external elements (i.e. hyperlinks, hypermedia).
	Lemmatisation	Refers to techniques used to normalise the form of a word to its lemmata.
	Image enhancement	Refers to techniques used to improve the appearance of an image by using techniques such as edge enhancement, grey-scale manipulation, smoothing, and sharpening.

Analysis	Method Term	Definition
	Searching/querying	Refers to the extraction of information from data by means of query languages.
	Collating	Refers to techniques used to bring together and compare two or more variant manuscripts of the same text.
	Collocating	Refers to the techniques used to detect patterns of words appearing together in a text.
	Concording	Refers to techniques used to generate indexes of words in a text.
	Content analysis	Refers to techniques used to analyse a text by breaking it into thematic and/or conceptual units following systematic rules.
	Parsing	Refers to the process of making a syntactic analysis of a text.
	Stemmatics/cladistics	Refers to techniques used to reconstruct the transmission of a text on the basis of relations between the various surviving manuscripts.
	Stylometrics	Refers to the techniques used for the quantitative examination of textual styles. Often used as a methods for authorship attribution studies.
	Topic Detection and Tracking	Refers to the discovery and the threading together of topically related material in streams of data by algorithmic means.
	Visual analysis/visualisation	Refers to techniques used to summarise, analyse and present data visually.
	Image feature measurement	Refers to the image analytical techniques used to measure parameters of an image object such as size, shapes, relative locations, textures, grey tones, colours, etc.
	Image segmentation	Refers to the image analytical techniques used to isolate and highlight features of interest.
Dissemination	Traditional/paper-based	Refers to the dissemination of a resource by traditional, usually paper-based, means, including publications of books, print editions, or paper-based journals.
	Electronic – internet	Refers to the displaying, presenting or hosting of a resource on the internet.
	Electronic – repository/archive	Refers to the dissemination of a resource via a digital repository, archive or library such as the AHDS.
	Electronic – other medium	Refers to the dissemination of a resource by electronic means other than internet dissemination or dissemination via a digital repository. This includes, for instance, the authoring and dissemination of CD-Roms or DVDs.

Resource type “Dataset/structured data and text”

	Method Term	Definition
Capture	Manual transcription	Data capture by means of a manual transcription/inputting process
	Scanning/photography - 2D	Data capture by means of a 2-dimensional scanner (I.e. flatbed scanner, film/slide scanner, drum scanner), or a digital camera.
	Text recognition	i.e. OCR, refers to the conversion of scanned in images of a text into text documents.
	Geophysical survey	Data capture by means of geophysical methods including seismic, magnetic, electromagnetic and gravity techniques.
	Digital motion capture	Capture of data on an object's movement using various methods including optical and magnetic techniques.
	Usage of existing digital data	Refers to the usage of data that already exists in digital form
Structure	Data modelling - flat/rectangular	Refers to the development of a theoretical framework by which data are structured for the use in a database system.
	Data modelling - relational	Refers to the development of a theoretical framework by which data are structured for the use in a database system.
	Data modelling - object oriented	Refers to the development of a theoretical framework by which data are structured for the use in a database system.
	Data modelling - network	Refers to the development of a theoretical framework by which data are structured for the use in a database system.
	Coding/standardisation	Refers to the process of translating ambiguous source data into standardized codes for data processing.
	Record linkages	Refers to techniques used to link records from different sources. The most common example is nominal record-linkage which is a method for linking people across time from different sources, based on their names and other attributes.
Analysis	Record linkages	Refers to techniques used to link records from different sources. The most common example is nominal record-linkage which is a method for linking people across time from different sources, based on their names and other attributes.
	Searching/querying	Refers to the extraction of information from a structured by means of query languages
	Visual analysis/Visualisation	Refers to techniques used to summarise, analyse and present data visually
	Descriptive statistics	Refers to the analysis of data by means of descriptive statistics
	Inference statistics	Refers to the analysis of data by means of inference statistics

Method Term

Data mining

Definition

Refers to the analysis of large amounts of data in order to extract new kinds of useful information such as implicit relationships between different pieces of information, patterns, etc.

Dissemination Electronic – internet

Refers to the displaying, presenting or hosting of a resource onto the internet.

Electronic – repository/archive

Refers to the dissemination of a resource via a digital repository, archive or library such as the AHDS.

Electronic – other medium

Refers to the dissemination of a resource by electronic means other than internet dissemination or dissemination via a digital repository. This includes, for instance, the authoring and dissemination of CD-Roms or DVDs.

Resource type “Image/graphics –still”

	Method Term	Definition
Capture	Scanning/photography - 2D	Data capture by means of a 2-dimensional scanner (i.e. flatbed scanner, film/slide scanner, drum scanner), or a digital camera.
	Digital moving image capture	Data capture by means of digital video cameras, webcams and TV cards.
	Digitising tablet/table	Data capture by means of a digitising table/tablet
	Digital remote sensing	Data capture by means of satellite-based sensors
	Heads-up digitising/Interactive tracing	Data capture by means of manually tracing spatial features of an image
	Usage of existing digital data	Refers to the usage of data that already exists in digital form
Structure	Photogrammetry	Refers to the techniques used to obtain reliable measurements or information from photographs
	Image restoration and rectification	Refers to techniques used to rectify known, measured or accurately surmised degradations of images, including geometric distortions, improper foci or repetitive noises.
	Image enhancement	Refers to techniques used to improve the appearance of an image by using techniques such as edge enhancement, grey-scale manipulate on, smoothing, and sharpening.
	2-d modelling - raster	Refers to the design of 2-dimensional representations/reconstructions of objects using a raster data model
	2-d modelling - vector	Refers to the design of 2-dimensional representations/reconstructions of objects using a vector data model
	Graphical rendering	The process of computing a graphical model's surface qualities, such as colour, shading, smoothness and texture.
Analysis	Content-based image searching/retrieval	Refers to techniques used to search and retrieve images by features of their content.
	Image segmentation	Refers to the image analytical techniques used to isolate and highlight features of interest
	Image feature measurement	Refers to the image analytical techniques used to measure parameters of an image object such as size, shapes, relative locations, textures, grey tones, colours, etc.
	Overlaying	Refers to the techniques used to produce a geometric intersection between two set of data to highlight features of interest.
Dissemination	Traditional/paper-based	Refers to the dissemination of a resource by traditional, usually paper-based, means, including publications of books, print editions, or paper-based journals.

Method Term

Electronic – internet

Electronic – repository/archive

Electronic – other medium

Definition

Refers to the displaying, presenting or hosting of a resource onto the internet.

Refers to the dissemination of a resource via a digital repository, archive or library such as the AHDS.

Refers to the dissemination of a resource by electronic means other than internet dissemination or dissemination via a digital repository. This includes, for instance, the authoring and dissemination of CD-Roms or DVDs.

Resource type “Image-moving”

	Method Term	Definition
Capture	Digital moving image capture	Data capture by means of digital video cameras, webcams and TV cards.
	Digital motion capture	Capture of data on an object's movement using various methods including optical and magnetic techniques
	Scanning/photography - 2D	Data capture by means of a 2-dimensional scanner (I.e. flatbed scanner, film/slide scanner, drum scanner), or a digital camera.
	Usage of existing digital data	Refers to the usage of data that already exists in digital form
Structure	Video editing	Refers to the techniques used to manipulate moving image data into a defined structure.
	Moving image compression	Refers to techniques used to eliminate redundant information of moving image files in order to reduce its size
	Digital animation	Refers to techniques used to create a timed sequence or series of graphic images or frames together to give the appearance of continuous movement.
Analysis	Content-based image searching/retrieval	Refers to techniques used to retrieve moving images by features of their content.
	Motion analysis	
Dissemination	Electronic – internet	Refers to the displaying, presenting or hosting of a resource on the internet.
	Electronic – repository/archive	Refers to the dissemination of a resource via a digital repository, archive or library such as the AHDS.
	Electronic – other medium	Refers to the dissemination of a resource by electronic means other than internet dissemination or dissemination via a digital repository. This includes, for instance, the authoring and dissemination of CD-Roms or DVDs.

Resource type “3-d object”

	Method Term	Definition
Capture	Scanning/photography - 2D	Data capture by means of a 2-dimensional scanner (i.e. flatbed scanner, film/slide scanner, drum scanner), or a digital camera.
	Scanning - 3D	Data capture by means of a 3-dimensional scanner
	Digitising tablet/tablet	Data capture by means of a digitising table/tablet
	Heads-up digitising/Interactive tracing	Data capture by means of manually tracing spatial features of an image
	Usage of existing digital data	Refers to the usage of data that already exists in digital form
	Structure	Photogrammetry
Image restoration and rectification		Refers to techniques used to rectify known, measured or accurately surmised degradations of images, including geometric distortions, improper foci or repetitive noises.
Image enhancement		Refers to techniques used to improve the appearance of an image by using techniques such as edge enhancement, grey-scale manipulation, smoothing, and sharpening.
3d modelling - vector		Refers to the design of 3-dimensional representations/reconstructions of objects or structures using a vector data model and specialised software such as CAD and GIS packages
3d modelling - vector - point cloud		Refers to the design of 3-dimensional representations/reconstructions of objects or structures using a vector data model and specialised software such as CAD and GIS packages.
3d modelling - vector - triangular network models		Refers to the design of 3-dimensional representations/reconstructions of objects or structures using a vector data model and specialised software such as CAD and GIS packages
3d modelling - vector - wireframe		Refers to the design of 3-dimensional representations/reconstructions of objects or structures using a vector data model and specialised software such as CAD and GIS packages
3d modelling - vector - surface		Refers to the design of 3-dimensional representations/reconstructions of objects or structures using a vector data model and specialised software such as CAD and GIS packages
3d modelling - vector - volumetric/solid		Refers to the design of 3-dimensional representations/reconstructions of objects or structures using a vector data model and specialised software such as CAD and GIS packages
3d modelling - interactive		i.e. Virtual Reality. Refers to the design of interactive 3-dimensional graphical representations/reconstructions of objects.
	Graphical rendering	The process of computing a graphical model's surface qualities, such as colour, shading, smoothness and texture.

	Method Term	Definition
Analysis	Design analysis - acoustic analysis	Refers to the behaviour and/or performance analysis of a design in different environments,
	Design analysis - stress analysis	Refers to the behaviour and/or performance analysis of a design in different environments,
	Design analysis - thermal analysis	Refers to the behaviour and/or performance analysis of a design in different environments,
	Virtual quantity surveying (volumetric analysis)	Refers to the mathematical determination of the volume of an object or structure.
	Overlaying	Refers to the techniques used to produce a geometric intersection between two set of data to highlight features of interest.
	Image segmentation	Refers to the image analytical techniques used to isolate and highlight features of interest
	Image feature measurement	Refers to the image analytical techniques used to measure parameters of an image object such as size, shapes, relative locations, textures, grey tones, colours, etc.
Dissemination	Electronic – internet	Refers to the displaying, presenting or hosting of a resource onto the internet.
	Electronic – repository/archive	Refers to the dissemination of a resource via a digital repository, archive or library such as the AHDS.
	Electronic – other medium	Refers to the dissemination of a resource by electronic means other than internet dissemination or dissemination via a digital repository. This includes, for instance, the authoring and dissemination of CD-Roms or DVDs.

Resource type “Spatial data”

	Method Term	Definition
Capture	GPS/total station surveys	Refers to the capturing of spatial information by means of Global Positioning Systems or Total Station equipment
	Digital remote sensing	Data capture by means of satellite-based sensors
	Geophysical survey	Data capture by means of geophysical methods including seismic, magnetic, electromagnetic and gravity techniques.
	Digital aerial photography	Data capture by means of digital cameras from an aerial platform
	Scanning/photography - 2D	Data capture by means of a 2-dimensional scanner (i.e. flatbed scanner, film/slide scanner, drum scanner), or a digital camera.
	Scanning - 3D	Data capture by means of a 3-dimensional scanner
	Digitising tablet/table	Data capture by means of a digitising table/tablet
	Heads-up digitising/Interactive tracing	Data capture by means of manually tracing spatial features of an image
	Usage of existing digital data	Refers to the usage of data that already exists in digital form
Structure	Geo-referencing/projection	Refers to techniques used to convert images from image coordinates to real-world coordinates
	Photogrammetry	Refers to the techniques used to obtain reliable measurements or information from photographs
	Image restoration and rectification	Refers to techniques used to rectify known, measured or accurately surmised degradations of images, including geometric distortions, improper foci or repetitive noises.
	Image enhancement	Refers to techniques used to improve the appearance of an image by using techniques such as edge enhancement, grey-scale manipulation, smoothing, and sharpening.
	2d modelling - raster	Refers to the design of 2-dimensional representations/reconstructions of objects or structures using a raster data model and specialised software such as CAD and GIS packages
	2d modelling - vector	Refers to the design of 2-dimensional representations/reconstructions of objects or structures using a vector data model and specialised software such as CAD and GIS packages
	3d modelling - vector	Refers to the design of 3-dimensional representations/reconstructions of objects or structures using a vector data model and specialised software such as CAD and GIS packages
	3d modelling - vector - point cloud	Refers to the design of 3-dimensional representations/reconstructions of objects or structures using a vector data model and specialised software such as CAD and GIS packages.
	3d modelling - vector - triangular network models	Refers to the design of 3-dimensional representations/reconstructions of objects or structures using a vector data model and specialised software such as CAD and GIS packages
	3d modelling - vector - wireframe	Refers to the design of 3-dimensional representations/reconstructions of objects or structures using a vector data model and specialised software such as CAD and GIS packages

Method Term

3d modelling - vector - surface

Definition

Refers to the design of 3-dimensional representations/reconstructions of objects or structures using a vector data model and specialised software such as CAD and GIS packages

3d modelling - vector - volumetric/solid

Refers to the design of 3-dimensional representations/reconstructions of objects or structures using a vector data model and specialised software such as CAD and GIS packages

3d modelling - interactive

i.e. Virtual Reality. Refers to the design of interactive 3-dimensional graphical representations/reconstructions of objects.

Graphical rendering

The process of computing a graphical model's surface qualities, such as colour, shading, smoothness and texture.

Spatial filtering/smoothing

A technique used to remove or reduce local noise or high frequency signal within spatial data, and therefore reveal the global pattern or trend.

Analysis

Accessibility analysis

Refers to the determination of an aggregate measure of how reachable locations are from a given location.

Cost/friction analysis

Refers to the determination of a distance measure based on the minimisation of friction or cost for a single path or an entire surface.

Kriging

Refers to a form of statistical modelling that interpolates data from a known set of sample points to a continuous surface.

Line-of-sight analysis

i.e. viewshed analysis. Refers to the determination of the intervisibility of different locations in a spatial model.

Network analysis

Refers to the analysis of flows along topologically linked data. For example, identification of the shortest path between two locations on a road network, or identification of the least expensive road to take for transporting goods between those locations.

Overlaying

Refers to the techniques used to produce a geometric intersection between two set of data to highlight features of interest.

Predictive spatial modelling

i.e. of archaeological site locations. Refers to the establishment of a set of testable hypotheses, which attempts to predict locations of interest in a spatial model.

Searching/querying

Refers to the extraction of information from data by means of query languages.

Spatial buffering

i.e. Proximity analysis. Refers to the determination of a zone of a specified distance around a spatial feature.

Spatial filtering/smoothing

A technique used to remove or reduce local noise or high frequency signal within spatial data, and therefore reveal the global pattern or trend.

Thiessen polygon analysis

Refers to the calculation of the influence of different centres of human activity based on a number of variables.

X-tent principle

Refers to the inference of the size of a centre's territory from settlement size, population and storage capacity.

Visual analysis/visualisation

Refers to techniques used to summarise, analyse and present data visually.

Dissemination Electronic – internet

Refers to the displaying, presenting or hosting of a resource onto the internet.

Electronic – repository/archive

Refers to the dissemination of a resource via a digital repository, archive or library such as the AHDS.

Electronic – other medium

Refers to the dissemination of a resource by electronic means other than internet dissemination or dissemination via a digital repository. This includes, for instance, the authoring and dissemination of CD-Roms or DVDs.

Resource type: “Sound/Audio”

	Method Term	Definition
Capture	Digital sound recording	Refers to data capture by means of sound recording equipment, involving an analogue-to-digital conversion
	Digital sound generation	Generation of sound by means of digital instruments
	Digital moving image capture	Data capture by means of digital video cameras, webcams and TV cards.
	Scanning/photography - 2D	Data capture by means of a 2-dimensional scanner (i.e. flatbed scanner, film/slide scanner, drum scanner), or a digital camera.
	Music recognition	i.e. OMR, refers to the capturing of musical semantics from a scanned in page of musical notations
	Usage of existing digital data	Refers to the usage of data that already exists in digital form
Structure	Sound editing	Refers to the techniques used to manipulate sound data into a defined structure.
	Sound compression	Refers to techniques used to eliminate redundant information of a sound files in order to reduce its size.
	Sound encoding	
	Sound encoding – MIDI	
Analysis	Content-based sound searching/retrieving	Refers to techniques used to retrieve sound by features of its content.
	Sound analysis	
Dissemination	Electronic – internet	Refers to the displaying, presenting or hosting of a resource onto the internet.
	Electronic – repository/archive	Refers to the dissemination of a resource via a digital repository, archive or library such as the AHDS.
	Electronic – other medium	Refers to the dissemination of a resource by electronic means other than internet dissemination or dissemination via a digital repository. This includes, for instance, the authoring and dissemination of CD-Roms or DVDs.