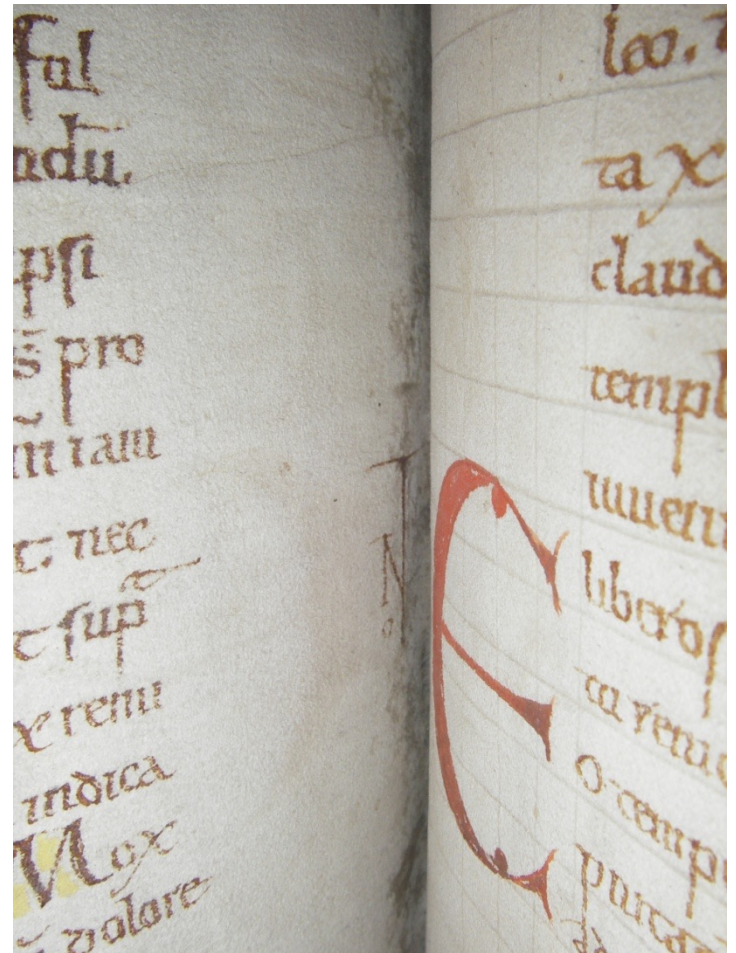


Cultural heritage digitisation

concepts
strategies
levels

Mats Dahlström
January 20, 2022



Concepts

- turning something from analog to digital
- digitalisation – digitisation
- digitalisering – digitisering
- digitalisat

WHY DIGITIZE?

Motives for digitizing cultural heritage material

- Preservation – reducing the wear and tear on the originals
- Make the material more accessible
- Make the material usable
- Make the material re-usable
- Develop and share technology and competence
- Enable new research
- Present the material in new ways and in new contexts
- Administrative, economical or political reasons
- Prestige and propaganda

The long chain of digitization

- Planning and budget
- Selecting, collecting, managing
- Document analysis and preparation
- Strategies
- Tools (selecting, acquiring, developing)
- Text and image capture (scanning, digital camera)
- OCR and other technologies
- Text encoding and other optimization, e.g. indexing
- Testing, validating
- Metadata
- Selecting delivery media and distribution forms/channels
- Publication and distribution
- Marketing
- Evaluation, documentation and reports
- Archiving, migrating
- Maintenance

Examples of overall strategies

- source media / new media “fidelity”
- long term / short term digitization
- first hand / second hand
- proactive / reactive digitization
- conservation / restoration digitization
- text digitization / image digitization

Text digitization versus image digitization

Linguistic text	Graphical and material artefacts (document)
Machine readable transcriptions	Digital facsimiles
Text capture (OCR)	Scanning, photography
Text encoding	Image editing
Text files	Image files
Might have a small "facsimile" dimension	Often concentrated to the document parts which carry text

Faces of William Blake – Project Gutenberg (PG)



gutenberg.org



THE ECHOING GREEN

The sun does arise,
And make happy the skies;
The merry bells ring
To welcome the Spring;
The skylark and thrush,
The birds of the bush,
Sing louder around
To the bells' cheerful sound;
While our sports shall be seen
On the echoing green.

Old John, with white hair,
Does laugh away care,
Sitting under the oak,
Among the old folk.
They laugh at our play,
And soon they all say,
'Such, such were the joys
When we all—girls and boys—
In our youth-time were seen
On the echoing green.'

Till the little ones, weary,
No more can be merry:
The sun does descend,
And our sports have an end.
Round the laps of their mothers
Many sisters and brothers,
Like birds in their nest,
Are ready for rest,
And sport no more seen
On the darkening green.

Faces of William Blake – The William Blake Archive (BA)

blakearchive.org

Songs of Innocence and Songs of Experience, by William Blake

THE WILLIAM BLAKE ARCHIVE GALLERY MODE

Songs of Innocence and of Experience Copy A (Printed 1795) (Selected)


Search IT G R ? 0 x

100 DPI 300 DPI

Songs of Innocence and of Experience Copy A

Songs of Innocence Copy B

Songs of Innocence Copy G



Object 6 (Bentley 6, Erdman 6, Keynes 6), 1789, British Museum

Object 14 (Bentley 6, Erdman 6, Keynes 6), 1789, Library Of Congress

Object 10 (Bentley 6, Erdman 6, Keynes 6), 1789, Yale Center For British

Add To Lightbox Rotate Magnify True Size Enlargement Diplomatic Transcription

Objects in Copy Objects from the Same Matrix Copy Information Electronic Edition Information

The Literary Bank, Strindberg section

August Strindberg, Verk i Litter x

litteraturbanken.se/#1/forfattare/StrindbergA/titlar

Appar bibl o kat blogg o hemsidor handel o nöjen hus o hem jobb media PP FB Google HB BSCW DIXIT gd tfu Övriga bokmärken

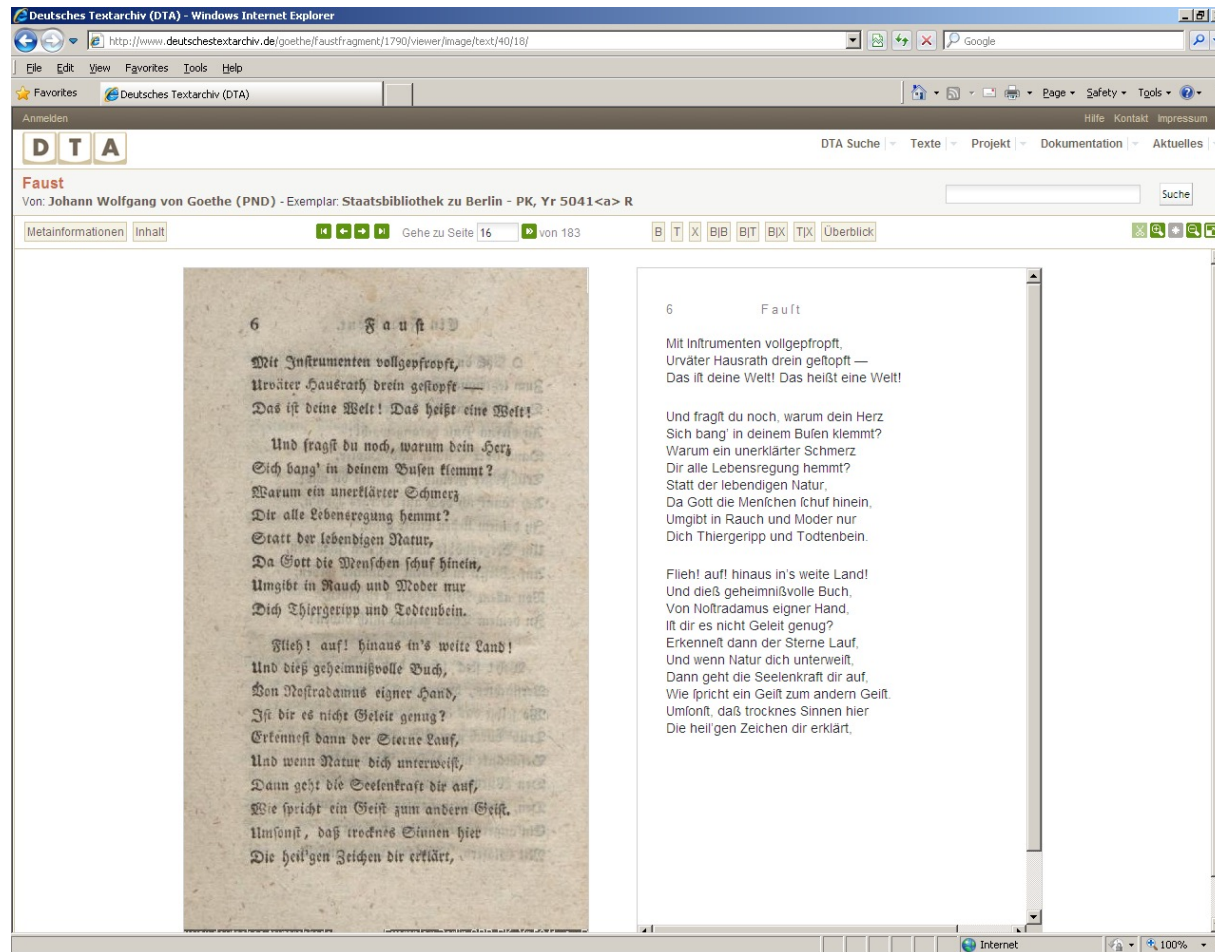
August Strindberg (1849-1912)

INTRODUKTION ★ VERK ★ MER ★ SÖK I TEXTERNA

Tillgängliga verk

BIBLIOTEKET	ETEXT EPUB	Abu Casems tofflor
EPUBFILER	ETEXT EPUB	Bland franska bönder
LJUDARKIVET	FAKSIMIL	Blomstermålningar och djurstycken [1888]
SNABBSÖKNING	FAKSIMIL	Brott och brott [ms]
SÖK I TEXTERNA	FAKSIMIL	Brända tomten [ms]
PRESENTATIONER	FAKSIMIL	Czarens Kurir eller Sågfilarens Hemligheter [annoterat exemplar]
NYTILLKOMMET	FAKSIMIL	Den starkare. En scen [1890]
LAGERLÖFARKIVET	ETEXT EPUB	Det nya riket
SKOLAN	FAKSIMIL	Det nya riket. Skildringar från Attentatens och Jubelfesternas Tidevarf [1882]
OM LB	FAKSIMIL	Dikter på vers och prosa [1883]
	ETEXT EPUB	Dikter på vers och prosa
	ETEXT EPUB	En blå bok I
	ETEXT EPUB	En blå bok II
	FAKSIMIL	En brevväxling mellan Nietzsche och Strindberg
	ETEXT EPUB	En dåres försvarstal

Digital image and text mode side by side (Deutsches Textarchiv)



The van Gogh letters

008 (008, 8): To Theo van Gogh. The Hague, Friday, 9 May 1873. — Vincent van Gogh Le...

Van Gogh Museum **huysen ing** Help | Quick Guide | Updates | Credits | Contact | Home

Vincent van Gogh
The Letters

by period
by correspondent
by place
with sketches

Search
keyword or number(s) >>
Advanced search
Search results

Van Gogh as a letter-writer
Correspondents
Biographical & historical context
Publication history

About this edition
Chronology
Concordance, lists, bibliography
Book edition

008 013 « 007 | 009 »

To Theo van Gogh. The Hague, Friday, 9 May 1873. SEARCH THIS LETTER PRINT

original text + line endings facsimile translation notes artworks

1873
den Haag 9 Mei 1873

008
Br. 1990: 008 | CL: 8
From: Vincent van Gogh
To: Theo van Gogh
Date: The Hague, Friday, 9 May 1873
[more...](#)

Waarde Theo,
Maandag morgen vertrek ik van Helvoirt naar Parijs & passeer Brussel om 2 uur 7 min.— 2 Als je kunt kom dan aan 't station, dat zal mij veel genoeg doen.
Voor ik het vergeet; gisteren 1873 heb ik aan Mevr. Tersteeg je portret laten zien 2 & die zou er zoo heel graag een hebben. Zou je soms kans zien er nog een te krijgen & mij te sturen.
Als het nu niet kan denk er dan later eens om.—
Theo je kunt niet begrijpen hoe lief allen hier voor mij zijn; 3 dat het mij spijlt van zooveel vrienden te moeten scheiden kan je wel denken.—
1873
Adieu kerel, groet Oom & Tante 4 hartelijk voor mij. tot ziens.
Vincent.
Informeer eens of je aan 't Station du Nord of du Midi moet wezen.— 5

original text + line endings facsimile translation notes artworks

original text + line endings facsimile translation notes artworks

1. Van Gogh was planning to be in Helvoirt by Saturday evening, 10 May to take leave of his parents (FR b2626); he left Helvoirt on Monday, 12 May. On his way to London he stopped off in Paris, where he spent nearly a week.

2. Maria Magdalena Alida Tersteeg-Pronk, wife of Van Gogh's superior Hermanus Gijbertus Tersteeg. For this portrait, cf. letter 5, n. 5.

3. Mrs van Gogh wrote to Theo on 11 May, when Vincent was in Helvoirt: 'It's wonderful to hear Vincent telling about how kind everyone was to him and about his leave-taking from The Hague. He was given an album and so many little presents and portraits, two very large, splendid portraits of Uncle and Aunt Haanebeek, Mr and Mrs Tersteeg and Betsy, Aunt Fietje and the girls, and others' (FR b2626).

4. Uncle Hein and Aunt Mietje in Brussels.

5. Van Gogh took the train that went to Paris via Rotterdam, Antwerp and Brussels. According to the 1873 timetable, the train arrived at 14.06 at Brussels North; the connection to Paris left from Brussels South at 14.35 and arrived in Paris at 21.05 (Brussels, Archives of the Belgian Railways).

top ↗

physical description →

mass digitization

	Primarily automated
	Breadth
	Minimized editing and metadata
	Clone
	Random and totality
	Text
	Quantitative

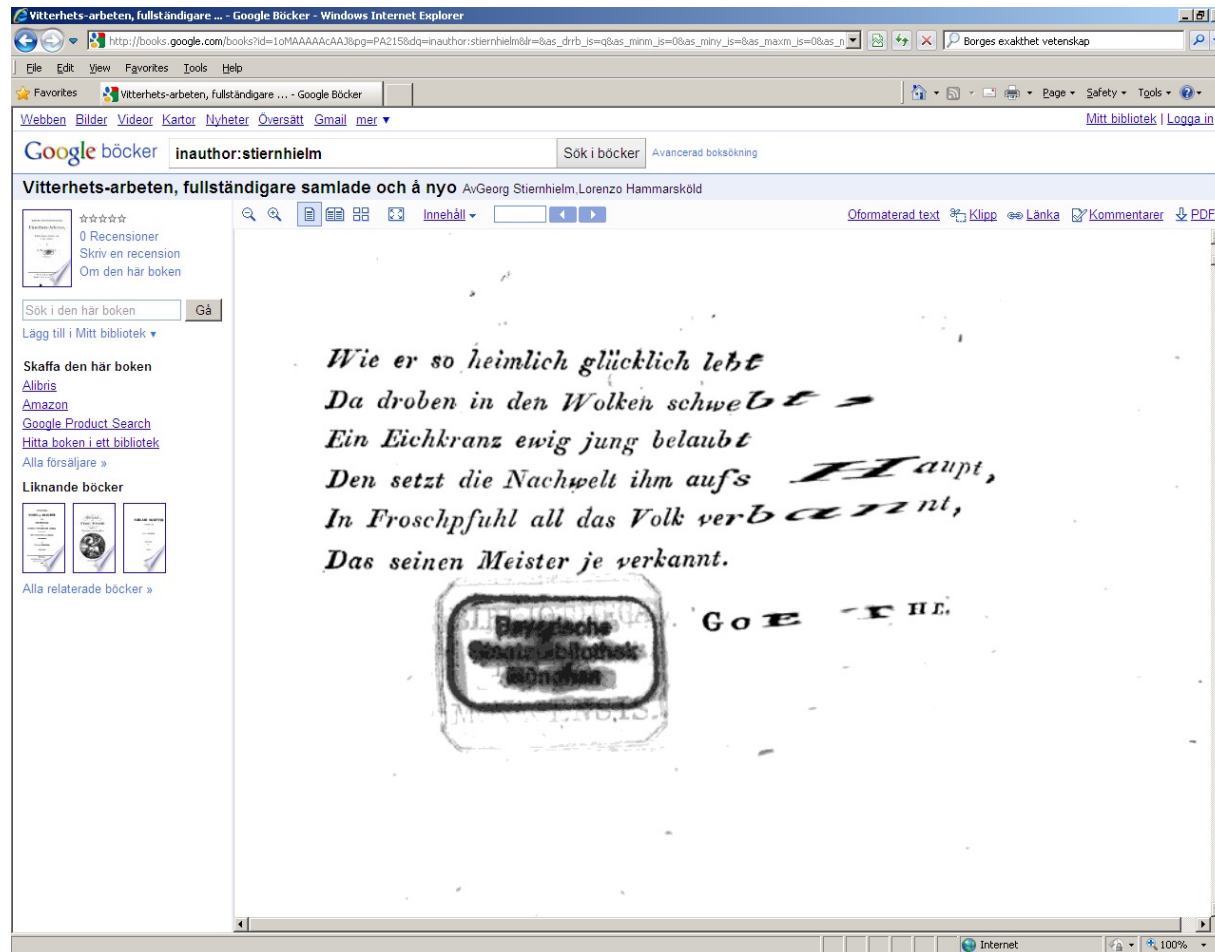
Google Books, Open Content Alliance (via Internet Archive),
HATHI Trust, Europeana, Gallica

Mass (or at least large-scale) digitisation projects

Mass digitization: criticism

- Poor scanning quality
- Transcription poor or even useless
- Often only simple, raw free text search
- Few and often erroneous metadata
- Rudimentary bibliographical information
- Sometimes difficult or even impossible to ascertain the provenance of the digital representation – what do we actually see on the screen?
- Non-transparent contracts and terms conc. copyright, re-use, long-term availability and maintenance of the digital files

An early Stiernhielm in Google Books



Scanned book on perception,
images and form

Critical vs mass digitization

Primarily manual	Primarily automated
Depth	Breadth
Maximized editing and metadata	Minimized editing and metadata
Interpretation	Clone
Critical selection	Random and totality
Document	Text
Qualitative	Quantitative

Codex Sinaiticus

"standard light"

Codex Sinaiticus - See The ...

www.codexsinaiticus.org/en/manuscript.aspx?book=34&chapter=2&lid=en&side=r&verse=19&zoomSlider=0

bibl o kat blogg o hemisor handel o nojen hus o hem jobb media Ullrichams kommuns... f FB PP H8 Google

Övriga bokmärken

Contact | Copyright

Advanced search →

HOME ABOUT CODEX SINAITICUS ABOUT THE PROJECT SEE THE MANUSCRIPT

GO TO: (Book)

choose a passage: Mark Chapter 2 Verse 19

choose a page: Quire: 76 Folio: 3 r

display options: ☒ Image ☒ Translation ☒ Transcription ☒ Physical Description

Mark, 1:35 - 2:19 library: BL folio: 218 scribe: A

Image Standard Light Raking Light i

Transcription view by verse view by page i

Translation Русский Ελληνικά Deutsch English i

1:35 And very early while yet night, he rose up and went out, and went away to a desert place, and there prayed.

36 And Simon and they that were with him followed after him.

37 And found him, and said to him: All seek for thee.

38 And he said to them: Let us go elsewhere to the next

Supported by DFG Arts & Humanities Research Council

Codex Sinaiticus

"raking light"

Codex Sinaiticus - See The ...

www.codexsinaiticus.org/en/manuscript.aspx?book=34&chapter=2&lid=en&side=r&verse=19&zoomSlider=0

bibl o kat blogg o hemsidor handel o röjnen hus o hem jobb media Ulicehamns kommuns... FB PP HB Google Övriga bokmärken

Contact | Copyright Advanced search →

HOME ABOUT CODEX SINAITICUS ABOUT THE PROJECT SEE THE MANUSCRIPT GO TO: (Book)

choose a passage: Mark Chapter 2 Verse 19 choose a page: Quire: 76 Folio: 3 display options: Image Translation Transcription Physical Description

Mark, 1:35 - 2:19 library: BL folio: 218 scribe: A

Image Standard Light Raking Light i

Transcription view by verse view by page i

Translation Русский Ελληνικά Deutsch English i

1:35 And very early while yet night, he rose up and went out, and went away to a desert place, and there prayed.

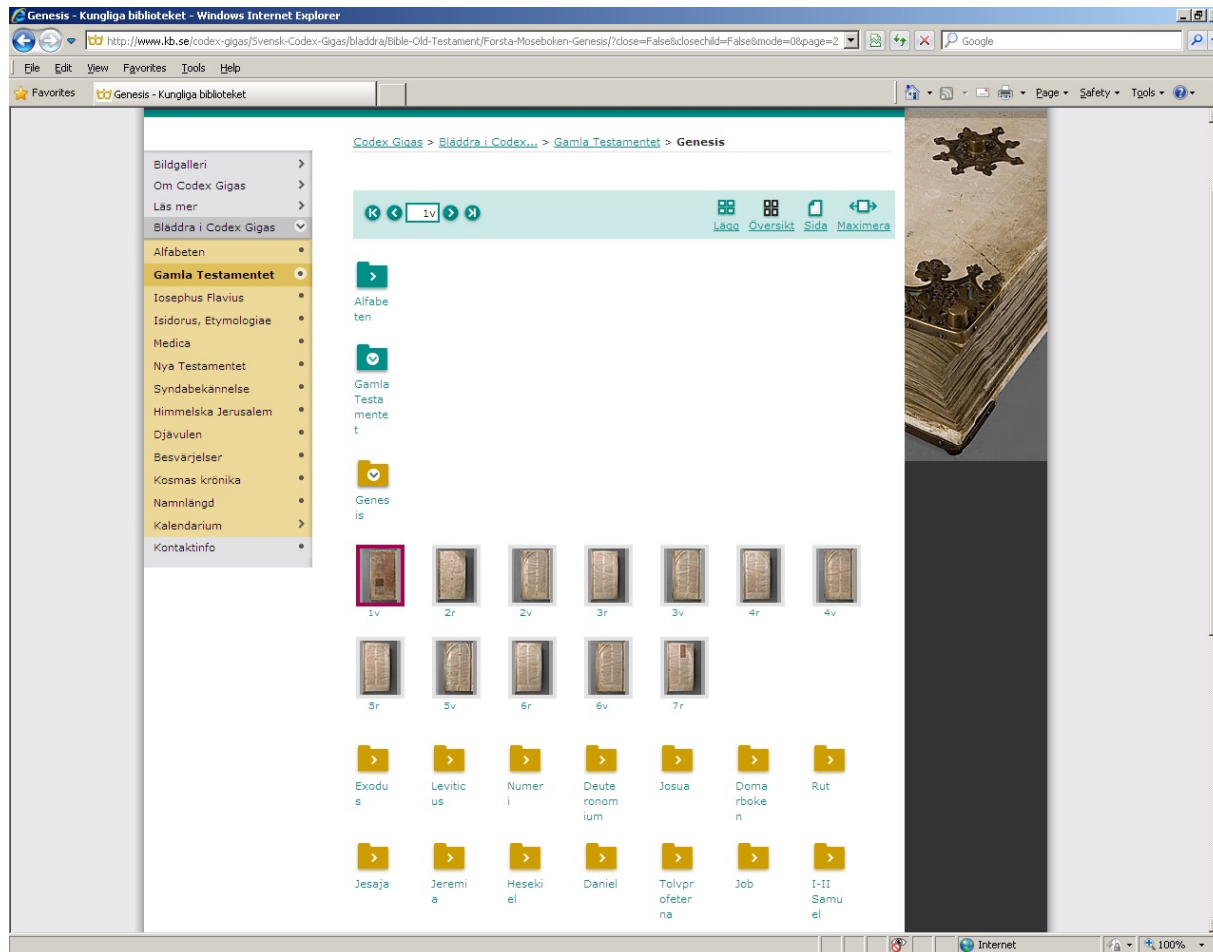
36 And Simon and they that were with him followed after him.

37 And found him, and said to him: All seek for thee.

38 And he said to them: Let us go elsewhere to the next

Supported by DFG Arts & Humanities Research Council ANU THE A.S. LINDSAY FOUNDATION RSS

Codex Gigas



Critical digitization: criticism

- Costly
- Slow
- Small audiences
- Necessitates skills in e.g. textual, historical or bibliographical scholarship
- Often neglects to produce machine readable transcriptions
- Sometimes lacks in documentation and "repeatability"
- The material has been explicitly interpreted and narrowly formed for a particular purpose
- A cloning ideal?

Lecture on scholarship and access

- advanced digitisation
- digitisation and scholarship
- access
- use
- re-use

Some references

- Dahlström, M. (2015). Critical Transmission. In: P. Svensson & D. T. Goldberg (eds.), *Between Humanities and the Digital*. Cambridge: Mass., The MIT Press, p. 467-481.
- Dahlström, M. and W. Dillen (2017). Litteraturbanken: the Swedish Literature Bank. *RIDE* 6. doi: 10.18716/ride.a.6.2. <https://ride.i-d-e.de/issues/issue-6/litteraturbanken-the-swedishliterature-bank/>
- Deegan, Marilyn & Kathryn Sutherland (2009). *Transferred Illusions: Digital Technology and the Forms of Print*. Aldershot: Ashgate.
- Duffy, Christina (2015). *Revealing the secrets of the burnt Magna Carta*. The British Library. <https://www.bl.uk/magna-carta/articles/revealing-the-secrets-of-the-burnt-magna-carta>
- Pierazzo, Elena (2014). Digital Documentary Editions and the Others. *Scholarly Editing*, 35. <http://scholarlyediting.org/2014/essays/essay.pierazzo.html>
- Prescott, Andrew & Lorna Hughes (2018). Why Do We Digitize? The Case for Slow Digitization. *Archive Journal*. <http://www.archivejournal.net/essays/why-do-we-digitize-the-case-for-slow-digitization/>