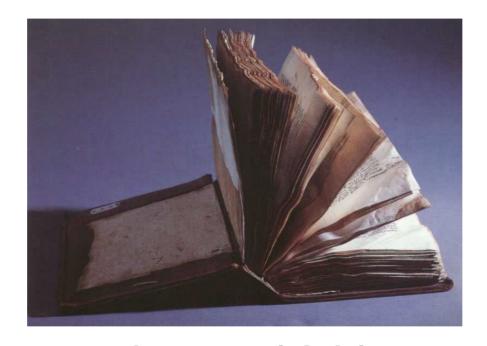
# The Metaphysics of Documents



Claus Huitfeldt DHC Kolloquium, Uni Köln 06.07.2023



## Structure of this talk

- Documents, texts, and abstract objects
- A problem: Trilemmas
- Some solutions to the trilemmas
- Digital documents
- Transcription
- Conclusion: Documents are created types



### "General" conclusion first:

- Metaphysics of documents is important to DH
  - We work with documents, and care about their:
    - Identity
    - Similarity
    - Retrieval
    - Format conversion
    - Version control
    - Authenticity



- In a wide sense:
  - Writings, images, video (film), music, monuments, buildings,...
- Here:
  - Fixations of natural verbal language in written form



- In a wide sense:
  - Writings, images, video (film), sound, monuments, buildings,...
- Here:
  - Fixations of natural verbal language in written form
- But what about:
  - Written documents containing images, video, etc.
  - Music notation, dance notation, math, architectural drawings
  - Signing language, Braille, Morse, semaphore



- In a wide sense:
  - Writings, images, video (film), sound, monuments, buildings,...
- Here:
  - Fixations of natural verbal language in written form
- But what about:
  - Written documents containing images, video, etc.
  - Music notation, dance notation, math, architectural drawings
  - Signing language, Braille, Morse, semaphore
    - Fine, but one thing at a time...



## It is complicated enough as it is:



enten prøvde nå yrkesrettleieren bent frem å overtale psykologen til å avgi en rapport som kunne sette døren på gløtt for et forsøk i retning av bilmekanikeropplæring for gutten.

Kanskje var det et slags instinkt, snarere enn fornuftige overveielser, som lå bak min beslutning i denne saken. I korthet er i alle tilfeller den videre historie denne:

Yrkesrettleieren fikk en uttalelse han kunne bruke til å søke gutten inn ved skolen. Gutten fikk plass, og klarte til alle utenforståendes overraskelse å gjennomføre skolegangen med et godkient resultat. En nærmere forklaring på dette var på sin plass. Og en forklaring fikk man. Yrkesrettleieren kunne fortelle at overhevisningen om at dette var veien for ham, var så sterk has vår venn at han ikke et gvehlikk tapte motet. All hans tid gikk med til å arbeide med fagene på skolen. Til og med de siste minutter før han falt i søyn om kvelden, ble anvendt til "studieformål", der han igjen og igjen gjennomgikk de plansjer og arbeidstegninger som han hadde spikret fast i taket over køven! Få skoler eller studier er så vanskelige at en

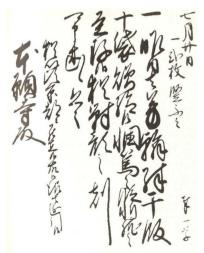
slik interesse og arbeidsmotivasjon ikke hol-

der det meste av veien. Tro kan virkelig flytte fiell, også i vrkesmessig sammenheng.

#### "Vi er ikke alle like"

Vi lever vel alle med en klar forståelse for at "det er forskiell på folk". Fra tidlig alder skjer det sammenligninger med andre, med våre søsken, skolekamerater, ja, i korthet nærmest alle som vi omgås eller er i slekt med. Og ikke alltid kommer vi fra slike sammenligninger med et "hederlig resultat". Nærmest til enhver tid er det noen som er flinkere eller snillere enn oss selv, og disse vil våre omgivelser gjerne sørge for å trekke frem, som gode eksempler på hva vi bør strebe etter å oppnå. Dermed vil selv den mest iherdige eller tålmodige siel kunne få en følelse av stadig å komme til kort, det fins hele tiden noen som er bedre.

Likevel er det nok så at vi mennesker til syvende og sist blir nokså forskjellige, fordi vi i starten gierne har våre spesielle anlegg eller særegenheter, men også fordi våre omgivelser har formet oss forskjellig. Noen vil slik vi før har nevnt, kunne gå rundt med en svært klar





1 bred 2 de melle 3 H-welle Smer

### Gloford. Pfalmus XXII DeLyra. 117

. macontracta fit fine ruga. Seed quibuidam, ordo harun gra
tiatum conucrius videur, fegentes
literam conucrius videur, f fed prius educaut fuper aqua refe-ctionis, fed prius alam mea conner. d. Deduair. Caf. Quarta, quòd fuper

f Nam & fi.Caj.Quinta, quod in vin

The face of the state of the st redinace cum fin.
(vel refor h virga tua. Caf. Sexta, correctio.
cillabit) Caffod vel virga, dillridio,qua con
uerti vitia. Baculus, gubernatio,
t ambuls qua fulfinet fideles.

amoust qua fuffine fideles.

i Tundi in cospedin. Caf. Septima,
menia domini, betat consuluato,
faturias fidet, cla ceclefitade qua

2.0x.114 Apotolous Qui manducat & bibli
indigne, indicium fibi manducat
& bibli.

ane diction.

Ane diction and distinct of the first clear for a final advert preparare, sid eft past final advert preparare, sid eft past final and distinct of the first clear first preparare and distinct of the first clear first preparare and the first preparare first preparare first preparare for the first preparare first preparare for the first preparate for the first preparare for th me. m Etcalix were: Caf. Nona, calix, Chri Pfalt 21-4 fti fanguis, qui ficinebriat vt fobri-firm in reddat.

me, reddat.

au) n Mew. Caf. Quia accipio. Vel, tuus, quia das. Menía & calix, caro & fan guis quamquám. vnum fint facramentum, fub diuerfis tamen fpecie hic ponuntur.

thabitabo o Et misfericordia. Cass. Decima, mise-

O Et supersonale de l'Obermanteticodia fubéquens ad cuñodiam, in longitudine dierum .

que precedit ad gram confredd .

Lampada p Et vi indult. Hiro-Hoc omnibus fuperioribus iungitur. Q.

que procedit ad gram confredd .

Lampada p Et vi indult. Hiro-Hoc omnibus fuperioribus iungitur. Q.

que l'annual de l'annual de

NICODE LYRA.

§ the or shown with distillance, proof forth it than to see at time.

1.40, 1.40 Mines for year where we the interest of the time to see at time.

1.40, 1.40 Mines for year where we the interest to see at an entity that an anear advan
time, plaint from the arterious proof to the see at the interest to the con
tion, year the contract of the proof to the contract to the contract

me renocando. a Et nihil mihi deerit. Ex quo enim me in terram fuo cultui deditam renocanit, teneo firmiter, quòd de villu comuenienti mibi & focietat mes prouidebit .

b În loco palcuz.i.bona fafficientis vel bene fafficitis ad vinendam

In pfal.xxii. Dominus regit me.

#### ADDITIO.

Fabulatio Ra. Sa. in hoc loco nullum habet fundamentum in fa Fabulatio Ra, Sain hoc Joco millum habet inndamentum in fa-cat feripatus, imbo centratium habit quod afferis, habeten in pfall-mo.lisji qui pfalfashus fiiti in deferro landge oil bi effet fugitiums, pout in allius intulo coninerur. « bisidem enperfeté distinar, quòd Danid ear in terta defera, innia, se insquofas unde apparet lali-ra fédicion RA. So dicentis, quod la le locus abundabé rinchius de fluminibut se huiufmodi. Que quidem fichio malitios chabet.

A a le line refuse a differ Prima gratia, locus paícus, sid el lérde line coloranis. Qui infain. sente ent le cas lifetime et à laire est point de la coloranis partia, aqua refectionis.

Descriptor, est conserve de l'existic gratia, aqua refectionis,
presentant que que de l'existe de la coloranis partia, aqua refectionis,
presentant que la coloranis partia que participation anni
presentant que l'existent de la coloranis partia de l'existent neces
presentant de la coloranis de l'existent de la coloranis de l'existent neces
presentant de l'existent de l'ex

a CaS. Quin finintalem Subdensiam credit Shi a Deo esta conferendam. b Sin feutro. e Africian Icharce. "Yel centar de friente." " & mihil mi'hi decrit, fin loco pascuze, s riei mei aliquid agermus contra botoré
nominis fai, va predillam est.

i Nam & si ambulauero in medio & c.
ivel vecuid est in periculo mortes ex perfecutione
chan & c.

Saulis intra regnam funm . menis)

K. Non timebo mala . ab eo excepticas . 1 fentie l Quomiam tu mecum es, languare de fluenten feufor fortis est in boc dirigit David ver- (vel place bares faum ad ipflow Deum). w Vinga ma de baculus tuus, id est defen a fård prins educauit neimann de. Et poft. f vr pod fit fincht & infts. b Clade. f Stout publication . Con-ferencie festis pracepta charteatis vell. od boc feientie pendunit, et eius pracepta docean. e Quiban parci

m Vinga ma de bacalus truss id est defen fontes tue ; de que desgrantar naminibus bacali est vinga . o Paralti in conspectu meo mensam, i victam sufficientem . p Aductius cos. [.S.and est cius coplices. am † conuertit. Deduxir me fuper fe- = 1 tou. b \$440,No nedizament. & the vites or quelle sign. mitas iustiriæ, propter nome fuum . h

Nam & fi + ambulauero + in medio a CAug, ries huius que ell embra monte. Cufa vel cibo , fed etiem de potu convenienti .

u Et mile.tu-fuble-me, providendo mibi umbræ mortis, non timebo mala, a a (Non timebo: nam. CAUE difeipina que eft ad oues, baculus ad grandiores fillos, qui ab animali vita confesant ad fairirantem.

at bible.

\*\*A freedy-file new margin, c. 6.5. End

\*\*Tourns not improveme. De backdistants, c. 6.5. End

\*\*Internation improveme. De backdistants, c. 6.5. End

\*\*Internation improvement. De backdistants, c. 6.5. End

\*\*In the quotient to mectum es. Virga tua & \_\_\_\_\_\_

[Sea-greates sea visitaria confoldat func.

The quotient function of the part is an estable size in the sea part of the part is an estable size in the sea part of the part of

y In lon.dicrum. Moraliter. Portel ex-scent poni pfal.ifle de bomine dente fugitnte à feculo inspanne à lon periculofo, & in-trante losson religionis, maxime mendi-caritis, confidera de divine proufficer & defenfione ab aduerfarijs, nom folion cor- † tu "pingualti in oleo caput meum, & ca-

\* a Et nibil mibi deeris. De necessaris en vizz. Primilm quertie regnum Dei & Manis, d'institutioni eius, & hac omnia adijetentur vobis. \* lix meus† inebrians † quam precla-• rus est. Et misericordia rua subsequed Super aquem refellionis, id est sapien calico

tar me, omnibus diebus vitæ mez. a ffablequeux meinon surem bie martem fed et eriam inba-biem. Eli ad precedentia fernande, & et inhebitem . b f Hen manufafte.

Et tut inhabitem in domo domini,

e Educatit me.non folum corporaliter, fed etiam spiritualiter.

f Inimate mean consertit. Ad le de fe culi vanitate.
g Deduxitme fuper femitas inflitig. Per
observantiam confiliorum, que funt

MORALITER

uore dilectionis. Inebriamini charifsimi . z. Quam praelarus eff. Effectioè , quia ducit ad claritatem patria . u. Es mifericordia tua fubfequetar me, omnibus diebus vita mea . Pet

augmentum gratie. x Es vimbabitem in domo domini, feilicer cœlefti fine fine .

cais eft ad easteandum verum findim boiss públ. in quo regimen eccidir fui fui Chrillo pultore demonstrainamode eccidirata in la terginnia regional regional pultore demonstrainamode eccidirata in la terginnia regional pultore demonstrainamode eccidirata in la terginnia regional pultore demonstrainamode eccidirata ecci



- An individual, concrete, medium-sized physical object.
  - Has shape, size, colour, weight, ...
  - Can be seen, heard, felt, smelled, tasted, ...

Do *digital* documents have weight, smell, taste...?

- Is relatively stable



- An individual, concrete, medium-sized physical object.
- Distinct from other such objects because:
  - It is an artefact (intentionally created by a human)

ChatGPT?

- It carries inscription in some writing system
- It has linguistic content, conveys meaning
- It is repeatable.



# Spoiler:

- Edge, corner and borderline cases not addressed:
  - Size: Projected on the Moon, nano-writing
  - Stability: Sky writing, smoke signals, rolling text boards
  - Perceivability: Invisible writing
  - Intentionality: 1,000,000 monkeys hammering on typewriters, pebbles on a beach, Broca's aphasia
- Let's see if we can handle the easier cases first.
  - Then, let's dive into the edge cases.



## Documents are repeatable

- The individuality of documents is often irrelevant
  - As opposed to e.g. paintings and sculptures

- Goodman: autographic vs allographic arts
- "If Mona Lisa is in the Louvre, where is Hamlet?"
  - Any individual document is functionally equivalent (for certain purposes) to other objects of a certain class
- Repeatability explained by instantiation:
  - Any document instantiates a text.
  - Any text may be instantiated by many documents.
  - A text is an abstract object.



### Documents and texts

- Documents instantiate the same text if they:
  - Have the same linguistic content
  - Express the same meaning
    - Physical or perceptual similarity of documents is irrelevant

# [Not considered here:]

- Instead of abstract objects, text may be seen as:
  - Mental states or events, mental representations
  - Brain states
  - Events
  - Actions, individual or collective
  - Social constructs



## Librarians regard documents as abstract objects

• The FRBR\* hierarchy:

Work (Text) Abstract

Expression (Translation)

Manifestation (Edition)

Item (Document) Concrete



<sup>\*</sup> Functional Requirements for Bibliographic Records

# Yay!! We are well into metaphysics!

- Or strictly speaking, ontology<sup>P</sup>
  - Though not ontology<sup>c</sup> as in computer science

Quine

- Ontology<sup>P</sup>: Which are the basic constituents of the world?
- Some candidates:
  - Slices of time and space
  - Objects, properties and relations
  - Events, agents and actions
  - Universals and particulars
  - Concreta and abstracta



## Concrete or abstract, particular or universal?

		Particular	Universal
		In at most one place at t	In many places at t
Concrete	In only p at t		
Abstract	Not in space or time		

(Adapted from Cynthia McDonald, Varieties of Things, 2005)



# Concrete or abstract, particular or universal?

		Particular	Universal
		In at most one p at t	Many in p at t
Concrete	In only p at t	Planets, trees, electrons	
Abstract	Not in space or time	Numbers, sets	Properties, relations

(Adapted from Cynthia McDonald, Varieties of Things, 2005)



## Concrete or abstract, particular or universal?

		Particular	Universal
		In at most one p at t	Many in p at t
Concrete	In only p at t	Planets, trees, electrons, documents	?
Abstract	Not in space or time	Numbers, sets, texts	Properties, relations

(Adapted from Cynthia McDonald, Varieties of Things, 2005)



## Texts as abstract objects

- As such, texts are on a par with numbers and sets:
  - They have all their intrinsic properties essentially,
  - are not in space and time,
  - are causally inert,
  - cannot be created or seize to exist,
  - cannot be changed.
  - They do have instances
- Think of numbers and (pure) sets:
  - The number 2 cannot change.
  - "New" prime numbers are discovered, not created.

Abstract objects belong to Frege's "third realm": not physical or mental, but abstract.



- 1) Texts are abstract objects
- 2) Abstract objects are not created
- 3) Texts are created



- 1) Texts are abstract objects
- 2) Abstract objects do not change
- 3) Texts change



- 1) Texts are abstract objects
- 2) Abstract objects do not cause anything
- 3) Texts cause things



- 1) Texts are abstract objects
- 2) Abstract objects do not cause anything
- 3) Texts cause things

### For "texts", read:

- musical works
- games (chess)
- computer programs
- concepts
- words



## Responses to the trilemma

- Authors do not create texts.
  - They discover, select or indicate them.
- To edit a document is not to change a text.
  - It is to change the document so it instantiates another text.
- Texts have no causal effects, only documents have.
- Texts are not abstract objects, but something else...
- Texts are "really" only mereological sums of documents.
- Texts do not exist.



# Texts as Timed abstract objects (TAO)

• Texts construed as functions from points of time to abstract objects (e.g. strings):

Time	String
$t_1$	I remember Verona
$t_2$	I remember, but dimly, Verona

- The abstract objects (strings) don't change.
- The text doesn't change (it is a function).
- But questions remain, e.g.:
  - What is the relation between the document and the function?
  - It is reasonable to view its history as "baked into" the text?



## Where has this brought us?

- One question we have not considered yet:
  - If texts are abstract objects, i.e. causally inert:
- How can they represent and transmit knowledge?
  - Text is, after all, often called "knowledge representation".
- This is not a knock-down argument, but a nice exit point:

# Exit (immutable) abstract objects!



# Digital documents

Notice the shift from "text" to "document" below.

- The OHCO thesis\*:
  - Texts are ordered hierarchies of content objects (OHCO).
  - SGML (XML) documents are OHCOs, too.
- XML markup captures the logical structure of documents.
- TEI, according to some (caricatures, perhaps):
  - We should distinguish the physical from the logical document.
  - Traditional layout is just presentational markup.
  - XML markup makes explicit the structure which is only implicit in traditional documents.

<sup>\*</sup> Coombs, Renear, DeRose (1987), 'Markup Systems and the Future of Scholarly Processing', Communications of the ACM, 30: 933-947.

## A (counter-)example: UnderDok

 An XML-based document system for the creation and maintenance of multi-lingual administrative documents in higher education.



#### **Course Description for FIL124**

#### Course name

Introduksjon til praktisk filosofi

Bokmål: Introduksjon til praktisk filosofi English: Introduction to practical philosophy

#### Course Credit Points

10 points.

#### Course Code

FIL124

#### Course Unit Level

Bachelor

#### Department

Department of Philosophy

#### Access to the Course Unit

The course is open to students admitted at the University of Bergen

#### Language of Instruction

Norwegian or English

#### Aim and Content

Philosophy is often divided into practical philosophy and theoretical philosophy. Practical philosophy includes such areas as ethics, aesthetics, political and social philosophy, philosophy of law, philosophy of religion, feminist philosophy, action theory and value theory. The distinction between practical philosophy and theoretical philosophy is, however, not always clear and is a matter of debate, and problems within practical philosophy are often relevant for questions in theoretical philosophy. It is therefore important that all who study philosophy have a solid knowledge of practical philosophy, even if they in their advanced studies choose to concentrate on theoretical philosophy.

FIL124 aims to give students an overview of important basic concepts, arguments and positions in practical philosophy. Although the main emphasis is on subjects from contemporary philosophy, it will often be appropriate to start with philosophical works and positions from previous time periods. After completion of the course, the students should be able to demonstrate insight into central theories and problems from within practical philosophy and to see their relevance and applicability for other contexts. The course provides a foundation for further studies in philosophy at the Bachelor level.

#### **Learning Outcomes**

#### Knowledge:

After taking the course, the students should have a good knowledge of important basic concepts, arguments and positions in practical philosophy.

#### Skills:

After taking the course, the students should be able to recognize and demonstrate insight into basic problems and arguments within practical philosophy in different contexts.

#### Competence:

The course provides a basis for further studies with the aim of attaining a B.A. in philosophy. In combination with other subjects and disciplines it can form part of an education which qualifies for teaching philosophy in high schools. The course can also serve as support for a deeper understanding of basic questions in connection with the study of other disciplines.

#### Overlap

#### Pre-requirements

#### Recommended previous knowledge

Approved first semester studies. A good proficiency in English is necessary since it ought to be assumed that much of the prescribed reading will be in English. Students are advised to take FIL124 in parallel with or after FIL120 (History of philosophy, from ancient times to the enlightmnment) and FIL121 (History of philosophy, from the enlightmnent to the 20th century)

#### **Teaching Methods**

Teaching is offered in the form of lectures and seminars. The seminars focus on student activities related to topics from the lectures. Seminars are compulsary.

#### Assessment methods

A 4-day Home examination of 3000- 5000 words (not including table of contents, list of references etc.). The paper should be submitted electronically via "Mi side".

Exams in the course are held every semester.

The compulsory activities have to be formally approved before one can take an exam in the course.

The compulsory requirements have to be formally approved in a semester in which there is teaching in the relevant course in order to register for the examination in a semester in which there is no teaching in that course.

#### Compulsory Requirements

Oral presentation in groups on the seminar

Students must attend at least two thirds of the seminars.

An approval of compulsory requirements is valid for three semesters, including the semester in which the approval is given.

#### Reading List

The reading list includes an introductory text book and selected excerpts from other texts.

#### **Grading Scale**

From A to F

#### **Teaching Term**

Fall.

#### Place of Teaching

Bergen

#### Course Unit Evaluation

The teaching will be evaluated from time to time.

#### **Contact Information**

Department of Philosophy Email: studierettleiar@fof.uib.no



### [Screenshot]

### **Contact Information**

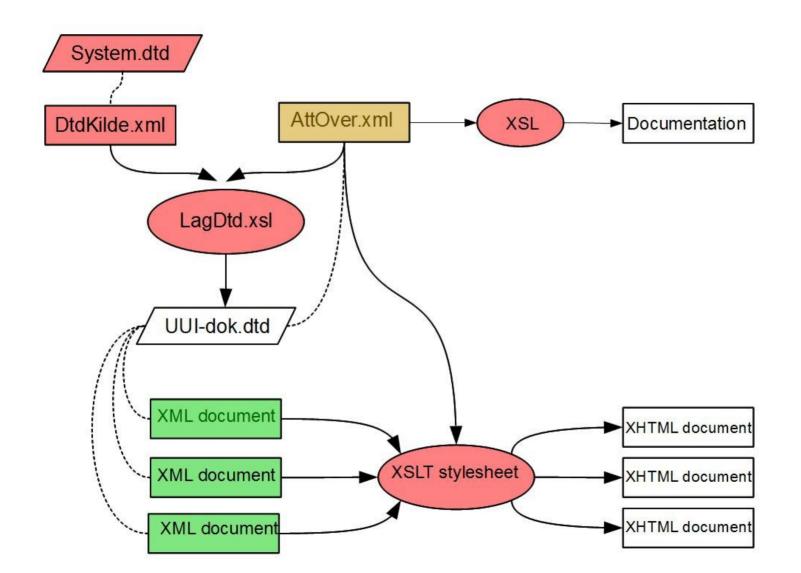
Department of Philosophy

Email: studierettleiar@fof.uib.no

FIL124.xml

<Kontaktinfo Er="FoF-kontakt"/>







### FIL124.xml

### <Kontaktinfo Er="FoF-kontakt"/>

### AttOver.xml

```
<Element id="Kontaktinfo">
 <Overskrift>
 <Nor>Kontaktinformasjon</Nor>
 <Eng>Contact Information</Eng>
 </Overskrift>
 <Attributter>
 <SpesialAttributter>
  <Verdi id="FoF-kontakt">
  <Norsk>
   <Avsnitt>Institutt for filosofi og førstesemesterstudier</Avsnitt>
   <Avsnitt>E-post: <Lenke>
    <Mål>studierettleiar@fof.uib.no</Mål>
    <Anker>studierettleiar@fof.uib.no</Anker>
    </l enke></Avsnitt>
  </Norsk>
  <English>
   <Avsnitt>Department of Philosophy</Avsnitt>
   <Avsnitt>Email: <Lenke>
    <Mål>studierettleiar@fof.uib.no</Mål>
    <Anker>studierettleiar@fof.uib.no</Anker>
    </l enke></Avsnitt>
  </English>
  <Kort/>
  </Verdi>
 </SpesialAttributter>
 </Attributter>
</Element>
```



### Std.xsl

```
<xsl:template match="Kontaktinfo">
    <xsl:call-template name="Overskrift"/>
    <xsl:call-template name="Attributt"/>
    <xsl:apply-templates/>
    </xsl:template>

<xsl:template name="Overskrift">
    <xsl:template name="Overskrift">
    <xsl:element name="h3">
     <xsl:call-template name="Overskrift-tekst"/>
    </xsl:element>
    </xsl:template>
```

### Gen.xsl

```
<xsl:template name="Overskrift-tekst">
 <xsl:apply-templates
     select="document('AttOver.xml')//*[@id=(name(current()))]/Overskrift"/>
</xsl:template>
<xsl:template name="Attributt">
  <xsl:if test="not((@Er='Ingen') or (@Er='Anna'))">
    <xsl:choose>
    <xsl:when test="((@Er='Anna') and ($Språk='Kort'))">
       <xsl:text> [...]</xsl:text>
    </xsl:when>
    <xsl:otherwise>
       <xsl:apply-templates select="document('AttOver.xml')
         //*/Element[@id=(name(current()))]
                                                    //*[@id=(current()/@Er)]"/>
    </xsl·otherwise>
    </xsl:choose>
  </xsl·if>
</xsl:template>
```



FIL124.xml

<Kontaktinfo Er="FoF-kontakt"/>

FIL124.xhtml

<h3>Contact Information</h3>

Department of Philosophy

Email: <a href="studierettleiar@fof.uib.no"> studierettleiar@fof.uib.no</a>

[Screenshot]

## **Contact Information**

Department of Philosophy

Email: studierettleiar@fof.uib.no



## Where is the document?

FIL124.xml

<Kontaktinfo Er="FoF-kontakt"/>

FIL124.xhtml

<h3>Contact Information</h3>

Department of Philosophy

Email: <a href="studierettleiar@fof.uib.no"> studierettleiar@fof.uib.no</a>

[Screenshot]

## **Contact Information**

Department of Philosophy

Email: studierettleiar@fof.uib.no



## Where is the document?

FIL124.xml

<Kontaktinfo Er="FoF-kontakt"/>

FIL124.xhtml

<h3>Contact Information</h3>

Department of Philosophy

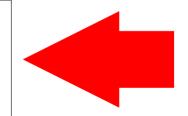
Email: <a href="studierettleiar@fof.uib.no"> studierettleiar@fof.uib.no</a>

[Screenshot]

### **Contact Information**

Department of Philosophy

Email: studierettleiar@fof.uib.no





## Where is the document?

FIL124.xml

<Kontaktinfo Er="FoF-kontakt"/>

FIL124.xhtml

<h3>Contact Information</h3>

Department of Philosophy

Email: <a href="studierettleiar@fof.uib.no"> studierettleiar@fof.uib.no</a>

[Screenshot]

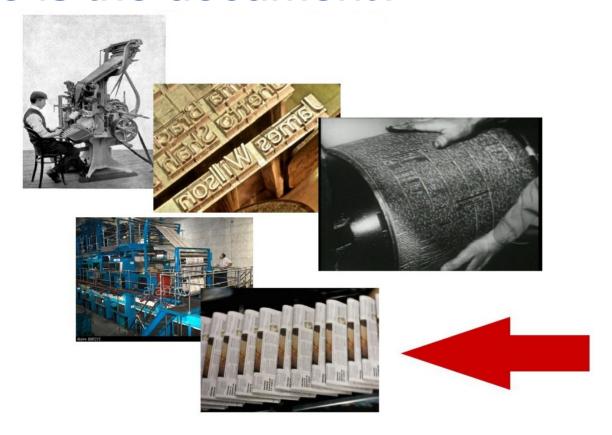
## **Contact Information**

Department of Philosophy

Email: studierettleiar@fof.uib.no

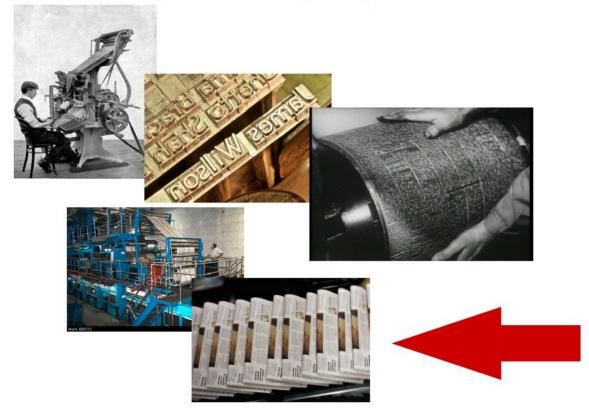
 Though in absence of the latter, we can hope to reconstruct it from the former. And info about production process often valuable.

## Where is the document?





## Where is the document?



• Though in absence of the latter, we can hope to reconstruct it from the former. And info about production process often valuable.

39

# And now to something *not* entirely different: **Transcription!**

- "Transcription" as in critical editing, not genetics\*
- Vander Meulen and Tanselle:
  - "... by transcription we mean the effort to report [...] precisely what the textual inscription of a manuscript consists of."
- So one would think transcribers know something about document identity criteria.
- The cardinal sin of transcription is "interpretation".
- Yet we hear that "all transcription is interpretational".

# **Transcription**

- Different kinds of transcription:
  - For diplomatic, critical, documentary, normalized, modernized, genetic, or synoptic editions etc.
- Levels of transcription:
  - Getting every word right, or also every letter?
  - If word -- what is a word and what is "right"?
  - If letter -- graphetic, allographic, graphemic?
  - Misspellings, slips of the pen, deletions, insertions, substitutions, abbreviations, different hands, etc.



- Based on C.S. Peirce's type-token distinction.
  - There are 14 "the"s on this page, but only one word "the" in the English language. (From memory.
     One type, many tokens.)
- We expand on Peirce:
  - Simple and compound tokens and types
  - Type *repertoires* (letters, words, sentences, ...)
- Goal: A formal model which can be proved consistent by some formal logic (FPC)

Sperberg-McQueen\* and Huitfeldt, forever forthcoming



- Documents are physical objects containing *marks*.
- Some of those marks are *tokens* instantiating *types* in some *type repertoire*.
- Documents may employ several type repertoires, together constituting a *type system*.
- Documents contain *sequences*, *sets or bags* of *simple* or *compound* tokens and types.



- The assignment of tokens to types is decided by a *reading*, which is a *relation* between documents, their tokens, and type systems.
- Roughly, if a document T is a transcription of another document E, there is some reading such that T and E constitute the same complex type in some type system.
- We may then say that T and E are T-similar.
- Thus, T-similarity is an equivalence relation.
- T-similarity is a *necessary* (not sufficient) condition for T to be a transcription of E.

- Why should we believe this?
  - The account cannot help us automatize transcription.
  - It cannot decide whether a transcription is correct or not.
  - But it does help identifying what transcribers disagree about, when they disagree.
  - And it does seem to suggest that the search for document identity criteria might fruitfully be replaced by negotiations over (exactly) what constitutes a correct reading of a document
- To transcribe is first to read, and then to write down what one has read.

   Patrick Sable from memory Patrick Sable 2012:

Patrick Sahle, from memory. Patrick Sahle 2012: "Transcription is reading written down again."



# But aren't types just abstract objects again?

 Yes, according to some, types, too, are immutable. Though types are not quite like numbers and sets:

Linda Wetzel

- They share more of their properties with their tokens than sets do with their members.
- There is no set of essential properties that needs to be shared by all tokens of a type
- So they are more like "impure sets", e.g. {Eiffel Tower}
- Our account is compatible with this view of types
  - But there are alternatives:



# But aren't types just abstract objects again?

• *Created types* can be created, change, and go out of existence.

Lee Walters 2013

- They are not in space, but in time.
- They are "universals tracing historical chains of (re)production" (cf. TAO above)
- A type exists iff it has at least one embodiment, i.e.
  - it has tokens
  - a recipe for its creation exists (XML?)
  - it is stored\*

No problems with this theory? Of course (and cmsmcq may have reservations). But we need to wrap up soon.

# Concluding, hand-waving remarks

- We may replace identity criteria based on
  - instantiation of "traditional" abstract objects
- with identity criteria based on
  - correctness of reading in terms of created types embedded within type systems
  - conventions and practices for tracing such types
- parts of this lends itself to formal modelling



# Concluding, hand-waving remarks

In general we should strive towards explication, formalization and objectivity. But, cum grano salis:

- There is an end to explication.
- There are limits to formalization.
- Interpretation is not inherently subjective.



# Thank you



## References (Not formal bibl. refs, but should be ok for web search.)

- Coombs, Renear, DeRose, "Markup Systems and the Future of Scholarly Text Processing"
- FRBR: "Functional Requirements for Bibliographic Records"
- Goodman, Nelson: "Languages of Art"
- Huitfeldt, Sperberg-McQueen: "What is transcription?"
- Huitfeldt, Marcoux, Sperberg-McQueen: "Extension of the type/token distinction to document structure"
- Huitfeldt, Vitali, Peroni: Documents as Timed Abstract Objects
- Huitfeldt: "UnderDok"
- McDonald, Cynthia: "Varieties of Things", 2005
- Sahle, "Modeling Transcription", wordpress 2012.
- Vander Meulen and Tanselle "A system of manuscript transcription."
- Quine: "On what there is"
- Walters, Lee: "Repeatable artworks as created types"
- Wetzel, Linda: "Types and tokens", 2009



# Extra! Extra! Is transcription interpretational?

- Example 1: Wittgenstein
- Example 2: The cat is on the mat
- Example 3: Uncertainty



munonggi





A and B volunteered to transcribe this snippet of writing. Neither of them had any clue as to what it might be. Let us see what happened.



munongji

1. Information provided: European language, Latin script.



1. Information provided: European language, Latin script.

munongji

A: munonyqi

B: wunouyqi



1. Information provided: European language, Latin script.

munongji

A: munonyqi

B: wunouyqi

2. Information provided: German language. NB: A doesn't know German



annongji

1. Information provided: European language, Latin script.

A: munonyqi B: wunouyqi

2. Information provided: German language.

NB: A doesn't know German

A: ofnmnbkr B: dfnmfbkr



munongji

1. Information provided: European language, Latin script.

A: munonyqi B: wunouyqi

2. Information provided: German language.

A: ofnmnbkr

NB: A doesn't know German

B: dfnmfbkr

3. Information provided: Wittgenstein, reversed alphabet.



munongji

1. Information provided: European language, Latin script.

A: munonyqi B: wunouyqi

2. Information provided: German language.

NB: A doesn't know German

A: ofnmnbkr

B: dfnmfbkr

3. Information provided: Wittgenstein, reversed alphabet.

A: ofnmnbkr B: offenbar



unions:

Information provided: European language, Latin script. 1.

A: munonyqi B: wunouyqi

Information provided: German language.

NB: A doesn't know German

A: ofnmnbkr

B: dfnmfbkr

3. Information provided: Wittgenstein, reversed alphabet. A: ofnmnbkr

B: offenbar

Agreed by both: The text actually reads... 4.

offenbar



munongzi

1. Information provided: European language, Latin script.

A: munonyqi B: wunouyqi

2. Information provided: German language.

A: ofnmnbkr

NB: A doesn't know German

B: dfnmfbkr

3. Information provided: Wittgenstein, reversed alphabet.

A: ofnmnbkr

B: offenbar

4. Agreed by both: The text actually reads...

offenbar

5. Agreed by both: The correct transcription is...

muuvnyzi



munongzi

1. Information provided: European language, Latin script.

A: munonyqi B: wunouyqi

2. Information provided: German language. NB: A doesn't know German

A: ofnmnbkr B: dfnmfbkr

3. Information provided: Wittgenstein, reversed alphabet.

A: ofnmnbkr B: offenbar

4. Agreed by both: The text actually reads...

offenbar

5. Agreed by both: The correct transcription is...

muuvnyzi

All of the steps 1-5 are *interpretational*. They do involve degrees of *uncertainty*. But in what sense are 4 and 5 *subjective*? Are there interpretational alternatives (subjective or not)?



## Example 2: The cat is on the mat

Exemplar

The cat is on the mat



Example 2: The cat is on the mat

Exemplar

The cat is on the mat

Transcription 1

The cat is on the mat

In what sense is this transcription interpretational? Is it the font? Is it the colour? Is it subjective?



Example 2: The cat is on the mat

Exemplar

The cat is on the mat

Transcription 1

The cat is on the mat

In what sense is this transcription interpretational? Is it the font? Is it the colour? Is it subjective?

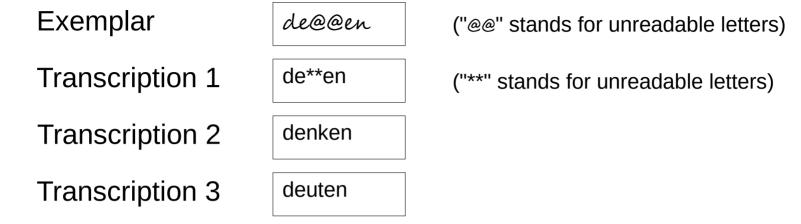
#### Transcription 2

The cub is on the mat.

Differences, yes. But are they interpretational? Are they subjective? – Not if scheme requires supplying missing full stops. And not if text is clearly about cats and not about cubs, so that "cubs" is a likely typo for "cats", and if typos are to be silently corrected.



#### Example 3: Uncertainty



Transcription 1 is less interpretational than 2 and 3. Yet transcribers 2 and 3 may agree that "@@" is unreadable. However, they supply readings, and those readings are based on different interpretations.

But isn't that just ordinary uncertainty, rather than subjectivity?

