

# Portfolio 3

## Indledning

Der er globalt set en bevægelse mod, at give digital adgang til kulturarvssamlinger, både de digitaliserede, og de digitalt fødte samlinger. Eksemplerne er mange, fra google books store digitalinges-projekt The Library Project<sup>1</sup>, over det gennemarbejdede australske Trove-api<sup>2</sup>, til fx Det Kgl Biblioteks mindre tilgængelige samlinger.

Dette ønske om åbenhed er en del af et større paradigme skifte indenfor alle grene af videnskaben, kaldet Open Science, der påvirker samtlige dele af den videnskabelige proces (Hampton, 2015). Fra indsamling og analyse, til deling af forskningsresultater.

Det er interessant i denne opgaves, er baggrunden for SMKs arbejde med deres åbne platform SMKOpen. For at se nærmere på SMKs motivation, er der taget udgangspunkt i Fecher & Friesike framework fra 2013, det beskriver fem skoler inden for Opens Science, hvor hver skole har forskellige indgangsvinkler og motivation med hvert sit fokus, også beskrevet i Fosters Open Science Training Handbook<sup>3</sup>.

Oversat til dansk er de femskoler:

1. Democratic School, der lægger vægt på, at fordeling af der er en ulige adgang til data, og at denne ulighed er et samfundsmæssig opgave at rette op på.
2. Pragmatic School, der mener at bedre viden opstår, når man giver mulighed for mere åbne og transparente workflows.
3. Infrastructur School, der har fokus på, at bedre platforme til deling og samarbejde give bedre og mere effektiv videnskabelse.
4. Public School, der mener at mere fri adgang til videnskabelige resultater vil give bedre impact. Herunder ligger også Citizen Science, hvor man inddrager offentligheden i vidensproduktionen, men også en bredere disseminering, fx via blogs.
5. Mesuament School, hvor opgøret med den måde man måler impacy i dag, og peger på alternative metrics, som en vej mod større åbenhed og bedre kvalitet.

Ser man på SMKs egen beskrivelse af deres motivation for at arbejde med åbne API'er i deres projektbeskrivelse<sup>4</sup>, skriver de sig ind under især to af Fecher & Friesikes skoler, nemlig den demokratiske skole, og skolen for infrastruktur. Først og fremmest lægger de stor vægt på, at give bred adgang til noget, som ellers kun en meget lille kerne af medarbejdere på SMK kan opleve, nemlig langt den største del af samlingen, der aldrig er på udstilling:

"SMK's samling er museets største aktiv, og består af et rigt og varieret udbud af blandt andet værker fra renæssancens klassiske europæere og en stor samling af dansk guldalderkunst. Af denne omfattende og righoldige samling, hænger kun ca. 0,6 procent af værkerne fremme. Med SMK Open bliver det muligt at tilgængeliggøre langt flere af samlingens værker."

Som i den demokratiske skole, er motivationen at give en mere lige adgang, til den fælles nationale kulturarv, ikke mindst den del af kunsten, der ligger i offentligt domæne.

Et andet fokus ligger på, at skabe en platform, der giver så god som mulig en adgang til det digitaliserede:

"Alle oplysninger stilles til rådighed via et website, hvorfra de digitaliserede værker vil kunne downloades i forskellige formater. Platformen gør det dermed både muligt for private brugere at hente og bruge enkelte værker, og for eksterne samarbejdspartnere, at tilgå API et og hente hele datasættet til integration med egne applikationer og systemer."

Her skriver de sig ind i en infrastruktur-tænkning, hvor forudsætningen for god åben videnskab netop er at skabe effektivt redskab til deling af data, og hermed effektivt redskab til videnskabelse.

Hvor Fecher & Friesikes fem skoler primært har et videnskabeligt perspektiv, er der i SMKs projektpapir også stor fokus på den helt brede offentlighed og særlige målgrupper er folkeskolen og kunstneriske udøvere.

Således indgår SMK i mange forskellige samarbejder med andre kulturinstitutioner som i Wiki Labs Kultur, bidrag til Hack for DK og med særligt fokus på unge i ULK - Unges Laboratorier for Kunst.

1. <https://support.google.com/books/partner/faq/3396243?hl=en>
2. <https://troveconsole.herokuapp.com>
3. <https://book.fosteropenscience.eu/en/02OpenScienceBasics/01OpenConceptsAndPrinciples.html>
4. [https://www.smk.dk/wp-content/uploads/2018/06/Projektbeskrivelse\\_SMK-Open.pdf](https://www.smk.dk/wp-content/uploads/2018/06/Projektbeskrivelse_SMK-Open.pdf)

1. ## API'en Til denne opgave er der udvalgt en API, der henter data fra den danske kunster Wilhelm Freddie's værker, med søgning på kunstneres navn og uden yderligere parametre: <https://api.smk.dk/api/v1/art/search/?keys=wilhelm%20freddie&offset=0&rows=202>

Søgningen giver et resultat på 202 elementer, der herefter er uploaded i code beautifier<sup>[5]</sup>.

## Typer af Metadata

Kigger man på Monsons opdeling i metadata typer (Monson s.98-90), er der først og fremmest en del deskriptive metadata som maler, størrelse, materiale og teknik brugt i fremstillingen. Disse oplysninger ligger bl.a. under "dimentions", "frame notes" og "documentation".

Af administrative data kan der nævnes id nummer, hvornår værket er digitaliseret og købt til samlingen, samt hvilken afdeling der er ansvarlig for digitaliseringen. Disse findes bla først i sættet. Der er også mere specielle data, som hvor og hvornår værket har været udstillet.

Strukturelle metadata er vanskeligere at få øje på i dette data sæt, da det i følge Monson typisk beskriver sammenhænge i komplekse datasæt, og der her henvises til et værk pr indføring. Der er enkelte noter, der henviser til andre værker som fx content\_description i id nummer 1180063298\_object "Jf. Freddie's egen kommentar til balletten, cit. i Læssøe 1996, p.113-114."

## Sammenligning med Dublin Core

Sammenligner man med Dublin Core, en standart af metadata skabt af OCLC i 1995 (Monson s.93), er der god overensstemmelse. Fra toppen af, er Title, Creator, Description er alle godt dækket ind og velbeskrevet. I det første værk er titelfeltet udfyldt som "Selvportræt, siddende på taburet, vendt mod venstre.", Creator er beskrevet i content\_person som Wilhelm Freddie, og Description er som nævnt overfor fordelt over flere felter, ligesom de er i Dublin Core hvor også Data, Type og format er beskrivende felter. I dette værk er der bla tale om en tegning på tyndt industripapir med målene 304 X 240.

Date er også godt dækket, både production\_date, altså skabelsesåret, som i dette tilfælde er 1945, og ligeledes digitaliseringsdato og anskaffelsesdato. Identifiser og Rights er også tydeligt beskrevet. Det sæt, der valgt i denne søgning er ikke i public domain, men stadig under kunstnerens copyright, da Wilhelm Freddie først døde i 1995, hvilket også er beskrevet i datasættet under production/creator\_date\_of\_death.

Dublin Core er kreeret for 25 år siden til at beskrive tekstligt materiale og derfor er der nogle felter, der ikke er relevante i denne sammenhæng. Det er felter Publisher, Relation og Source og disse fremgår heller ikke i SMKs metadata, i dette sæt.

## Analyse

Der er en del data, der kan bruges til statiske beregninger. Mest interessant er frembringelsesdato, hvor en tese kan være at frembringe en tidslinie over kunstnerens produktion, fx omkring og efter anden verdenskrig og udstillingsdatoer, gerne kombineret med udstillingssteder, der vil kunne sige noget om kunstnerens indflydelse geografisk og periodisk.

[^5]<https://codebeautify.org/jsonviewer>

In [231]:

```
### Opgave 3
##Data importeres

# import the requests module
import requests
# import pandas
import pandas as pd
# import numpy
import numpy as np
# import json normalize module
from pandas.io.json import json_normalize

api_search_url = 'https://api.smk.dk/api/v1/art/search/?'

#her bygges opsætningen
params = {
    'keys': 'wilhelm+freddie' ,
    'rows': '2000',
    'encoding': 'json',
    'items': 'items',
}

# henter svar fra url på de parametre, der er bedt om
response = requests.get(api_search_url, params=params)

print(response)
json = response.json()
df = json_normalize(json['items'])
```

<Response [200]>

In [239]:

```
# størrelse af datasæt
df.shape
```

Out[239]:

(202, 15)

In [240]:

```
# Visning af hele datasættet i en dataframe
df.head()
```

Out[240]:

	indkøbsdato	created	documentation	deltaget i udstilling	frame_notes	id	notes	object_history_note	udstilles
0	2003-12-31	2019-08-07T07:35:35Z	NaN	NaN	[Bagklædning: false, Mikroklimaramme: false]	1180072995_object	NaN	[Wilhelm Fredgies dødsbo]	False
1	1976-12-30	2019-08-07T05:39:01Z	NaN	NaN	[Bagklædning: false, Mikroklimaramme: false]	1180043061_object	NaN	NaN	False
2	2004-01-20	2019-08-07T06:44:05Z	[{'title': 'Clinch: danske kunstnere markerer ...	[{'exhibition': 'Clinch!', 'date_start': '2004...	[Bagklædning: false, Mikroklimaramme: false]	1180060327_object	NaN	[Dansk privatsamler]	True
3	1996-12-31	2019-08-07T03:37:39Z	[{'title': 'Wilhelm Freddie- den evige oprører...	[{'exhibition': 'Modern Danish Art', 'date_sta...	[Bagklædning: false, Mikroklimaramme: false]	1180004557_object	NaN	NaN	False
4	1989-12-31	2019-08-07T04:29:33Z	[{'title': 'Wilhelm Freddie: Galerie Asbæk, Kø...	[{'exhibition': 'Freddie', 'date_start': '1989...	[Bagklædning: false, Mikroklimaramme: false]	1180021751_object	[Betegnet på bagsiden for oven med tusch: Wilh...	NaN	True

In [233]:

```
# for at få et overblik over værdier, og især hvilke kolonner, der ikke har værdier,
#laves et overblik med isnul funktionen
print(df.isnull().sum())
```

acquisition_date	199
acquisition_date_precision	9
alternative_images	158
colors	117
content_description	160
content_person	199
content_subject	201
created	0
credit_line	0
current_location_name	185
dimensions	3
distinguishing_features	182
documentation	122
exhibitions	167
frame_notes	0
has_image	0
id	0
iiif_manifest	0
image_cropped	162
image_height	124
image_iiif_id	124
image_iiif_info	124

image_mime_type	124
image_native	117
image_orientation	124
image_size	124
image_thumbnail	117
image_width	124
inscriptions	25
labels	201
materials	183
modified	0
notes	167
number_of_parts	77
object_history_note	58
object_names	10
object_number	0
object_url	0
on_display	0
part_of	128
parts	154
production	0
production_date	1
production_dates_notes	190
public_domain	0
related_objects	158
responsible_department	9
rights	0
techniques	0
titles	0
work_status	156

dtype: int64

In [234]:

```
# her droppes de kolonner er har nul værdier og dem, der ikke er interessant i denne sammenhæng.
df.drop(columns = ["alternative_images",
                    "acquisition_date",
                    "current_location_name",
                    "colors",
                    "content_description",
                    "content_person",
                    "content_subject",
                    "credit_line",
                    "has_image",
                    "inscriptions",
                    "labels",
                    "materials",
                    "modified",
                    "part_of",
                    "parts",
                    "related_objects",
                    "responsible_department",
                    "techniques",
                    "dimensions",
                    "distinguishing_features",
                    "dimensions",
                    "iiif_manifest",
                    "image_cropped",
                    "image_height",
                    "image_iiif_id",
                    "image_iiif_info",
                    "image_mime_type",
                    "image_native",
                    "image_orientation",
                    "image_size",
                    "image_thumbnail",
                    "image_width",
                    "number_of_parts" ,
                    "object_names",
                    "object_number" ,
                    "object_url",
                    "production_dates_notes"], inplace=True)
df.head()
```

Out [234]:

	acquisition_date_precision	created	documentation	exhibitions	frame_notes	id	notes	object_history_
0	2003-12-31	2019-08-07T07:35:35Z	NaN	NaN	[Bagklædning: false, Mikroklimaramme: false]	1180072995_object	NaN	[Wilhelm Freddies dødsbo]
1	1976-12-30	2019-08-07T05:39:01Z	NaN	NaN	[Bagklædning: false, Mikroklimaramme: false]	1180043061_object	NaN	
2	2004-01-20	2019-08-07T06:44:05Z	['title': 'Clinch: danske kunstnere markerer ...	['exhibition': 'Clinch!', 'date_start': '2004...	[Bagklædning: false, Mikroklimaramme: false]	1180060327_object	NaN	[Dansk privatsamler]
3	1996-12-31	2019-08-07T03:37:39Z	['title': 'Wilhelm Freddie- den evige oprører...	['exhibition': 'Modern Danish Art', 'date_start': '1989...	[Bagklædning: false, Mikroklimaramme: false]	1180004557_object	NaN	
4	1989-12-31	2019-08-07T04:29:33Z	['title': 'Wilhelm Freddie: Galerie Asbæk, København ...	['exhibition': 'Freddie', 'date_start': '1989...	[Bagklædning: false, Mikroklimaramme: false]	1180021751_object	[Betegnet på bagsiden for oven med tusch: Wilh...	

In [235]:

```
# Her ses alle tilbageværende kolonne overskrifter og alle data-typer, der primært består af objekter, men også boolske, her True/False og floats
df.dtypes
```

Out[235]:

```
acquisition_date_precision    object
created                       object
documentation                  object
exhibitions                   object
frame_notes                   object
id                             object
notes                         object
object_history_note           object
on_display                    bool
production                    object
production_date               object
public_domain                 bool
rights                        object
titles                        object
work_status                   object
dtype: object
```

In [243]:

```
df.head()
```

Out[243]:

	indkøbsdato	created	documentation	deltaget i udstilling	frame_notes	id	notes	object_history_note	udstilles
0	2003-12-31	2019-08-07T07:35:35Z	NaN	NaN	[Bagklædning: false, Mikroklimaramme: false]	1180072995_object	NaN	[Wilhelm Freddies dødsbo]	False
1	1976-12-30	2019-08-07T05:39:01Z	NaN	NaN	[Bagklædning: false, Mikroklimaramme: false]	1180043061_object	NaN	NaN	False
2	2004-01-20	2019-08-07T06:44:05Z	['title': 'Clinch: danske kunstnere markerer ...	['exhibition': 'Clinch!', 'date_start': '2004...	[Bagklædning: false, Mikroklimaramme: false]	1180060327_object	NaN	[Dansk privatsamler]	True
3	1996-12-31	2019-08-07T03:37:39Z	['title': 'Wilhelm Freddie- den evige oprører...	['exhibition': 'Modern Danish Art', 'date_start': '1989...	[Bagklædning: false, Mikroklimaramme: false]	1180004557_object	NaN		
4	1989-12-31	2019-08-07T04:29:33Z	['title': 'Wilhelm Freddie: Galerie Asbæk, København ...	['exhibition': 'Freddie', 'date_start': '1989...	[Bagklædning: false, Mikroklimaramme: false]	1180021751_object	[Betegnet på bagsiden for oven med tusch: Wilh...		

3	1996-12-31	2019-08-07T03:37:39Z	Freddie- den	documentation	deltaet i udstilling	Mikroklimate: false, Mikroklimaramme: false]	1180004557_object	NaN	NaN	False
	indkøbsdato	created	documentation	deltaet i udstilling	frame_notes		id	notes	object_history_note	udstilles
4	1989-12-31	2019-08-07T04:29:33Z	[[{'title': 'Wilhelm Freddie: Galerie Asbæk, Kø...	[[{'exhibition': 'Freddie', 'date_start': '1989...	[Bagklædning: false, Mikroklimaramme: false]	1180021751_object	[Betegnet på bagsiden for oven med tusch: Wilh...	NaN	True	

In [237]:

```
#Her omdøbes udvalgte kolonner
df.rename(
    columns={
        "acquisition_date_precision": "indkøbsdato",
        "on_display" : "udstilles",
        "exhibitions" : "deltaet i udstilling"
    },
    inplace=True)
df.head()
```

Out[237]:

	indkøbsdato	created	documentation	deltaet i udstilling	frame_notes	id	notes	object_history_note	udstilles
0	2003-12-31	2019-08-07T07:35:35Z	NaN	NaN	[Bagklædning: false, Mikroklimaramme: false]	1180072995_object	NaN	[Wilhelm Freddie dødsbo]	False
1	1976-12-30	2019-08-07T05:39:01Z	NaN	NaN	[Bagklædning: false, Mikroklimaramme: false]	1180043061_object	NaN	NaN	False
2	2004-01-20	2019-08-07T06:44:05Z	[[{'title': 'Clinch: danske kunstnere markerer ...	[[{'exhibition': 'Clinch!', 'date_start': '2004...	[Bagklædning: false, Mikroklimaramme: false]	1180060327_object	NaN	[Dansk privatsamler]	True
3	1996-12-31	2019-08-07T03:37:39Z	[[{'title': 'Wilhelm Freddie- den evige oprører...	[[{'exhibition': 'Modern Danish Art', 'date_sta...	[Bagklædning: false, Mikroklimaramme: false]	1180004557_object	NaN	NaN	False
4	1989-12-31	2019-08-07T04:29:33Z	[[{'title': 'Wilhelm Freddie: Galerie Asbæk, Kø...	[[{'exhibition': 'Freddie', 'date_start': '1989...	[Bagklædning: false, Mikroklimaramme: false]	1180021751_object	[Betegnet på bagsiden for oven med tusch: Wilh...	NaN	True

In [ ]:

In [278]:

```
# ændrer datatypen object til float
c = 'indkøbsdato'
for c in df.columns:
    df[c] = pd.to_numeric(df[c], errors='coerce')
```

In [279]:

```
df.dtypes
```

Out[279]:

```
indkøbsdato      float64
created          float64
documentation    float64
deltaet i udstilling float64
```

```
deltaget_i_udstilling    float64
frame_notes              float64
id                       float64
notes                   float64
object_history_note      float64
udstilles                bool
production              float64
production_date          float64
public_domain            bool
rights                  float64
titles                  float64
work_status              float64
dtype: object
```

## Videre arbejde

Analysér der kunne være interessante arbejde videre med i denne notebook kunne være:

- Lave en grafisk visualisering over geografiske data fra kolonnen 'deltaget i udstilling.
- At sammenkøre denne visualisering med årstals data fra samme kolonne.
- At lave en graf, der viste en opsummerede antal produktioner over år, for at vise kunsterens produktionshastighed fordelt på perioder.
- Det kunne også være interessant at plotte kunsterens alder ind på denne tidslinie, for at give et billede af produktionen gennem livet.

Man kunne lave ovenstående, og mange andre analyser, på det ovenover beskrevne data, og de ville vise det billede, som ovenstående data indeholder, men det er vigtigt at forholde sig den bias, der er i alle samlinger. Man må tage forbehold i resultaterne for, hvor stor en del af kunsterens produktion der ejes af SMK, og hvor stor en del af denne samling, der er digitaliseret.

Fecher, B. & Friesike, S. (May 30, 2013) Open Science: One Term, Five Schools of Thought. RatSWDWP 218. Available at SSRN: <https://ssrn.com/abstract=2272036> or <http://dx.doi.org/10.2139/ssrn.2272036>

Monson, J. D. (2017). Getting started with digital collections: scaling to fit your organization. Chicago: ALA Editions, an imprint of the American Library Association.