

KANTTEKENINGEN BIJ EEN DATABANK

De afgelopen 2 jaar besteedde BAM veel aandacht aan het in kaart brengen van Mediakunst, Beeldende Kunst en Audiovisuele Kunst in Vlaanderen. De gegevens werden door medewerkers van BAM met zorg verzameld en ingevoerd in een databank die via <http://bamar.be> publiek te raadplegen is. Met name de sectie Mediakunst leverde een rijke hoeveelheid data op, en al bladerend door de website beginnen we ons af te vragen hoe die gegevens beter te begrijpen. De hieropvolgende pagina's vormen de neerslag van onze expeditie.

// Deelnemers

Nicolas: programmeur en data-activist (OSP)
Femke: ontwerper, kunstenaar en rapporteur (OSP)
Pierre: ontwerper, typograaf, cartograaf (OSP)
Harrison: ontwerper, typograaf (OSP)
Michael: computer wetenschapper gespecialiseerd in interactieve lees- en redactiesystemen
Thomas: ontwerper, betrokken bij de denktank over de toekomst van de BAM-databank
Didier: programmeur, ontwikkelt de BAM databank en website (Système D)

.....

KAART 1: FLOWCHARTS

Het is verleidelijk om direkt met grafieken, kaarten en netwerkvisualisaties te beginnen, maar ons eerste werk bestaat er uit inzicht te krijgen in de manier waarop de gegevens in de BAM-databank zijn opgeslagen. Dat is niet alleen uit technisch oogpunt noodzakelijk, maar het geeft ons ook een beter beeld van de 'data-praktijk', de manier waarop de gegevens in werkelijkheid gehanteerd worden.

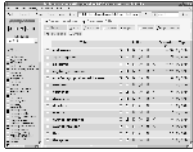


Eennot.html
19.7 KB

Didier heeft ons een *SQL-Dump* gestuurd, een letterlijke kopie van de BAM-gegevens die we ieder op onze eigen computers in een locale MYSQL-databank importeren. Nu kunnen we zelf complexe zoekopdrachten uitvoeren en ook de verschillende scripts die we

programmeren, direct toegang geven tot de gegevens.

```
// Overzichtstabel in  
PHP-MyAdmin
```



Maar hoe is de databank eigenlijk precies gestructureerd? [Nicolas](#) print alle verschillende tabellen uit om te zien hoe ze zijn georganiseerd. Welke gegevens verwijzen naar elkaar, welke tabellen zijn eigenlijk in gebruik?



[Michael](#) schrijft een klein script in zijn favoriete programmeertaal Python, dat een lijst genereert van alle tabellen in de databank, georganiseerd volgens het aantal *rows* dat in gebruik is.

(...)

```
c2.execute("select  
count(*) from  
"+tablename)  
count =  
int(c2.fetchone()[0])  
c2.close()
```

```
# print  
tablename, (50 -  
len(tablename)) * " ",  
count
```

```
tablerows[tablename] =  
count
```

(...)

Het volgende overzicht is het resultaat:

```
relations_projects_  
works 0  
mailing_templates 0  
mailing_sending 0  
relations_projects_  
organisations 0  
relations_projects_pa  
ges 0  
relations_projects_  
keywords 0  
relations_persons_  
projects 0  
relations_events_  
events 0  
relations_persons_  
dossiers 1  
relations_projects_  
files 1  
relations_dossiers_  
works 1  
banners 1  
events_reeksen 2  
relations_events_  
works 2  
menu_items 2  
relations_events_  
organisations 3  
relations_events_  
addresses 4
```

(...)

```
organisations 657  
relations_  
newsletters_pages 784
```

addresses 1026
relations_persons_
categories 1096
relations_
organisations_
categories 1156
mailing_cdata 1611
relations_
newsategories_
pages 1646
pages 1793
subsidies 1828
relations_persons_
works 2044
works 2080
relations_works_
images 2141
images 2284
ag_items 4370
ag_item_categories_
5027
mailing_users 8836
mailing_confirmation
9326

Het valt op dat sommige tabellen die in eerste instantie belangrijk lijken, in feite leeg zijn:

relations_persons_projects heeft 0 rows, en anderen waarvan we niet meteen weten waar ze voor dienen zoals ag_item_categories, bevatten heel veel data.

In antwoord op onze vragen e-mailt Didier een schematische voorstelling van de verschillende tabellen en hun verbindingen.



De *flowchart* diende oorspronkelijk als ontwerp voor de databank, en helpt nu de opzet van het systeem te communiceren. We transponeren de gegevens over het feitelijk gebruik van de databank naar de *flowchart* om te begrijpen hoe deze functioneert. De databank is tot leven gekomen!

.....

KAART 2: TAGCLOUDS

Wat zijn tagclouds

Plugin: waarom en wat is dat.

Over inkscape, uitleggen

Over queries, uitleggen



Als *testquery*, gebruikt Michael de volgende zoekopdracht:

```
SELECT first_name,  
count( * )  
FROM persons  
GROUP BY first_name  
ORDER BY 'count( * )'  
DESC
```

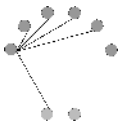
Het resultaat is interessanter dan we verwachten:

[profielen](#)
[samenwerking](#)

Hoe verhouden die termen zich eigenlijk tot de keywords in de database?

[Nicolas](#) begint met te bepalen wat de populairste 14 termen zijn. Daarvoor moeten we op verschillende plaatsen kijken; verschillende tabellen maken gebruik van de tabel keywords.

// Flowchart van alle tabellen die met de tabel keywords zijn verbonden

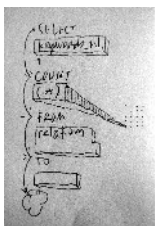


We besluiten dat de keywords gerelateerd aan organisaties, plus die aan personen er in dit geval het meest toe doen:

```
SELECT keyword_nl,  
COUNT( * )  
FROM  
relations_persons_keywor  
ds, keywords  
WHERE keywords.id =  
keyword_id  
GROUP BY person_id  
ORDER BY `COUNT( * )`  
DESC  
LIMIT 0 , 14
```

en

```
SELECT keyword_nl,  
COUNT( * )  
FROM  
relations_organisations_  
keywords, keywords  
WHERE keywords.id =  
keyword_id  
GROUP BY organisation_id  
ORDER BY `COUNT( * )`  
DESC  
LIMIT 0 , 14
```



Er zijn duidelijk verschillen tussen de keywords die favoriet zijn bij personen, en die bij organisaties:

// KEYWORDS VAN PERSONEN

- [activisme](#) 50
- body 39
- antropologie 28
- antropologie 19
- [activisme](#) 18
- [artificiële](#)
- [intelligente](#) 17
- commons 17
- apparatus 16
- cartografie / mapping 14
- [artificiële](#)
- [intelligente](#) 13
- body 13
- [artificiële](#)
- [intelligente](#) 11
- antropologie 10
- archieff 8

// KEYWORDS VAN ORGANISATIES

antropologie 42
activisme 39
bewaking 35
bewaking 34
archief 31
activisme 17
activisme 16
virtuele gemeenschap
(community) 14
cinema 12
archief 11
apparatus 10
activisme 10
body 8
cinema 6

Maar in beide gevallen staat activisme hoog op de agenda! Is dat eigenlijk wel correct? En waarom staat het 4 keer op de lijst?

Foute zoekopdracht!

We beginnen opnieuw:

QUERY

Dat ziet er overtuigender uit.

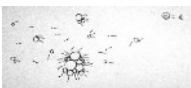
Wie hebben eigenlijk activisme als keyword gekozen?



We kunnen er niet meteen achterkomen hoe de verschillende zoekopdrachten bijelkaar op te tellen (dat moet niet te moeilijk zijn maar we hebben er geen geduld voor), dus [Nicolas](#) stapt over op PHP.

.....

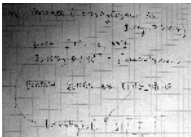
KAART 4: STEDEN



Waarom deze kaart?



Wat kun je er aan aflezen?



Precies uitleggen hoe het werkt.

Verskillende schetsen

.....

KAART 5: VLAANDEREN INTERNATIONAL

Het probleem van Mediakunst localiseren.

Verskillende gegevens tellen mee.

Het probleem van steden
zoeken; interferentie in
de kaart.

Hoe werkt de ip-
localisatie?
Waar vind je deze
gegevens?

Geolocatie per blok...