



# Artificiële Intelligentie en Relevance Ranking

ANET



# Ranking en context (I)

- Vraag:
  - Gezocht op "pearl"
  - Topresultaat? En Waarom?

**A**

# Programming Pearls



Second Edition

**B**

# PEARL JAM

twenty

**C**

## Britain, Japan and Pearl Harbor

Avoiding war in East Asia, 1936-41

Antony Best

Copyrighted Material

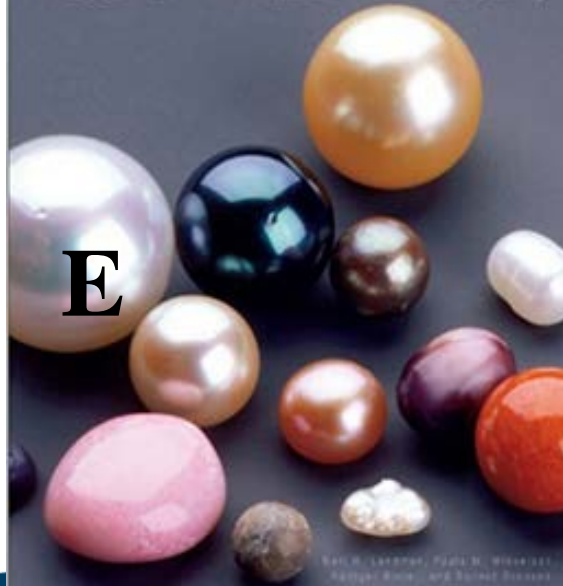
# DIVING for PEARLS

*A Thinking Journey with Hannah Arendt*

**D**

Front Cover

## Pearls: A Natural History



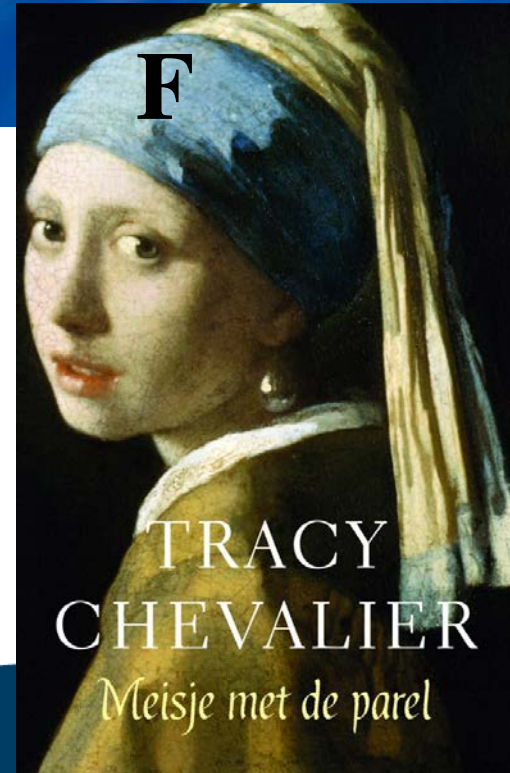
**E**

Van N. Lorenzen, Paul M. Whitsett, Katja Kohn, and Robert Roedel

American Museum of Natural History • The Field Museum

**F**

TRACY CHEVALIER  
*Meisje met de parel*





## Ranking en context (II)

- Vraag?
  - Topresultaat voor iemand anders? En Waarom?



# Ranking en context (III)

- Vraag?
  - Wat met nummer 3 tot en met 6?
    - Interesse?
    - Ranking?



# Ranking en context (IV)

- Besluit:
  - De 'juiste' ranking bestaat niet
    - voor iedereen anders
    - een momentopname
    - moeilijk in te schatten voor iemand anders



# Ranking en context (V)

- Doel
  - De resultaten van zoekactie tonen in een volgorde
  - Zoeker bekijkt slechts een klein deel
  - Ranking: dat wat de gebruiker **wil** – en wat voorhanden is – eerst tonen



# Ranking en context (VI)

- Brocade en Context
  - Aantal contexten vooraf vastleggen
    - Beschrijven in Brocade
      - # voorbeelden
  - VRAAG:
    - Hoe een gebruiker koppelen aan een context?





# Indexering en Ranking (I)

- Google kijkt naar ...
  - Aantal linken van buitenaf naar een webpagina
  - 'https' in plaats van 'http'
  - 1 afbeelding op de pagina
  - Meta tags in de webpagina
  - Inhoud
  - Commerciële belangen
  - Uw context (dynamisch)



# Indexering en Ranking (II)

- TD-IDF
  - Term frequency–inverse document frequency
    - Basis van praktisch alle ranking mechanismen in statische search indexen
    - Beloont zoekwoorden die veel voorkomen in een document maar niet in de rest van het corpus
    - Uitstekend voor full text (web pagina's), minder voor databanken



# Indexering en Ranking (III)

- Index boek

## Index

5S 50, 209  
5x waarom 77, 78, 124, 126, 195  
A3-denken 90, 168  
Andon 89, 95, 183  
Ballé 96  
Beers, Bas van 72, 94, 105  
Benchmarking 54, 121, 123, 149, 198  
Black belt 12, 64, 89, 137  
Borging 28, 50, 172, 192  
Brown paper sessie 50, 197  
Business-case 34, 36, 171  
Coachings-Kata 100  
Commitment 50, 56, 58, 96, 142  
Communicatie 58, 60, 122, 134, 135  
Continu verbeteren 24, 27, 45, 72, 97, 124, 130, 148, 185  
Dagstarts 149, 154, 182, 192  
Dankbaar, Ben 74  
Deming, William 71, 75, 78, 190  
Denken in mogelijkheden 124, 155  
Doelen 46, 51, 88, 92, 93, 112, 114, 136, 165, 194, 200, 207  
Doelstellingenboom 91, 118, 201  
Doelwaarde 119  
Draagvlak 50, 56, 58, 96, 142  
ERP 138  
Est. Emiel van 99, 100  
– Strategisch 199  
– Tactisch 196  
Ishikawa, Kaoru 71, 76  
Jones, Dan 71, 74  
Juran, Joseph 70, 74  
Just-in-time 70, 77  
Kaizen 26, 29, 47, 71, 76, 98, 138, 149, 151, 171, 192, 197  
Kanban 12, 44, 77, 89  
Kernkwaliteit 103  
Kernkwadrant 104  
Kernoorzaak 77, 102, 125, 199  
Keten 179, 28, 39, 118  
Klant 15, 21, 22, 32, 44, 57, 96, 115, 123, 185  
Klantwaarde 32, 33, 34, 106  
Kostenreductie 21, 52, 185  
Kritische Performance Indicator 112, 206  
– Beschrijven 119  
– Eigenschappen 112  
Kritische succesfactoren 50, 130  
Kwaliteit 11, 74, 75, 115, 152  
– Controles 46, 135, 152  
Lean  
– Bedrijfsbreed 104, 206, 212  
– Cultuur 12, 14, 23, 69, 74, 113, 130, 137, 150, 154, 164, 211  
– Dienstverlening 12, 22, 32, 73, 74, 116  
– Duurzaamheid 71, 72  
– Elke dag 148, 158

Federatie van  
Organisaties op het gebied van het  
Bibliotheek-  
Informatie- en  
Dokumentatiewezen (FOBID)

# REGELS VOOR DE TITELBESCHRIJVING

③

regels voor  
catalogusbouw

## REGISTER

De cijfers verwijzen naar de paragrafen.  
C-paragrafen (catalogustechniek) betreffen hoofdstuk I, K-paragrafen (keuzeregels) betreffen hoofdstuk II, V-paragrafen (vormregels) betreffen hoofdstuk III.

Termen met een asterisk (\*) zijn opgenomen in de Lijst van begrippen (Bijlage 1).

aangenomen titel, zie: uniforme titel  
aanvullende bijdragen K 7, 9, 20  
aanvulling tot voornamen V 1, 10  
academische titels, zie: titels en predikaten  
academische disputaties K 24  
achternaam V 5, 7, 37, 43, 59, 72, 73  
achtervoegsels V 5, 9, 10  
acroniemen V 106  
adellijke namen V 5, 53–56  
adellijke titels V 5, 7, 10, 24, 53–56  
advertenties K 36, 38  
afbeeldingen, zie: illustratiemateriaal  
afhankelijke corporatie, zie: corporatie, gelede  
afkortingen V 94, 106, 109  
afzonderlijke bijdragen K 40–41  
akten K 36  
alfabet, zie: schrijfwijzen  
alfabetische catalogus: functies C 2, 14, 15  
alfabetische rangschikking C 4, 6, 7, 15  
algemene toeschrijving, zie: toeschrijving  
alternatieve hoofdwoorden, zie: hoofdwoord  
alternatieve systemen van catalogusbouw C 1, 14–18  
Amerikaanse namen met voorvoegsels V 22  
Amerikaanse samengestelde namen V 32  
„analytical entry”, zie: onderdeelbeschrijving  
andere titel dan de hoofdtitel C 13; K 41, 47  
annotatie m.b.t. de inhoud K 40  
anonieme werken\* K 59–60  
atlassen K 11, 15, 39, 44  
atypische overheid, zie: corporatie  
auteur, „de” K 1–6  
auteur niet in aangewezen bron K 54  
auteur niet in publikatie K 53, 58  
auteur+titelverdubbelingen C 6, 15; K 5, 67; V 96  
auteur+titelverwijzingen K 34, 58, 60; V 3, 96, 97  
auteursaanduidingen\* K 48, 50–51; V 5, 76–77, 78, 79  
auteursvermelding, zie: verantwoordelijkheidsvermelding  
auteurswerk\* K 2, passim; V 3, 96–100  
automatisering C 14, 18



# Indexering en Ranking (IV)

- Databanken
  - Weinig inhoud; redundante inhoud (titel, abstract en onderwerpsontsluiting zijn grotendeels equivalent)
- Discovery Services → bijkomende moeilijkheid
  - heterogeen in 'natuur' en 'structuur'



# Indexering en Ranking (V)

- Scope
  - Zoeken in heterogene data (heterogeen in 'natuur' en 'structuur')
  - Geen processing tijdens het zoeken zelf
  - Correctheid en compleetheid
  - Ondersteuning van multiple contexten
  - A.I.
    - (Toepassing van **Kasparov's Law**)



# A.I. (I)

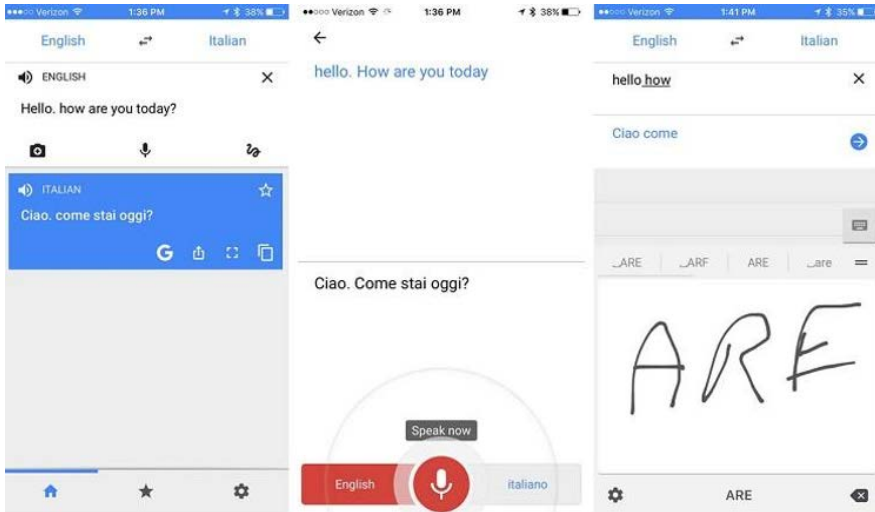


Sophia is a humanoid robot with unique features developed by Hanson Robotics Limited





# A.I. (II)

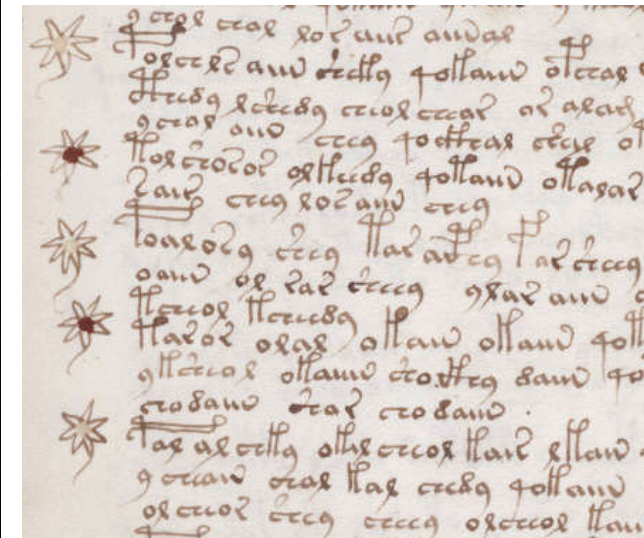






# A.I. en bibliotheken (III)

## Voynich manuscript (Beinecke MS 408)





# A.I. en bibliotheken (IV)

## Voynich manuscript

| Feiten          |  |
|-----------------|--|
| Naam            | Wilfrid <u>Voynich</u> , <u>Poolse boekhandelaar</u> (1912)  |
| Perkament       | 1404-1438 ( <sup>14</sup> C)   |
| Omvang          | ≈ 23,5 cm × 16,2 cm × 5 cm ; 240/272 pp. = 18/20 katernen  |
| Controverse     |  |
| Eigenaars       | Keizer Rudolf II (1552-1612), John Dee (astroloog Elizabeth I), ...  |
| Auteur          | Albertus Magnus, Roger Bacon, <u>Athanasius Kircher</u> , <u>Voynich</u> , ...   |
| Taal en schrift | <ul style="list-style-type: none"><li>• Code</li><li>• Kunsttaal</li><li>• Nonsens (vervalsing of 'glossolalie')</li><li>• Natuurlijke taal (in kunstalfabet)</li></ul>            |
| Oplossingen     | <ul style="list-style-type: none"><li>• 1921 (Griekse steno), 1943 (Latijnse code), 1978 (Oekraïens zonder klinkers), 1987 (Creools Vlaams), 2014 (Aziatische taal), ...</li></ul> |



## A.I. en bibliotheken (V)

- Hauer & Kondrak (University of Alberta, 2017)
- AI training
  - ontcijferen van brontaal voor versleutelde tekst (97% / 380 talen)
  - ontcijferen van anagramversleuteling van tekst (93% / 50 teksten / 5 talen)
- Voynich manuscript = (mogelijk) Hebreeuws in alfabetisch geordende anagrammen (bv. 80% lijn 1 = Hebreeuws)



## A.I. (VI)

- Schrikt af, inspireert en ... beroert ons
- **Kasparov's Law**
  - Best mogelijke resultaten door combinatie van:
    - Sterke AI
    - Goede menselijke operator met enig inzicht in de materie
    - Optimale procedures
- = Basis van de ranking software



# Ranking in Brocade (I)

- Procedures en vakkennis
  - Voorbereiding
    - Vastleggen **CONTEXTEN** (beleid)
    - Ranking scores geven aan alle records in Brocade (experten)
      - **ASPECTEN** vastleggen + wegen
        - » catalografisch, leen, aankoop, impala, ...
        - Voeden een **genetisch algoritme**



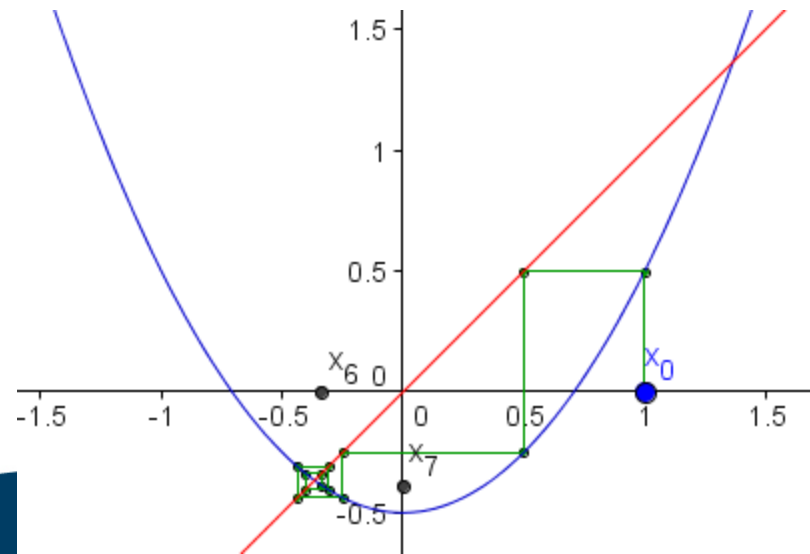
## Ranking in Brocade (II)

- **DOMEINEN** (bevat de **ASPECTEN**) :
  - <https://dev.anet.be/menu/rankdomain>
  - Lijst van boeken → visualisatie scores/ranking



# Ranking in Brocade (II)

- A.I.
  - Gewichten worden voortdurend bijgestuurd
    - Steeds andere ranking → getoetst aan een statistisch corpus → beste blijft behouden
      - Ideale ranking voor een bepaald doelpubliek





## Opdracht (xx)

- Gegeven een lijst van x aantal boeken over xxx en de doelgroepen
  - Kunsthistorisch onderzoek
  - Lerarenopleiding
- Welke aspecten kunnen invloed hebben op ranking score
  - Catalografie / Leen / Acquisitie





## Opdracht (xx)

- Gegeven een lijst van 55 boeken
  - Kunsthistorisch onderzoek
  - Lerarenopleiding
- Welke aspecten moeten meespelen bij bepaling van ranking score
  - Catalografisch / Leen / Acquisitie



# Doorbomen

- Ranking, A.I. en de GDPR?
- ...



# Literatuur (I)

- Nogal wiskundig:
  - Artificial Intelligence: A Modern Approach / Stuart Russell & Peter Norvig – ISBN 978-0136042594
  - Deep Learning / Ian Goodfellow, Yoshua Bengio, Aaron Courville



## Literatuur (II)

- Voor elke bibliotheek:
  - The rise of the robots / Martin Ford – ISBN 978-0465059997
  - Thinking Machines: The Quest for Artificial Intelligence / Luke Dormehl – ISBN 978-0143130581
  - Deep Thinking: Where Machine Intelligence Ends and Human Creativity Begins / Gary Kasparov – ISBN 978-1610397865