

# Collaborative Authoring on Open Content

with wikis

***Ronald Beelaard (rbeelaard AT gmail DOT com)***  
Wikimedia Nederland

# Wikipedia FAQ

---

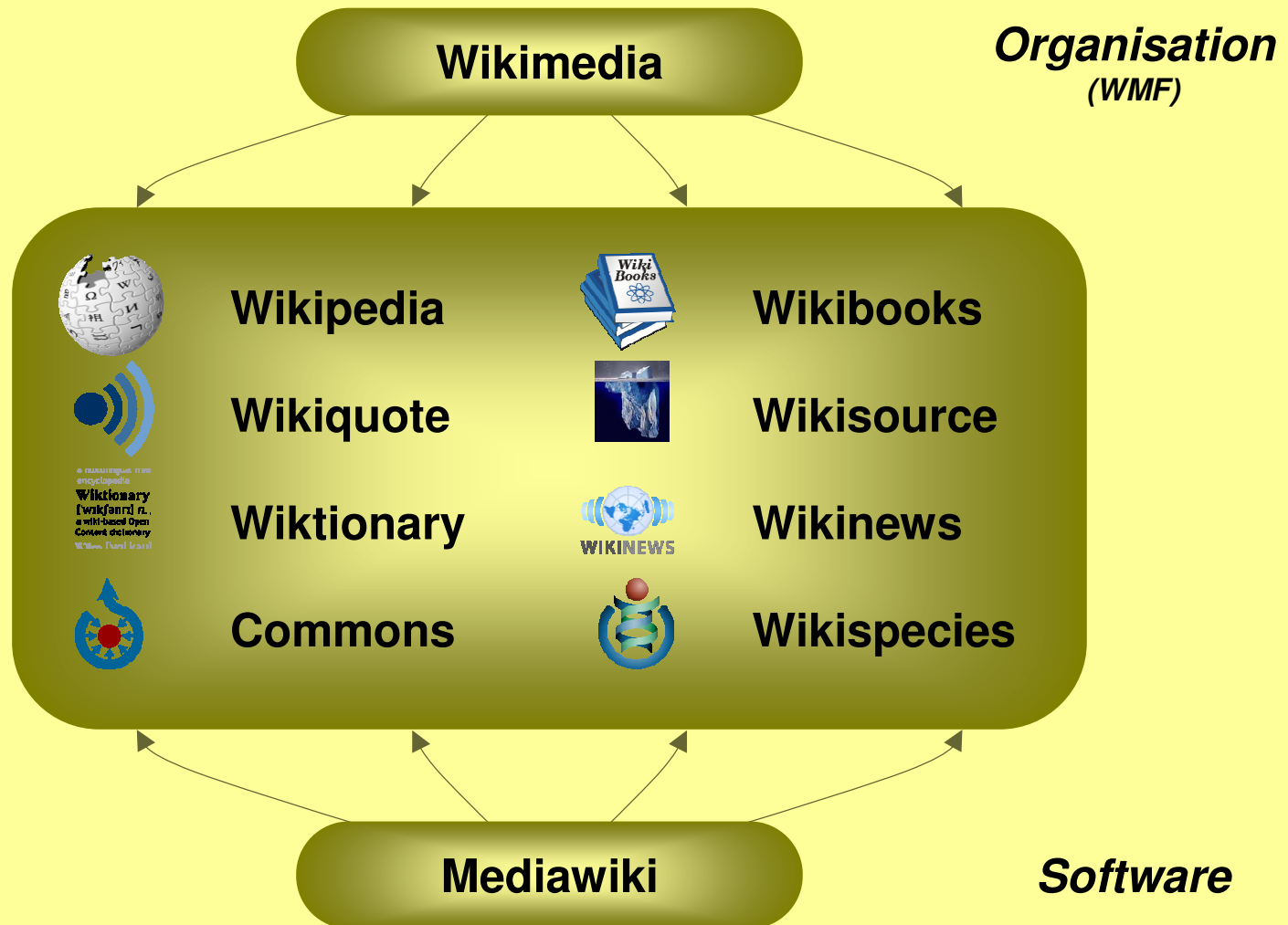
- **Wikipedia is a free encyclopedia**
  - Yes, you can edit (nearly) everything
  - Can be with or without registration
- **There is no editorial board**
  - Sysops / Moderators are trusted users with a few extra knobs (= capabilities) serving the community
- **Everyone is equal**
  - The community decides



- **Wikimedia Foundation Inc. (Florida, USA)**
  - Servers, Bandwidth
  - Fund raising
- **Local chapters**
  - Independent
  - In NL
    - Vereniging Wikimedia Nederland (internal focus)
    - Stichting Wikimedia Nederland (external focus)

***Organisations have no control of any of the projects***

# Terminology



# Objective of Wikimedia projects

---

*Provide free and free accessible information  
in as many languages as possible*

- **Examples**
  - Wikipedia is an free encyclopaedia
  - Wikibooks is a free set of informative books
- **Free as in**
  - Free beer
  - Free speech
- **However**
  - NPOV (Neutral Point Of View)

# The Essence of Wikipedia

---

## *Collaborative Authoring*

- Community
- Growth
- Quality / Reliability
- Vandalism

## *Open Content*

- Open source → content
- Wikimedia objectives
- Licensing

## *Enabling factors*

- Software platform
- Servers - Bandwidth
- System management
- Organisation

# Reliable Information

- **Long ago**
  - It was in that newspaper, that book, etc.
  - It was on the (radio) news
  - Info in an encyclopaedia had "authority"
  
- **Now – Malmberg Research** (18-04-06)
  - Secondary education (12-17 years)
  - Info source: 92% internet, 4% books
  - Don't care about info reliability: 82%



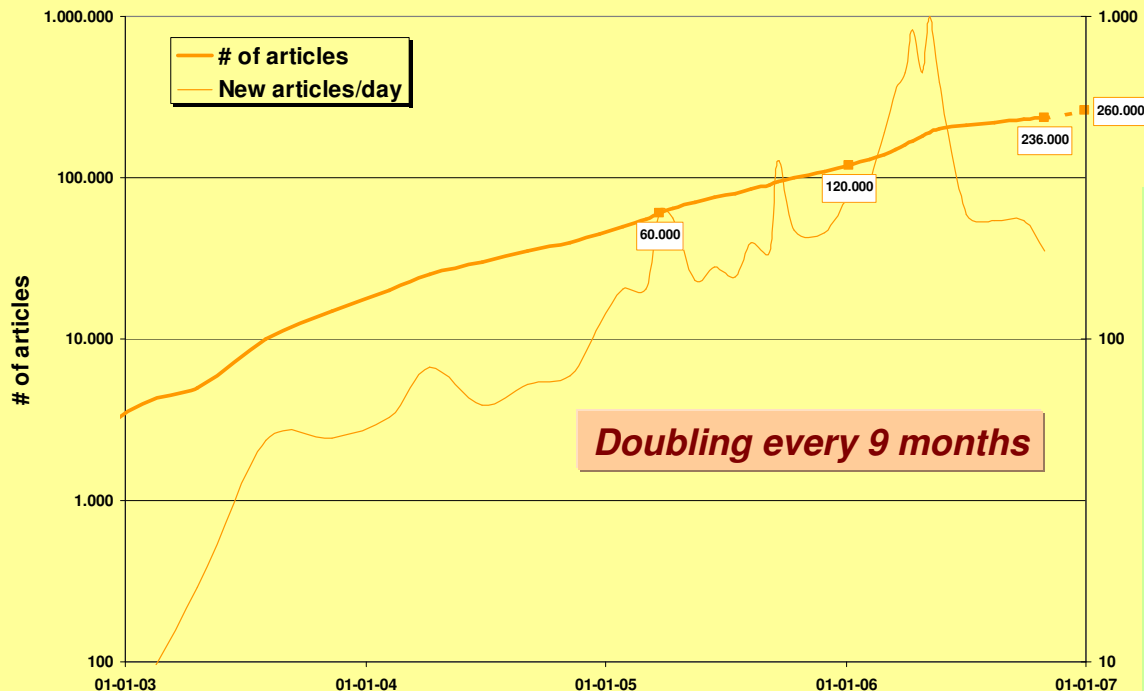
***New concepts on true information needed***



- **42 scientific articles**
  - From Wikipedia (en: version) and Britannica
  - Reviewed by (real ?) experts in the field
- **Result**
  - Wikipedia: avg. 4 errors
  - Britannica: avg. 3 errors
- **Reactions**
  - Britannica                      22-3-06
  - Nature                              23-3-06

***Britannica also contains errors***

# Wikipedia Growth (NL)



RAB - Wikimedia - sheet 10

## Groei

Aangezien de groeicurve op een logaritmische schaal min of meer een rechte is, neemt het aantal artikelen in werkelijkheid exponentieel toe.

## Nieuwe artikelen per dag

Men kan duidelijk de verhoogde activiteit zien van z.g. bots. Hiermee worden vaak vlak voor het bereiken van een belangrijke mijlpaal op grote schaal nieuwe artikelen aangemaakt, waarbij de gegevens afkomstig zijn van een database. Dit is een verschijnsel dat ook in anderstalige Wikipedia's te zien valt.

## Verwachte ontwikkeling

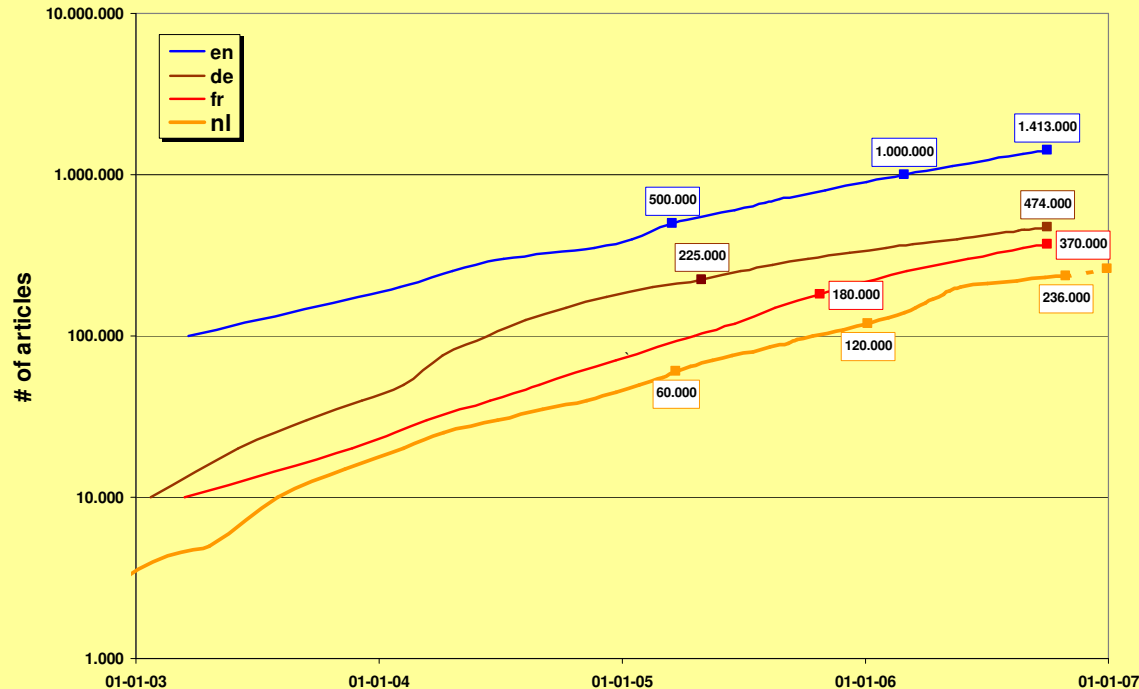
De huidige trend extrapolierend, is te verwachten dat de Nederlandse versie van Wikipedia tegen het eind van 2006 ruim een kwart miljoen artikelen zal bevatten. Ter vergelijking: Encarta bevat ongeveer 60.000 artikelen.

## Wanneer gaat deze "wet" gelden?

De geschiedenis (hier niet volledig getoond) geeft aan dat vanaf een of enkele tienduizenden artikelen deze wetmatigheid van autonome groei begint te gelden.



# Wikipedia Growth (other languages)



RAB - Wikimedia - sheet 11

## Groei van anderstalige Wikipedia's

De Franse en Engelse versie vertonen een soortgelijke ontwikkeling als de Nederlandse versie. De Franse versie groeit iets harder, de Engelse iets langzamer.

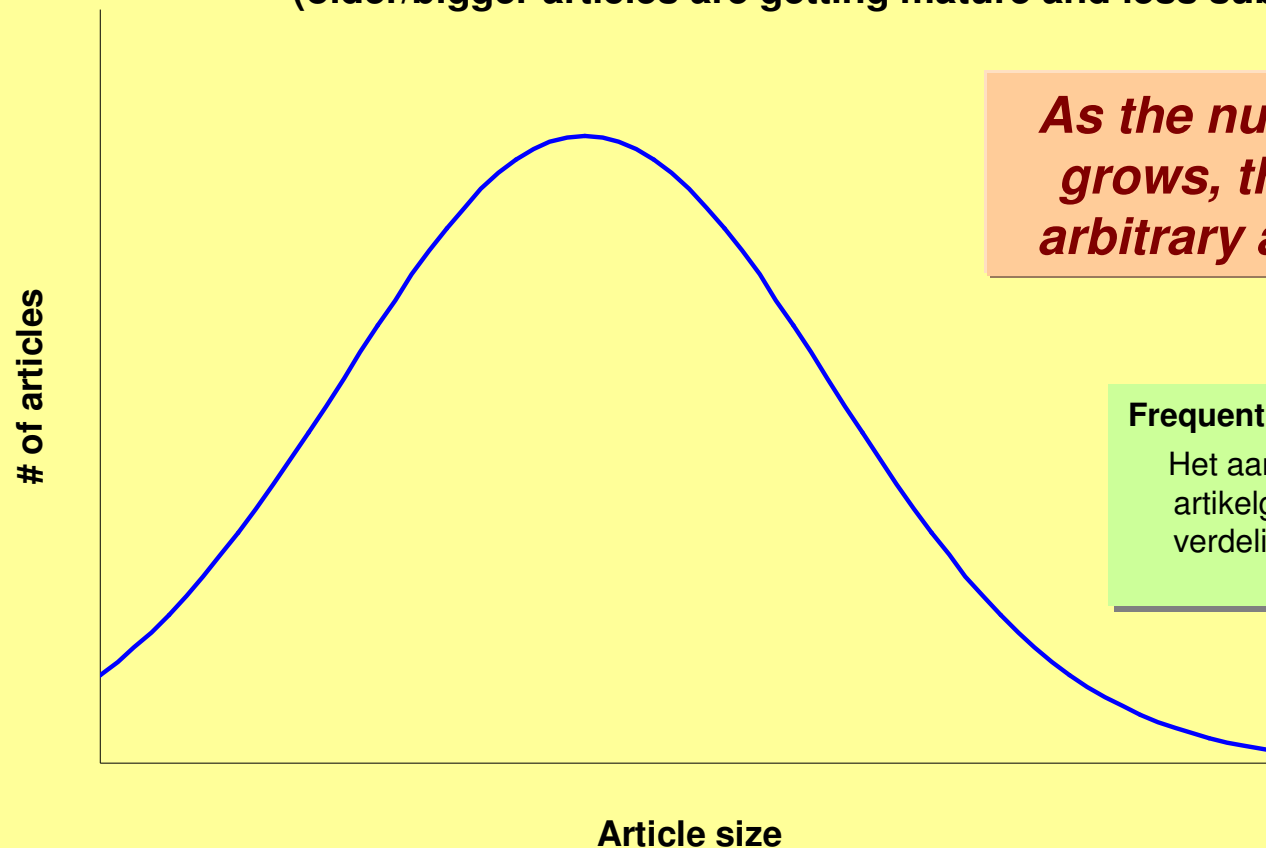
De Duitse versie vertoont een iets ander gedrag. In de eerste helft van zijn bestaan is de groei een stuk sterker geweest dan die in de tweede helft. De oorzaak hiervan is niet bekend.

## Grenzen van de groei

Uit de trends van de verschillende taalversies is nog geen enkel signaal af te leiden waaruit geconcludeerd kan worden, dat er ergens een grens aan de groei zou zijn.

# Quality and Quantity

- **Article size is a measure for amount of information**
- **Amount of information is a measure for quality**  
(older/bigger articles are getting mature and less subject to change)



*As the number of articles grows, the quality of an arbitrary article improves*

## Frequentieverdeling

Het aantal artikelen als functie van de artikelgrootte volgt een soort normale verdeling



# Confusion

- **Other language versions**

- en: 16 Aug
- de: 15 Aug
- fr: 21 Aug

- **Google**

- Various results

- **Israeli authorities**

- Government: 2 Aug
  - <http://www.pmo.gov.il>
- Knesset: 2 Aug
  - <http://www.knesset.gov.il>



ERPRINS

Uw profiel

Encyclopedie | Wereldatlas | Multimedia

Prins  Zoeken


## Peres, Shimon

Encyclopedieartikel

Zoeken in dit artikel | Printvriendelijke pagina bekijken | Dit per e-mail verzenden

**Peres, Shimon**, oorspronkelijk Peresky (ook: Sjimor tegenwoordig Wit-Rusland, 21 aug. 1923), Israëlisch politicus en premier van Israël in 1977, 1985-1986 en 1995-1996.

Peres ontmoette in 1924 met zijn ouders naar Palestina.



ENCYCLOPEDIA  
BRITANNICA online

Sharing knowledge since 1768

Home | Blog | Editorial Board

Search  Britannica Online GO

This Article's Table of Contents

- Introduction
- Additional Reading

Print this Table of Contents

Shopping

## Peres, Shimon

Concise Encyclopedia Article

PRINT PAGE • PRINT ARTICLE • E-MAIL ARTICLE • CITE ARTICLE • LINK

You are viewing a brief summary article. Members can log in to view the full article.

### Shimon Peres

originally **Shimon Peresky**  
born Aug. 16, 1923, Woyn, Pol.

## In Summary

---

- 1. Collaborative Authoring on Open Content works and results in (relatively) reliable information**
- 2. Even when anonymous authors are contributing**
- 3. Quantity (diversity of subjects) and Quality is continuously improving**
- 4. Above a certain threshold further growth is an autonomous process**

- **Users of all kind**
- **Colleagues**
  - People you know IRL
  - People you think to know
  - User names (nicknames)
    - Active
    - Less active
  - Anonymous users (IP addresses)
- **People and nicknames are different phenomena**
  - Argumentation / reasoning
  - Main influence: Anonymity



# Wikipedia Community (NL)

---

- **Age**
    - Users 12 – 65+
    - Sysops 14 – 62
  - **Geography**
    - The Netherlands and the Flanders
    - But also: US, CH, ZA, TH
  - **Specialties (examples)**
    - Religion
    - Chess
    - Animals
    - Motorcycles
  - **Users**
    - Registered: > 70.000
    - Reasonably active: 500 - 1000
-

# Mediawiki projects are "open"

---

- **Self promotion / biased biographies**
- **Data falsification**
- **Edit wars**
- **Copyright issues**
- **Vandalism**
  - Various kind

# Patrolling Edits

---

- **Anyone can do that**
  - More / less experienced users
- **Check on**
  - Ordinary vandalism
  - Encyclopaedic relevance
  - NPOV
  - Copyright infringements
  - Spamming
  - Fancruft
- **Tools**
  - List of Recent Changes (rc)
  - List of new articles
  - Live tools
  - Robots
- **Procedures**
  - List of articles patrolled
  - Deletion list
- **Special provisions**
  - Article protection
  - User blocking

- **Text**

- GFDL license (derived from GNU open source license)
  - Pressing *Save* implies copy left
  - May be re-used, provided same license and source quoted
  - May be re-used for all purposes (incl. commercial)
- Requires
  - Validity check for major additions

- **Images**

- Various licenses possible (also CC variants a.o.)
  - Is part of article (GFDL)
  - Prone to recurring disputes

# There is more ...

---

## Largest wikis

### Wikia

### Wikikids

## Enterprise wikis

## Government (a.o. CIA)

### Links

#### Largest wikis

[http://meta.wikimedia.org/wiki/List\\_of\\_largest\\_wikis](http://meta.wikimedia.org/wiki/List_of_largest_wikis)

#### Wikia

[http://www.wikia.com/wiki/List\\_of\\_Wikia](http://www.wikia.com/wiki/List_of_Wikia)

#### Wikikids

<http://kennisnet.wikia.com/wikikids/wiki/Hoofdpagina>

# Future ?

# Value / Cost (an economic perspective)

---

## *Value*

- **Community**
  - Key to building content
- **Content (virtual value)**
  - # of edits × avg. edit time × hourly rate
  
  - # of words × avg. time × hourly rate

## *Costs*

- **Infrastructure**
  - Servers & bandwidth
  - Co-location
- **Software**
  - Open source
- **Organisation**
  - Voluntary based
  - Now/recent 4-5 FTE's

**Paid by small donations**

***A revival of the "new economy" hype of the late 90's ?***

## *Strength*

- Open character
- Diversity of contributors
- Dynamics

## *Weakness*

- Open character
- Sepsis academic world

## *Opportunity*

- Success Wikipedia
- Media attention
- High internet ranking

## *Threat*

- Traditional publishers
- Infiltration





WIKIMEDIA  
NEDERLAND

# ‘Volks’ Wikipedia heeft grote toekomst, ook voor serieuze informatievoorziening

Dolph Kohnstamm

Emeritus hoogleraar psychologie  
en redacteur van het Nieuw Cultureel  
Woordenboek

Een van de oprichters van internetencyclopedie Wikipedia wil een variant daarop maken met een grotere rol voor deskundigen. Maar er zijn betere manieren om de zwakke punten van Wikipedia te corrigeren. Maak bijvoorbeeld anonieme bijdragen onmogelijk.

De twee Amerikaanse oprichters van internetencyclopedie Wikipedia, Wales en Sanger, hebben ruzie. Sanger heeft aangekondigd een verbeterde versie van Wikipedia te gaan ontwikkelen: Citizendium, een compendium van kennis voor *citizens*. Waarom doet hij dat? Heeft hij kans van slagen? Ik denk van niet. Trouwens, voor de Nederlandstalige versie van Wikipedia zal dat voorlopig niets uitmaken. Maar wel moeten hier meer mensen met kennis van zaken zich voor Wikipedia gaan inzetten. Want dat is het waard.

## Waarom Citizendium?

De bezwaren die tegen Wikipedia aangevoerd kunnen worden, zijn bekend.

VHIC presentatie d.d. 3-11-06 – page 25

## Heeft Citizendium kans van slagen?

Sanger kan om te beginnen alle anderhalf miljoen (!) artikelen die nu al op de Engelstalige Wikipedia staan als basis-materiaal kapen van Wikipedia. Dat volgt uit de *Free Documentation License*, waaronder alles valt wat in Wikipedia staat, teksten zowel als illustraties. Deze licentie houdt in dat alles vrij van auteursrecht is en dus vrijelijk mag wor-

den overgenomen en gewijzigd (maar wel met bronvermelding). De hergebruiker is alleen verplicht datgene wat hij of zij overneemt ook weer vrijelijk beschikbaar te stellen voor nieuwe hergebruikers.

Maar na het kopiëren van het materiaal begint voor Citizendium het proces van kwaliteitsverbetering. Let wel, we spreken dan alleen nog van de Engelstalige versie van Wikipedia; aan Citizendiums in andere talen denkt Sanger nog niet.

Zal hij voldoende mensen vinden die zonder betaling – want geld is er niet, tenzij hij zou besluiten advertenties te gaan toelaten – dit verbeterwerk voor Citizendium willen gaan doen? Ik denk van niet. Zelf werk ik mee aan de Nederlandstalige versie van Wikipedia, en ik heb ervaren dat het verbeteren van bestaande artikelen een lastig en ergerniswekkend karwei

te allen tijde herzien mogen worden, alleen wel door tevoren goedgekeurde medewerkers.

Bovendien zal Citizendium het enthousiasme missen van alle jonge doehet-zelf mensen, die aan het ideaal van een gratis en altijd up-to-date encyclopedie willen meewerken, in een sfeer van samenwerking en wederzijdse hulp. Citizendium zal het daarentegen moeten hebben van druk bezette oudere deskundigen, die het als een zure verplichting zullen ervaren voor nop de

de anoniemi, die ook geen gebruikerspagina openen waarop ze zich bekendmaken, moeten geweerd worden.

Deze week nog zag ik weer dat een zogeheten scrievandaal in een kwartier tijd uit tal van artikelen hele stukken had geschrapt. Gelukkig was een oppassende vandalisbestrijder, ene Willem, zo alert om deze vernielingen vrijwel dezelfde dag nog ongedaan te maken. Zo werkt die gemeenschap nu eenmaal. Allerlei mensen geven hun vrije tijd aan het bewaken en opschonen van

accepteren dat anderen die zich met het artikel bemoeien, waaraan men juist zo hard gewerkt heeft, dikwijls veel jonger zijn, studeren of zelfs nog op school zitten.

Ik raad collega's die aan Wikipedia willen gaan meewerken, aan om op hun 'gebruikerspagina' duidelijk te zijn over hun deskundigheid. Mijn ervaring is dat men op Wikipedia respect heeft voor mensen die tonen kennis van zaken te hebben, bijvoorbeeld door aan artikelen goede links naar externe bronnen mee te geven.

Waar Noord-Nederlands sprekenden ook even aan moeten wennen, is dat Zuid-Nederlands sprekenden, Vlamingen dus, andere woorden en zinsconstructies kunnen gebruiken dan wij hierboven gewend zijn. Zelf verander ik bij gelegenheid het Vlaams maar in ABN, maar je moet oppassen niet op gevoelige tenen te gaan staan.

Mijn ervaring is dat de mensen die veel aan Wikipedia meewerken – dat zijn er veel meer dan alleen de 'moderatoren' – redelijk en van goede wil zijn, bereid tot overleg op de 'overlegpagina' van het betreffende artikel of op de overlegpagina van de betreffende auteur, mits men zelf ook inschikkelijk en respectvol opereert.

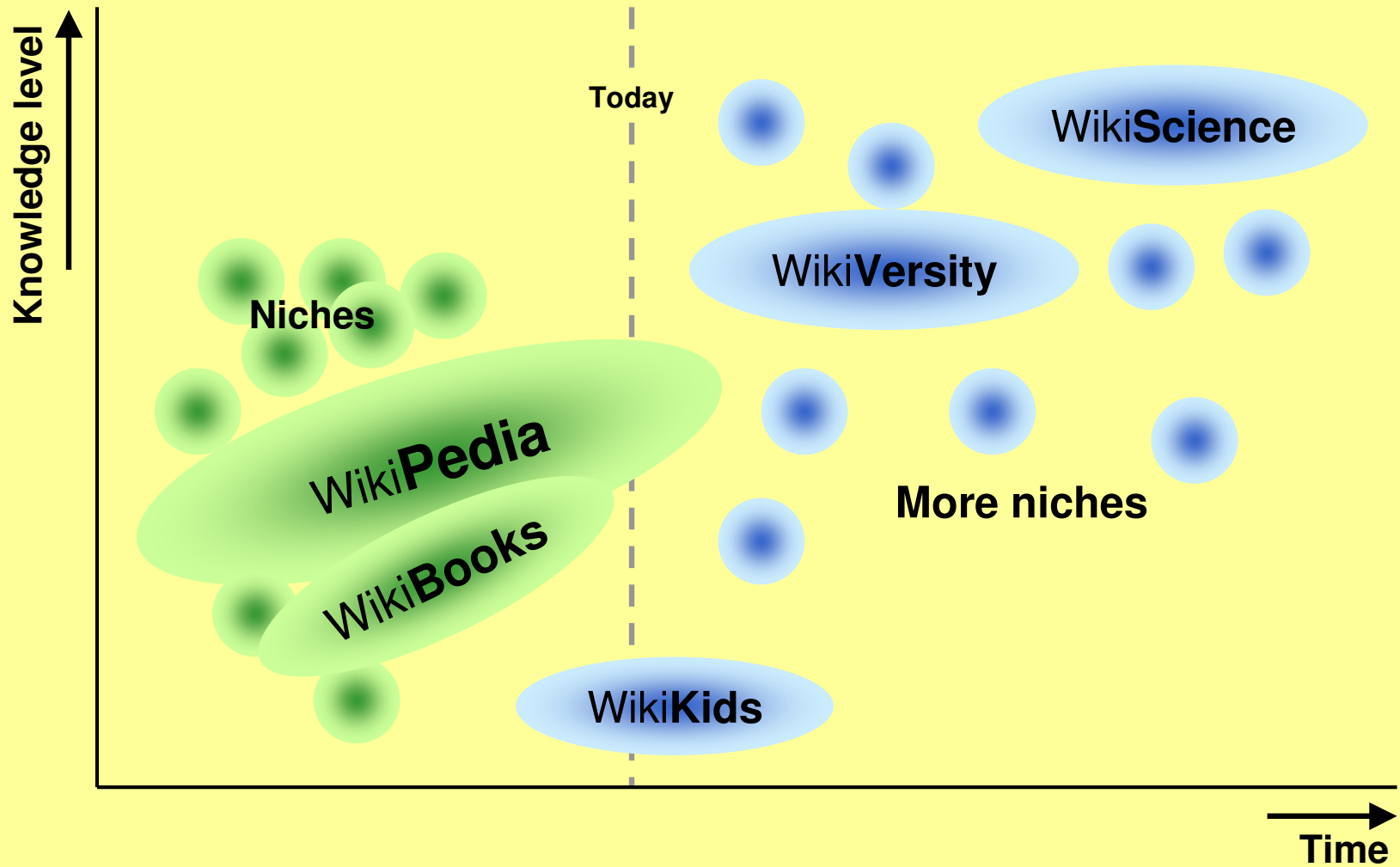
Ook voor het aantrekkelijker maken van artikelen door illustraties toe te voegen zijn meer vrijwilligers nodig. De techniek van het meewerken is in korte tijd onder de knie te krijgen en al doende leert men bij.

Men moet zich niet laten afschrikken door het feit dat op Wikipedia veel artikelen staan die normaal niet in een encyclopedie gevonden worden, zoals over vele honderden soorten motorfiets-





# The Knowledge Landscape



**A project such as Wikipedia is a subculture  
of something bigger on the internet  
that is just in it's infancy**

**That is what internet originally was  
meant for**