VK Multimedia Information Systems

Mathias Lux, mlux@itec.uni-klu.ac.at

Dienstags, 16.oo Uhr c.t., E.1.42

Agenda

- Topics & Goals
- Modalities & Examination
- Schedule

- What is **Information**?
- What are **Information Systems**?
- The Information Overload
- Current state in MuMe Consumption

C.V.

- Technische Mathematik an der TU Graz
- Doktoratsstudium Telematik

- 98-01 Entwicklung von Web-Applikationen
- 01-06 Know-Center in Graz (KPlus)
- 05-06 Ass. am KMI / TU Graz
- 06- ... Ass. am ITEC / Uni Klagenfurt

Course Topics

Multimedia Databases

Multimedia Management

Social Media Sharing

Video Analysis

Metadata

Data Mining

Digital Audio

Information Retrieval

Social Networks

Image Processing

Retrieval Evaluation

Goals I

Basic (and a little more) understanding of

- Multimedia Retrieval
- Multimedia Analysis
  - Images in the spatial domain
  - Audio & video processing
- Multimedia Databases & Metadata

Goals II

- Overview on State of the Art
  - Who is who in research
  - What to read if I want to know more?
  - Available tools in development
  - De facto & de jure standards

Goals III

- Providing a solid base for
  - Further research,
  - Consulting and
  - Practical development
- in the area of
  - Multimedia information
  - Multimedia information systems
- And: Hands on experience!

Modalities

Multimedia Information Systems ist eine
„prüfungsimmanente Lehrveranstaltung“

The grade comes from

- Some few mandatory exercises / readings
- Ongoing collaboration
- A mid term and a final project

Modalities

Mid term projects are

- The same for everyone
- A simple VIR system + evaluation

Final projects are

- Practical implementations of MMIS
- Research work & studies

Projects topics will be

- ... assigned after Easter holidays
- ... assigned to groups or individual students

Team Work

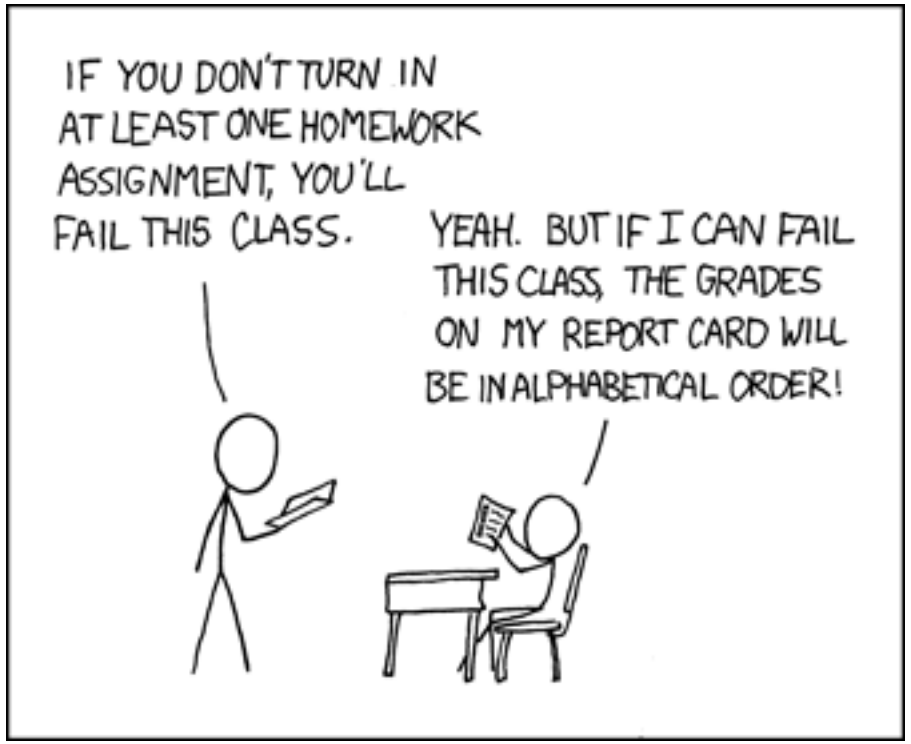
- Preferably teams of **2** students
- TEAM = „Toll-Ein-Anderer-Machts“ ?

Grades

Grade is derived from:

- 1/4 Exercises
  - Pen & paper
  - Readings
- 1/2 Project
  - ¼ each
  - Implementation
- 1/4 Presentation

http://www.xkcd.com



Schedule

- Introduction, Motivation, Information Theory & Systems
- Information Retrieval
- Web based IR, PageRank, HITS
- Network Analysis & Social Networks (Guest Lecture)
- Multimedia Metadata
- Image Analysis and Content Based Image Retrieval
- Audio & Sound Analysis
- Video Information Systems
- Multimedia Databases

Questions?

Any questions regarding organizational issues left?

Agenda

- What is **Information**?
- What are **Information Systems**?
- The Information Overload
- Current state in MuMe Consumption

What is Information?

Definition von Aamondt und Nygard (1995) :

- Data
- Information
- Knowledge

Aamodt, A. & Nygard, M. “Different roles and mutual dependencies of data, information, and knowledge - an AI perspective on their integration” *Data Knowl. Eng., Elsevier Science Publishers B. V.*, **1995**, *16*, 191-222

Data
Aamondt und Nygard (1995)

Data are syntactic entities

- Patterns without meaning
- Input to an interpretation process

Example:

- Bits & Bytes of a JPEG encoded image

Information
Aamondt und Nygard (1995)

Information is interpreted data

- Information is data with meaning
- Output from interpretation
- Input to knowledge based process

Example:

- Decoded (and displayed) JPEG image

Knowledge
Aamondt und Nygard (1995)

Knowledge is learned information

- Incorporated in an agents (software / human) reasoning resources
- Ready for active use
- Output of learning process

Example:

- There is a dog shown on the JPEG image

What is Information?

Definition of Zeleny (1987):

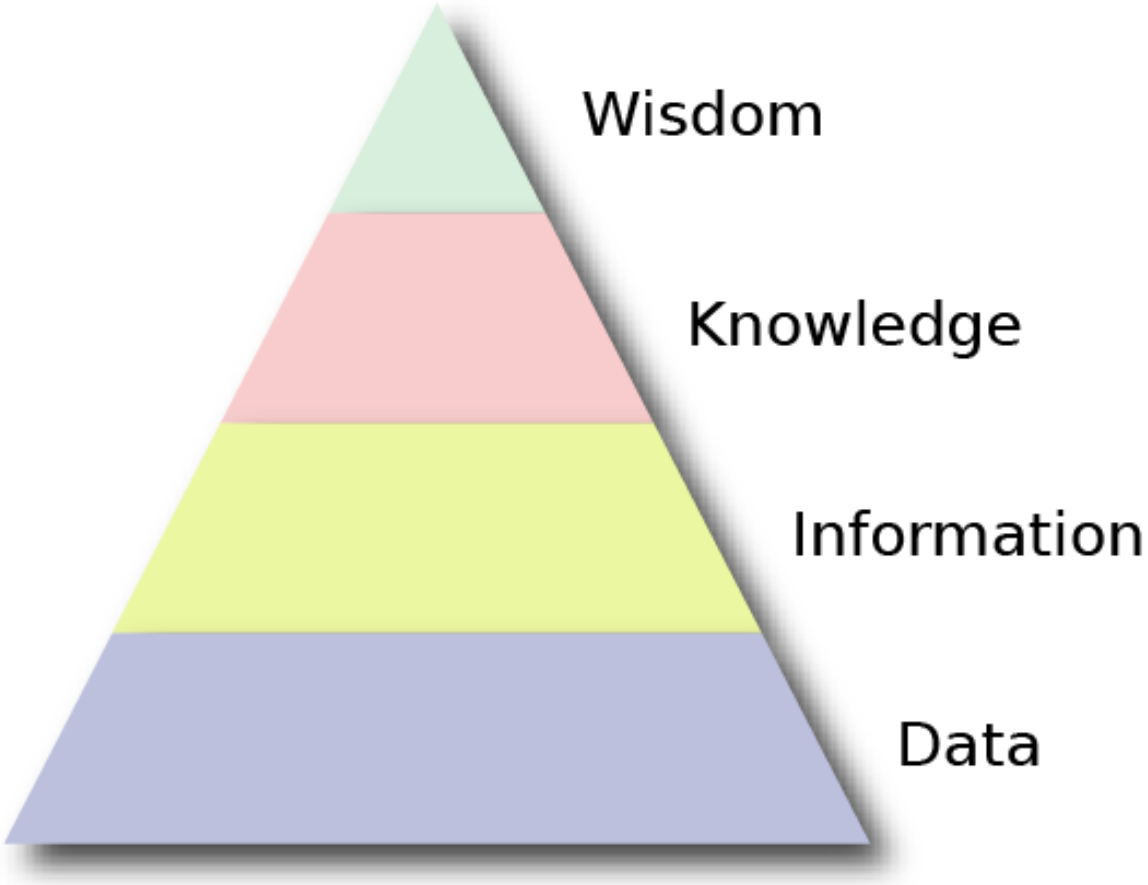
- Data
- Information
- Knowledge
- Wisdom

Zeleny, M. “Management Support Systems: Towards Integrated Knowledge Management” Human Systems Management, 1987, 7, 59-70

What is Information?

Image originally published in the December 1982 issue of THE FUTURIST, taken from http://www-personal.si.umich.edu/~nsharma/dikw_origin.htm

The DIKW Hierarchy

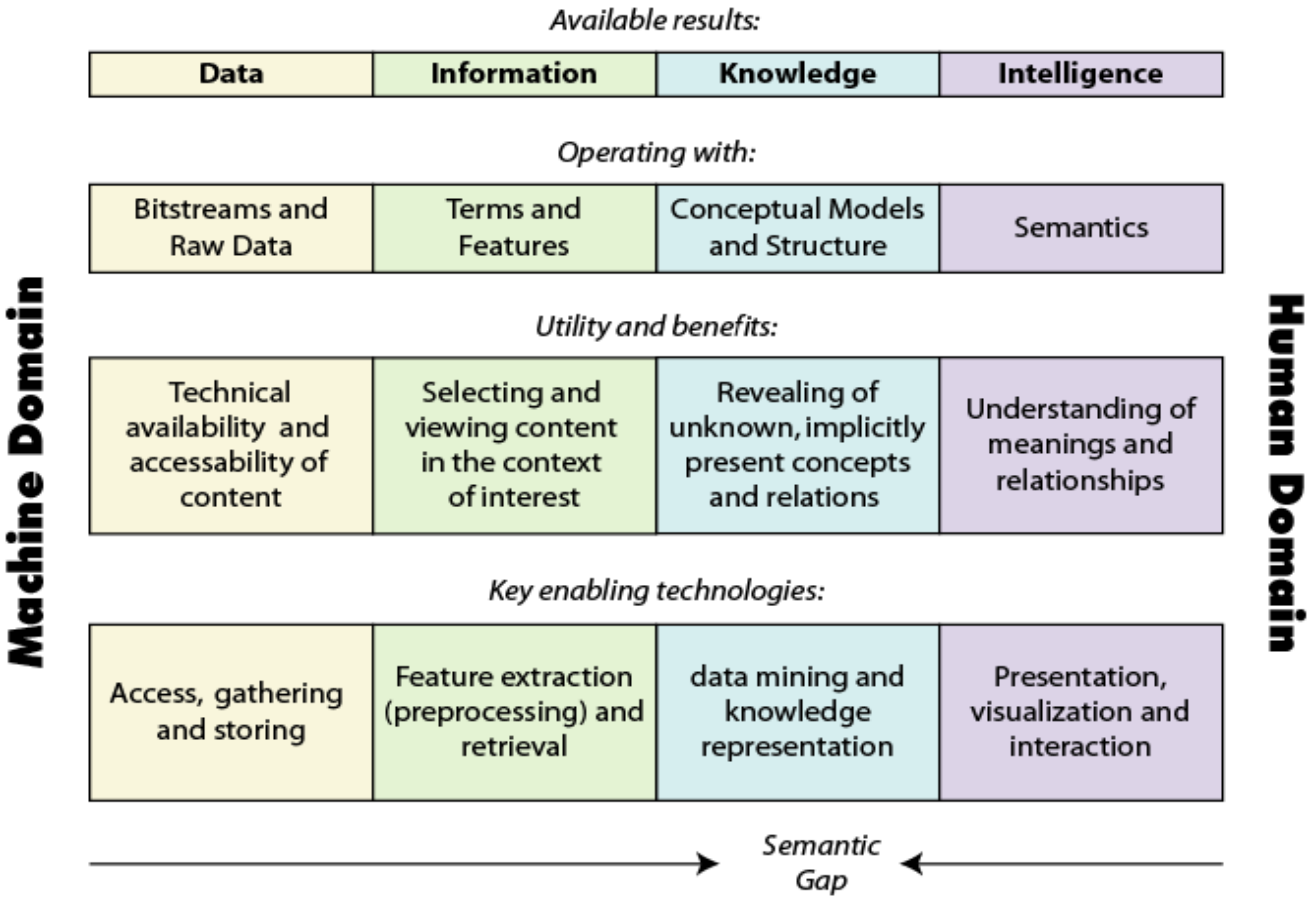


The DIKW Hierarchy

Definition of the DIKW levels:

Data	Know nothing
Information	Know what
Knowledge	Know how
Wisdom	Know why

Modified DIKW (IBM)



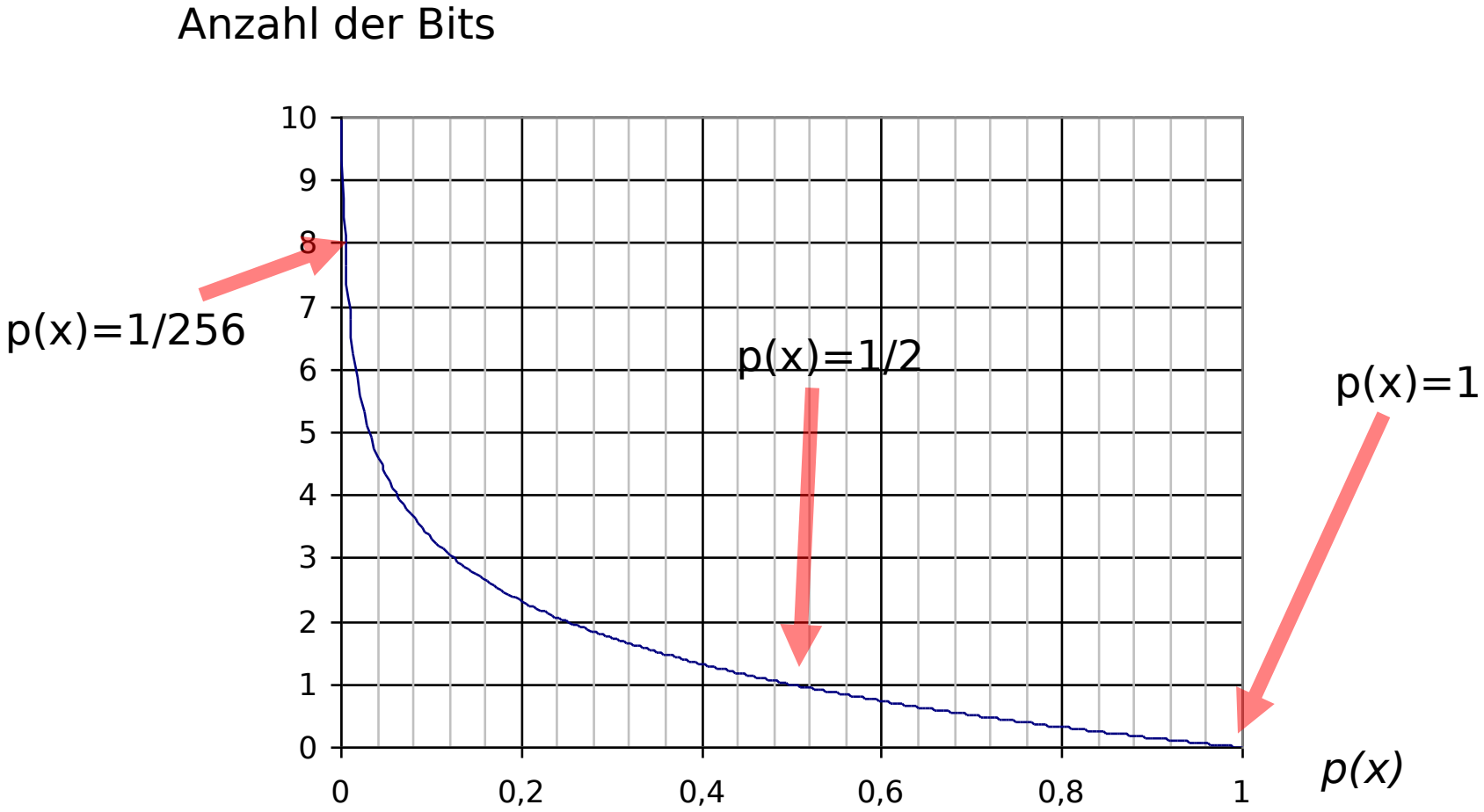
What is Information?

Shannon's Information Theory

- Problem: Communication over a noisy channel
- Fundamental finding:
  - Information Content (measured in bits) of an Event (e.g. letter) depends on the entropy (probability of occurrence)

$$I(A_n) = \log_2\left(\frac{1}{P(A_n)}\right) = -\log_2(P(A_n))$$

Shannon's Information Theory



Grice's Maxims of Conversation

- As informative as required
- As correct as possible
- Relevant to the aims of the conversation
- Contribution should be clear, unambiguous and concise

Haupmann, A. G. & Witbrock, M. J. "Story Segmentation and Detection of Commercials in Broadcast News Video" ADL '98: Proceedings of the Advances in Digital Libraries Conference, IEEE Computer Society, 1998

Agenda

- What is **Information**?
- What are **Information Systems**?
- Information Overload
- Current state in MuMe Consumption

What are Information Systems?

- Systems for handling information
  - Collect, Store & Organize
  - Process, Disseminate & Transmit
- Three main parts in these systems
  - People,
  - Machines &
  - Methods

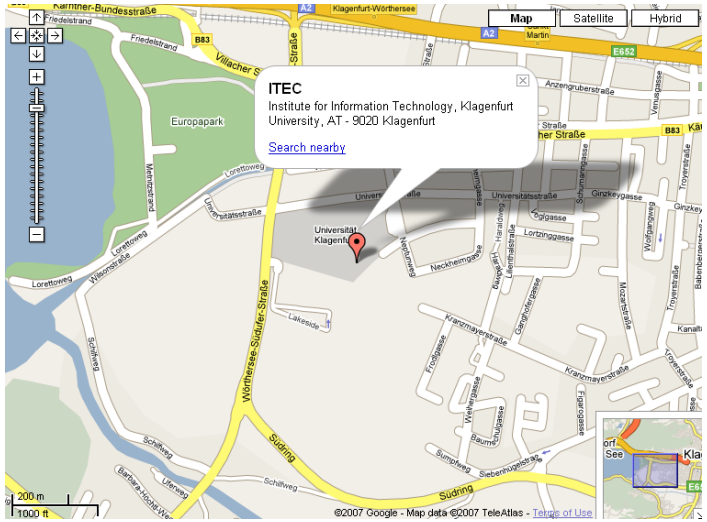
MEMEX

Memory **Ex**tender – Vannevar Bush

- Published in 1945 (Atlantic Monthly)
- An electromechanical device for
  - Viewing books and films
  - Adding information and comments
  - Interlinking information
  - Browsing links
- MEMEX is an early hypertext system.

Geographic IS

- Focus on **Spatially Referenced Data**
  - Coordinates, Height
  - Distance, Inclusion, Neighbouring
  - Hierarchical organisation



taken from Google Maps

Multimedia IS

- Focus on Multimedia Data & Metadata
  - Storage, Transmission
  - Search & Retrieval
  - Organization & Dissemination
- Media types
  - Textual / Visual / Auditive / Haptic / Olfactory
  - Rastered or Rendered / Modelled
    - Midi vs. MP3
    - VRML vs. PNG
    - LASER vs. MPEG-2

Agenda

- What is **Information**?
- What are **Information Systems**?
- Information Overload
- Current state in MuMe Consumption

Information Overload

- 5 Exabytes of new information in 2002.
  - 92% of the new information was stored on magnetic media, mostly in hard disks.
  - That’s 800 MB per person on the globe
  - That’s 37.000 times the LoC
  - That’s 30% more than in 1999

Lyman, Peter and Hal R. Varian, "How Much Information", 2003. Retrieved from http://www.sims.berkeley.edu/how-much-info-2003 on [2007-02-07]

Information Overload

18 Exabytes of new information in information flow in 2002

- 98% of new information generated by phone calls
- Most radio and TV broadcast content is **not new** information.
  - ~ 70 out of 320 million h of radio is new, that's 3.500 TB
  - ~ 31 out of 123 million h of TV are new, that's 70.000 TB

Information Overload

Information Flow (ctd.)

- IM: 5 billion messages / day or 274 TB p.a.
- Email: 400.000 TB p.a.
- P2P: growing, but not yet estimated,
  - significant traffic has been observed on different backbones
  - Ranges from 20%-60% are mentioned

Information Usage

An average American adult:

- Telephone - 16.17 hours a month
- Radio - 90 hours a month
- TV - 131 hours a month
- 53% of the U.S. uses the Internet
  - ~ 25.5 h / month at home
  - ~ 74.5 h / month at work

Agenda

- What is **Information**?
- What are **Information Systems**?
- Information Overload
- Current state in MuMe Consumption

Current State in MuMe Consumption

- Digital Photography
  - Still images
- Digital Video in General
  - Streaming and download

Digital Imaging Devices (global)

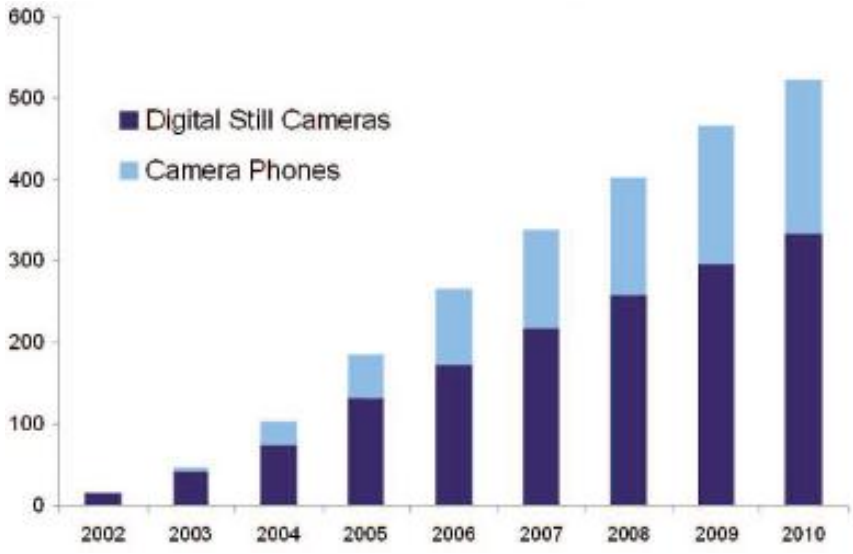
- First question: How many devices exist?

Device	# in 2006
digital cameras	$400 * 10^6$
camera phones	$600 * 10^6$

Source: IDC Study “Expanding Digital Universe” http://www.emc.com/about/destination/digital_universe/

Number of Digital Photos (global)

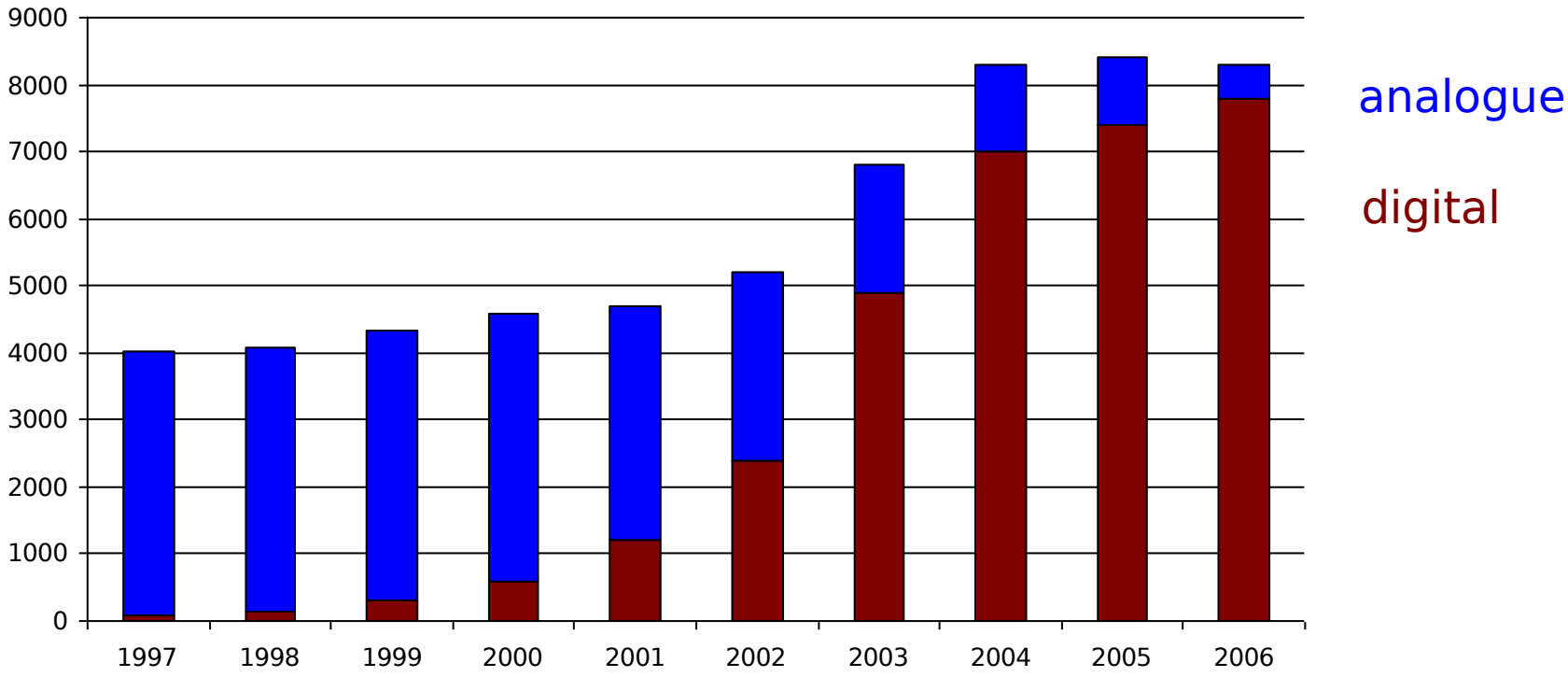
- Estimate 2006
  - > 150 billion photos from cameras
  - > 100 billion photos from camera phones
- Forecast 2010
  - > 500 billion photos
  - + increased resolution



Source: IDC Study “Expanding Digital Universe” http://www.emc.com/about/destination/digital_universe/

Digital Imaging Devices (Germany)

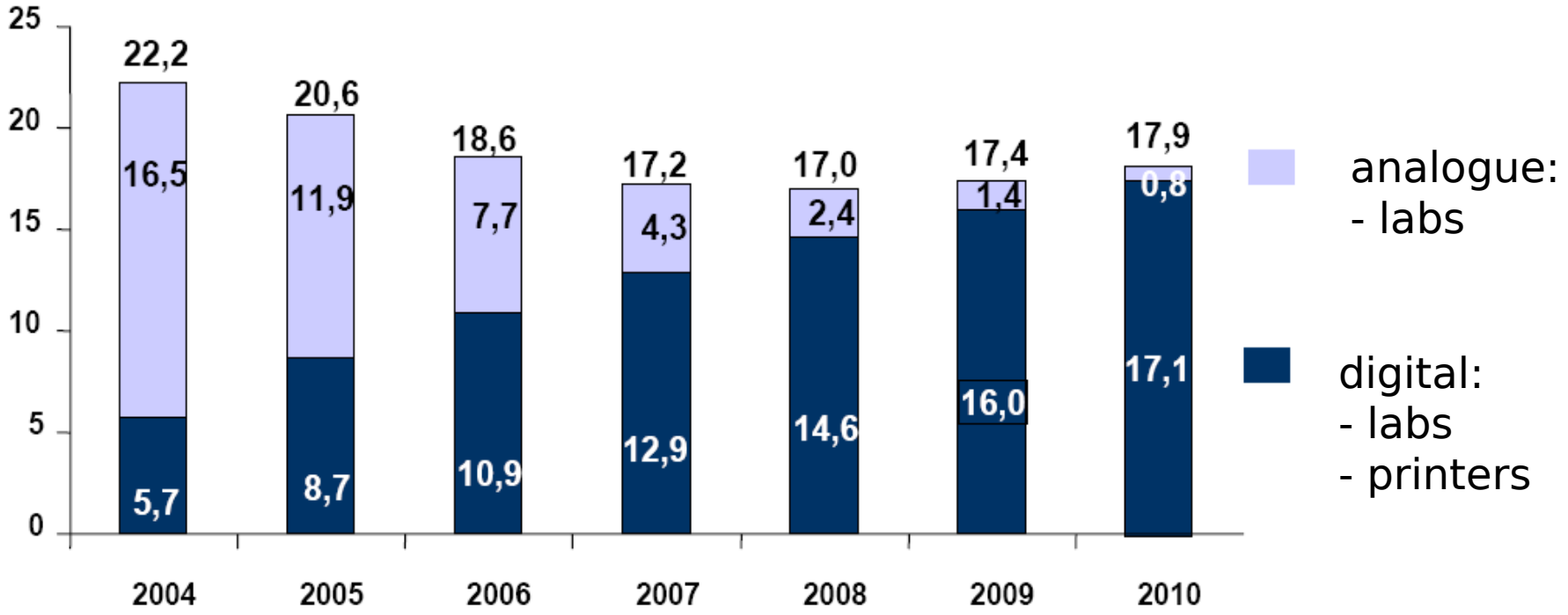
Still image cameras sold in Germany (thousands)



Source: Cewe Factbook, http://www.cewecolor.de

Photo prints market (Western Europe)

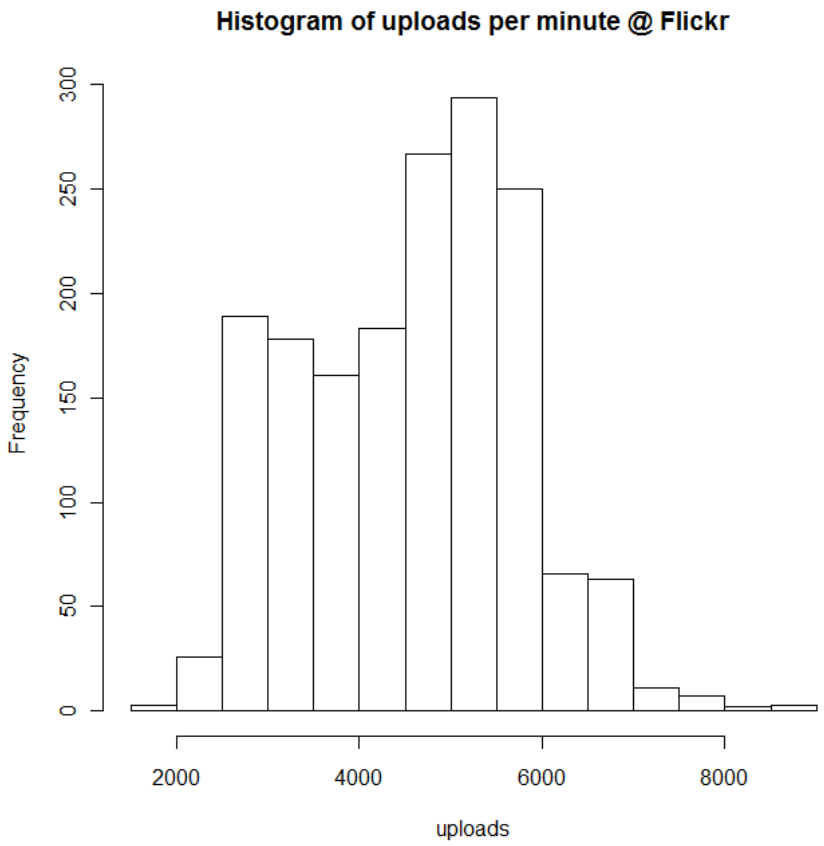
- Photo prints forecast (in billions)



Source: Cewe Factbook, http://www.cewecolor.de

Flickr

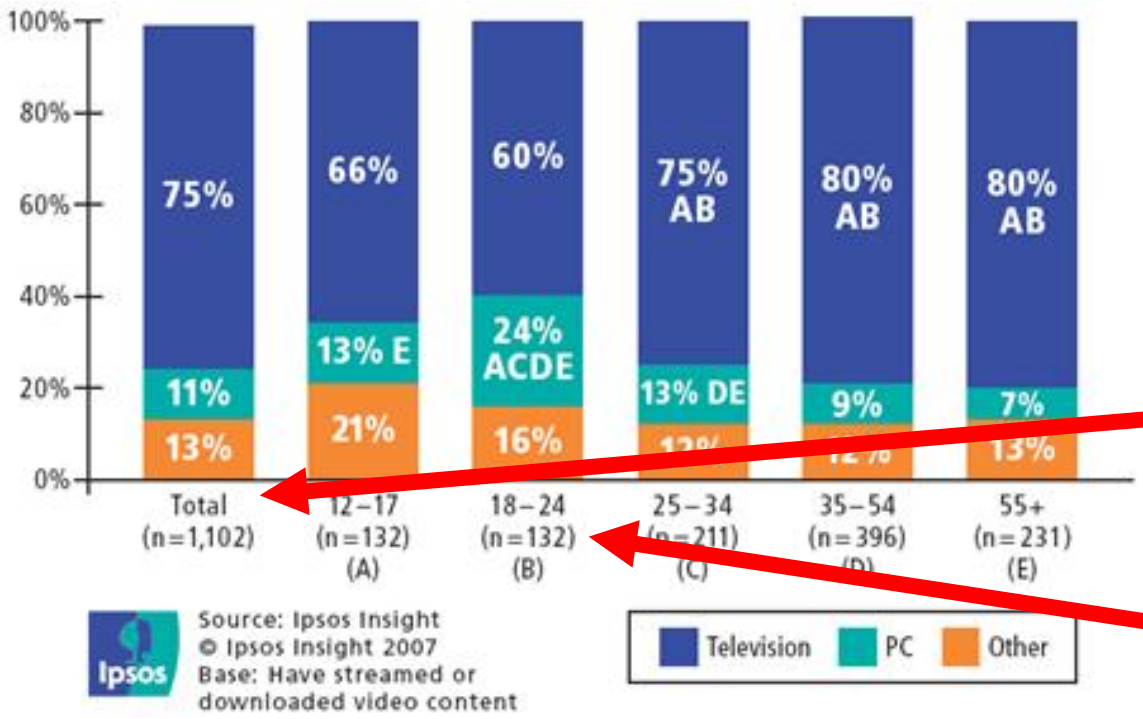
- 3 billionth upload in November 2008
- Currently ~4,800 uploads per minute



Content

- What is **Information**?
- What are **Information Systems**?
- Information Overload
- Current state in MuMe Consumption
  - Digital Photography
  - Digital Video in General

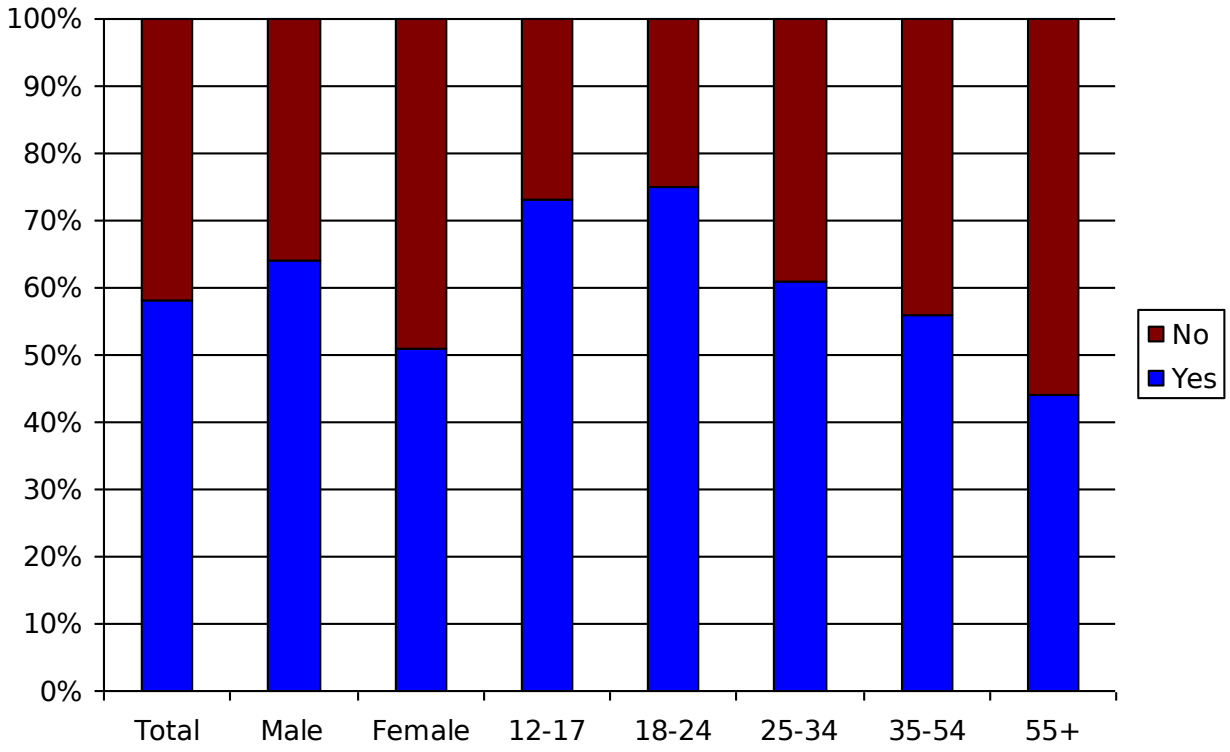
Where video content is watched? (US, 2007)



- Among people having downloaded / streamed content
- 11% in total watched on PC
- 24% in age 18-24

Source: Ipsos Insight's 2007 MOTION Study - http://www.ipsosinsight.com/pressrelease.aspx?id=3500

Ever Streamed a File Off of the Internet? (US, 2007)



- 75% of the 12-24 years old.
- More than half in total.

Source: Ipsos Insight's 2007 MOTION Study - http://www.ipsosinsight.com/pressrelease.aspx?id=3500

Short vs. Long Clips (US, 2007)

- Short clips are preferred
  - 3/4 of streamers have streamed short news or sports clips
  - 2/3 of streamers have streamed amateur or homemade clips
- Also due to YouTube ...
  - 40% of the streamers use YouTube
  - Common restrictions in video size & length

Source: Ipsos Insight's 2007 MOTION Study - http://www.ipsosinsight.com/pressrelease.aspx?id=3500

Most Common Barrieres for payed DLs (US, 2007)

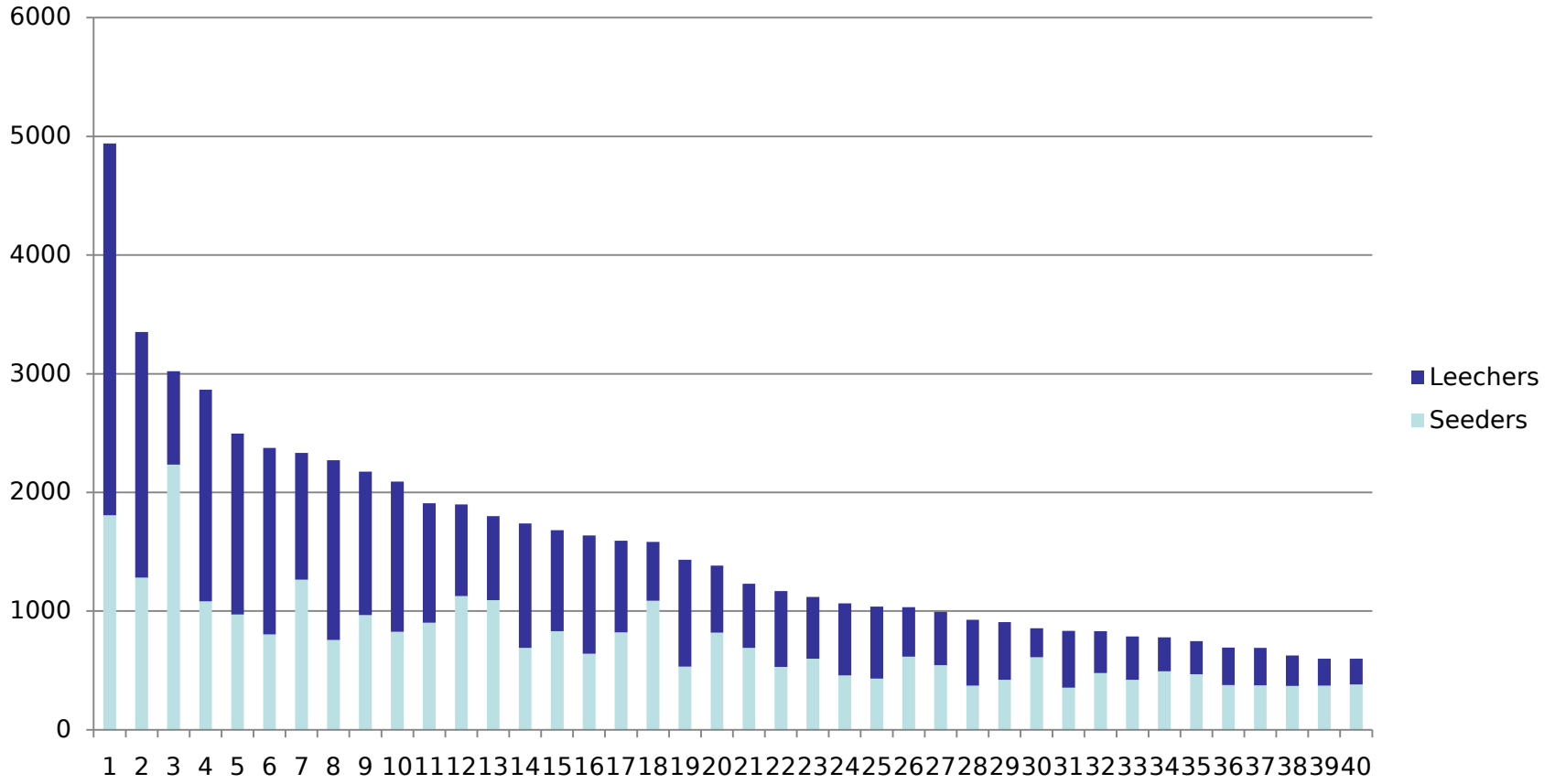
- Unwillingness to pay for this content,
- Difficulty or inability to
  - burn these files onto DVD
  - watch this content on their living room TV.

But there are also benefits in the future:

- 24/7 access
- reasonable fees or free (ad supported)

Bittorrent Movies

(query: jaybob, 2010-03-09)



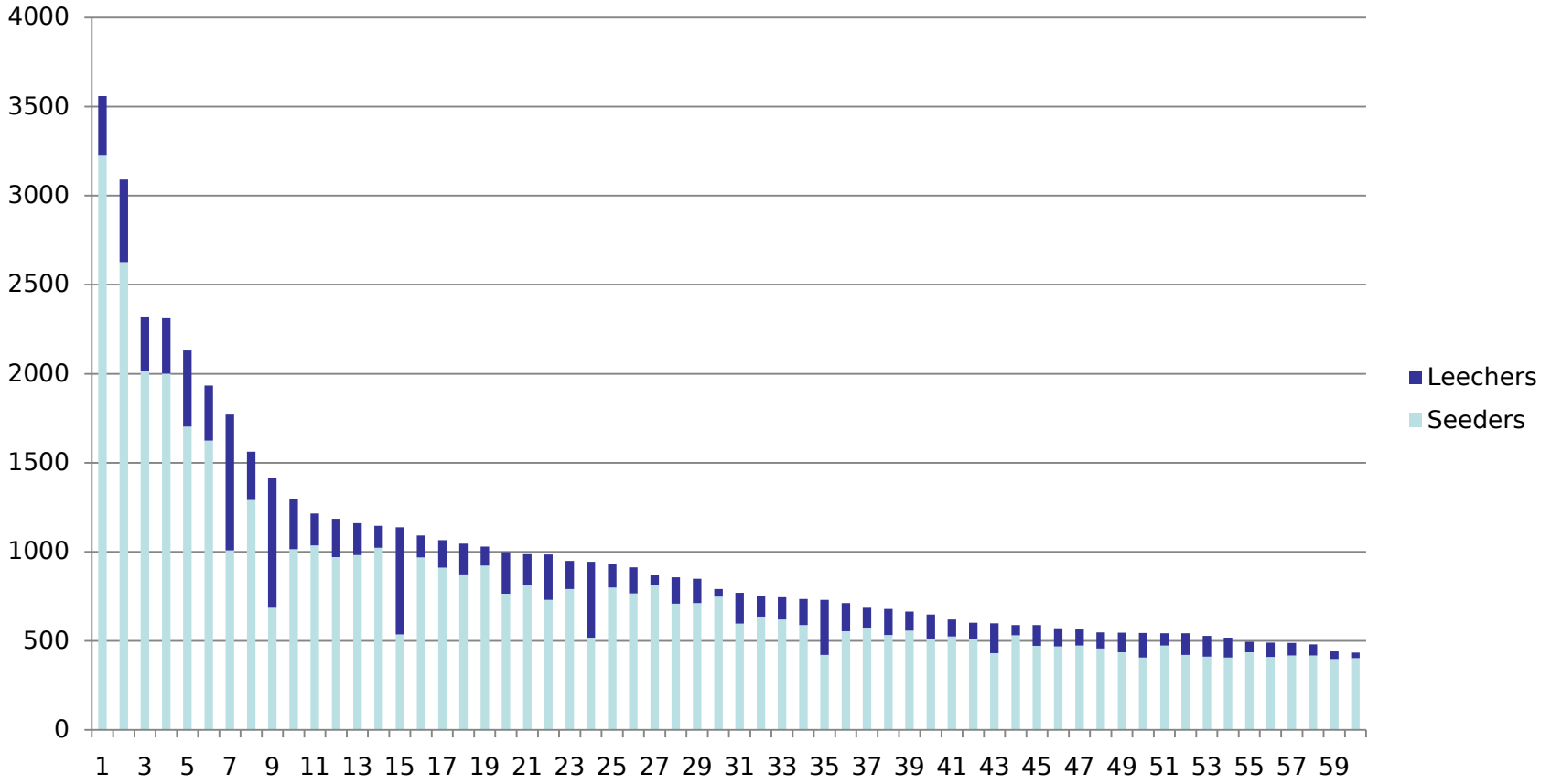
Bittorrent Movies

(query: jaybob, 2010-03-09)

	Seeders	Leechers
+25 40[DL] New Moon {2009} DVDRIP. Jaybob	1808	3133
+38 46[DL] Ninja Assassin {2009} DVDRIP. Jaybob	1281	2071
+177 268[DL] Paranormal Activity {2009} DVDRIP. Jaybob	2235	787
+15 16[DL] The Men Who Stares At Goats {2009} DVDRIP. Jaybob	1081	1785
+19 38[DL] Legion {2010} DVDRIP. Jaybob	970	1526
+14 13[DL] Did You Hear About The Morgans {2009} DVDRIP. Jaybob	804	1571
+109 65[DL] The Time Traveler's Wife {2009} DVDRIP. Jaybob	1265	1067
+15 19[DL] The Fourth Kind {2009} DVDRIP. Jaybob	757	1513
+15 17[DL] Up in the Air {2009} DVDRIP Jaybob	967	1207
+10 9[DL] Precious {2009} DVDRIP.Jaybob	825	1265
+59 46[DL] Cirque du Freak The Vampire's Assistant {2009} DVDRIP Jay	902	1006
+64 47[DL] The Informant! {2009} DVDRIP Jaybob	1126	772
+113 164[DL] 2012 {2009} DVDRIP. Jaybob	1091	709
+6 7[DL] Astro Boy {2009} DVDRIP.Jaybob	689	1049
+30 25[DL] Old Dogs {2009} DVDRIP Jaybob	831	850
+6 9[DL] The Descent Part 2 {2009} DVDRIP.Jaybob	641	995
+77 47[DL] Where The Wild Things Are {2009} DVDRIP. Jaybob	820	773
+179 125[DL] 9 {2009} DVDRIP. Jaybob	1087	496
+11 9[DL] Fantastic Mr Fox {2009} DVDRIP. Jaybob	532	900

Bittorrent Movies

(query: axxo, 2010-03-09)



Bittorrent Movies

(query: axxo, 2010-03-09)

	Seeders	Leechers
+857 831Fantasy	3229	331
+48 30Whip It[2009]DvDrip-aXXo	2627	464
+724 515Action	2016	305
+46 21Fame[2009][Extended Edition]DvDrip-aXXo	2001	311
+359 295Iron.Man[2008]DvDrip-aXXo	1703	427
+263 165300[2006]DvDrip[Eng]-aXXo	1624	309
+7 11The.Hurt.Locker[2008]DvDrip-aXXo	1008	762
+382 176Sex.Drive[2008][Unrated.Edition]DvDrip-aXXo	1291	270
+80 59Star.Wars.I-VI.Collection.DvDrip-aXXo	685	731
+32 13Action	1014	283
+422 219Animation	1035	179
+171 73Tropic.Thunder[2008]DvDrip-aXXo	969	216
+113 86Orphan[2009]DvDrip-aXXo	982	178
+462 269Pineapple.Express[2008]DvDrip-aXXo	1022	124
+66 53Alien.Quadrilogy.DvDrip-aXXo	536	601
+406 242Role.Models[2008][Unrated.Edition]DvDrip-aXXo	968	124
+383 202Hancock[2008]DvDrip-aXXo	910	155
+418 287Wanted[2008]DvDrip-aXXo	872	174
+399 253Step.Brothers[2008][Unrated.Edition]DvDrip-aXXo	922	107

Readings

C.J. van Rijsbergen: *Information Retrieval – Introduction*, London, Butterworth, 1979

Available on homepage: mmis08 / multimedia

Your task:

- Read + answer questions
- Send me an **email** with the answers until next course.

Questions:

- What is the difference between Data Retrieval and Information Retrieval?
- What does “relevance” mean in the context of Information Retrieval?

Thanks ...

... for your attention!