



Foto: Thomas Josek

# Basisinformationstechnologie II

Sommersemester 2023. Zusammenfassung II.  
*Basierend auf Jan Wieners' Folien*

# Ausblick: Seminarthemen BIT II (SoSem 2023)

- Betriebssysteme
  - Verknüpfung Hard- und Software, Aufgaben von Betriebssystemen, Prozesse, Multitasking, Speicher- und Dateiverwaltung
- Rechnerkommunikation / Computernetzwerke
  - Themen u.a.: Hardwareaspekte, Übertragungstechnik, Verbindungsarten, Dienste, Protokolle, Ports, Schichtenmodelle: ISO/OSI vs. TCP / IP Modell, ...
- Text, Datenmodellierung, Metadaten(standards)
  - Themen u.a.: XML, PDF, DocX & Co., Metadatenstandards im BAM-Sektor, Semantic Web, RDF, CIDOC CRM
- Bild
  - Themen u.a.: Farbmischung, Bildbearbeitungs-/verarbeitungsalgorithmen, Kompression, ...
- Künstliche Intelligenz
  - Geschichte
  - Maschinelles Lernen, Q-Learning
  - KI in Computer- und Videogames

## Betriebssysteme

- Metaphern
  - „Datei“
  - „Ordner“
  - „Desktop“
- Dateiverwaltung
- Prozessverwaltung
- (Arbeits)Speicherverwaltung

# Blickwinkel auf OS (Tanenbaum)

## Bottom-Up

„Die Aufgabe eines Betriebssystems besteht darin, eine geordnete und kontrollierte Zuteilung der Prozessoren, Speicher und Ein-/Ausgabegeräte an die um sie konkurrierenden Programme durchzuführen“

(vgl. Tanenbaum: Moderne Betriebssysteme. 2009, S. 35 ff.)

→ Ein Betriebssystem übernimmt die Verwaltung aller Bestandteile eines komplexen Systems

## Top-Down

Betriebssysteme verwandeln [...] etwas Hässliches in etwas Schönes [...].“

(vgl. Tanenbaum: Moderne Betriebssysteme. 2009, S. 34)

→ Programmiersicht: Ein Betriebssystem stellt Abstraktionen für Anwendungsprogramme bereit

## (Interface / GUI)

### Bereitstellung elementarer Dienste und Ressourcen

- **Dienste**, die (vereinfachten) Zugriff auf Ressourcen ermöglichen.
- **Ressourcen**
  - Rechenzeit (CPU / GPU)
  - Hauptspeicher
  - Festplatten- / Wechselmedienspeicherplatz
  - Interne Geräte wie Erweiterungskarten, Onboard-Komponenten (z.B. Onboard Soundkarte, Onboard LAN)
  - Externe Geräte, z.B. über USB mit dem Rechner verbunden: Scanner, Drucker

# Dateizugriff: Adressierungsarten

## Absolute Adressierung

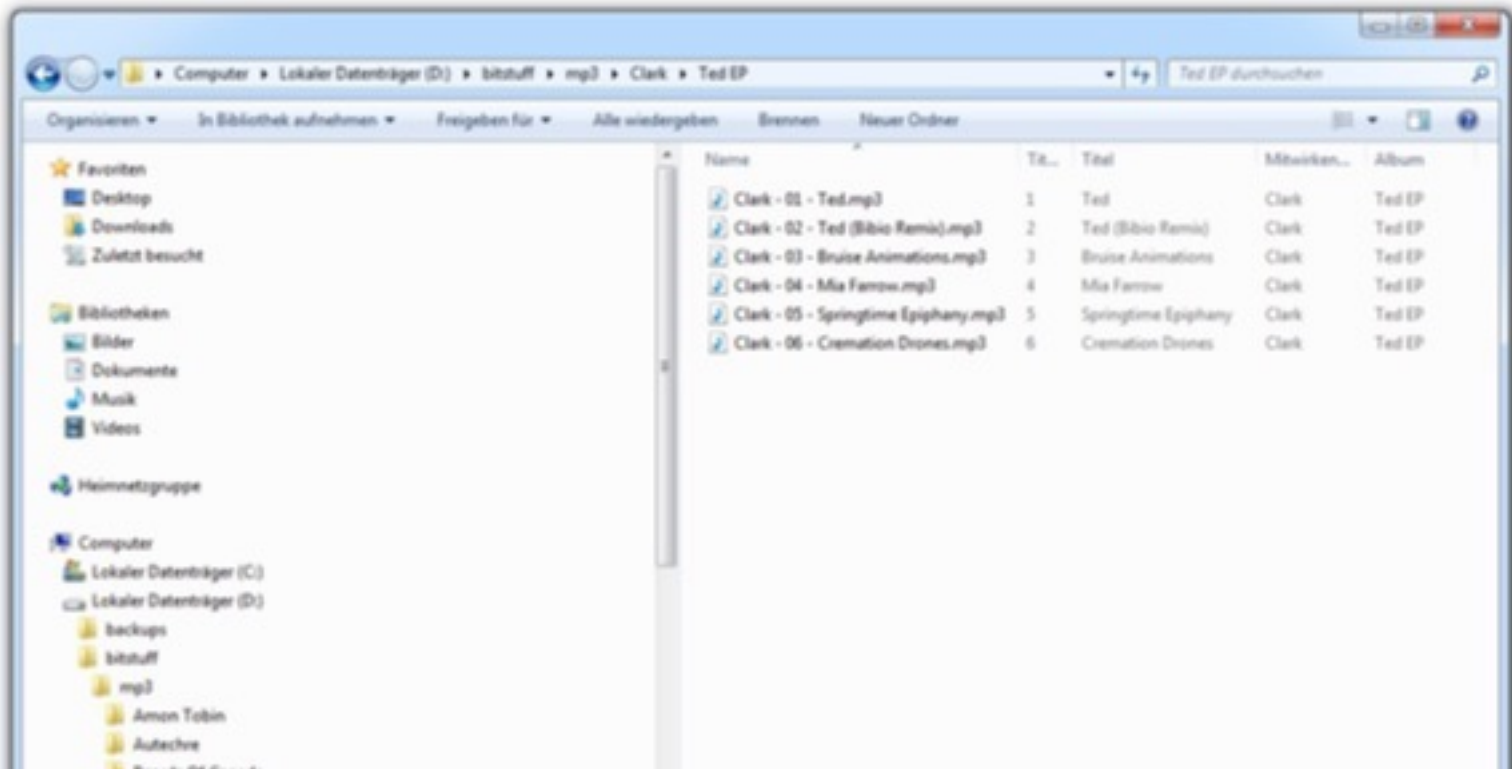
D:\bitstuff\mp3\Clark\Ted EP\Clark - 01 - Ted.mp3

## Relative Adressierung

./Clark – 01 - Ted.mp3 (dasselbe Verzeichnis)

../\*.mp3 (eine Ebene / Ordner höher)

../../\*.htm (zwei Ebenen höher)



# NTFS: MFT

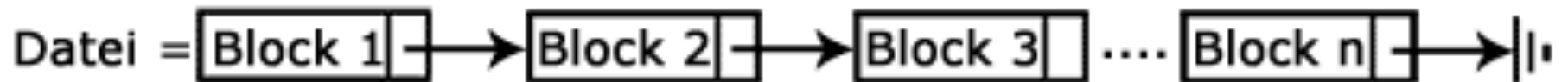
Wichtigste Datenstruktur jedes Volumes:  
Masterdateitabelle (MFT, Master File Table)

- Lineare Folge von festgelegten 1-KB-Einträgen
- Jeder MFT-Datensatz beschreibt eine Datei oder ein Verzeichnis
  - → Enthält die Attribute der Datei (i.e. Name, Zeitstempel, Liste von Adressen, die angeben, wo auf der Platte die zugehörigen Blöcke stehen)
- MFT ist selber eine Datei, kann somit irgendwo auf der Partition liegen

# Dateioperationen

Grundmenge an (elementaren) Operationen zur Verwaltung von Dateien:

- Neu
- Löschen (unlink)
- Kopieren
- Umbenennen
- Verschieben



→ „Einfach verkettete Liste“



# Betriebssystem: Zentrale Aufgaben

## Dateiverwaltung ✓

- Abstraktionen / Schnittstellen für Zugriff auf Low-Level Funktionen (z.B. der Festplatte) ✓
- Blöcke / Cluster ✓
- Fragmentierung ✓

Prozessverwaltung / Ressourcenverwaltung /  
Zeitplanung → Scheduling

Speicherverwaltung

# Prozess

**Prozess** = „Aktivität jedweder Art“

Prozess umfasst:

- Ein Programm
  - Eingaben
  - Ausgaben
  - Zustand
- 
- Mehrere Prozesse können sich einen einzelnen Prozessor teilen (Analogie: Kuchenbäcker) oder auf mehrere Prozessoren verteilt werden
  - Eine **Schedulingstrategie** entscheidet, wann die Arbeit an einem Prozess unterbrochen und ein anderer Prozess bedient wird.

# Prozessunterbrechung

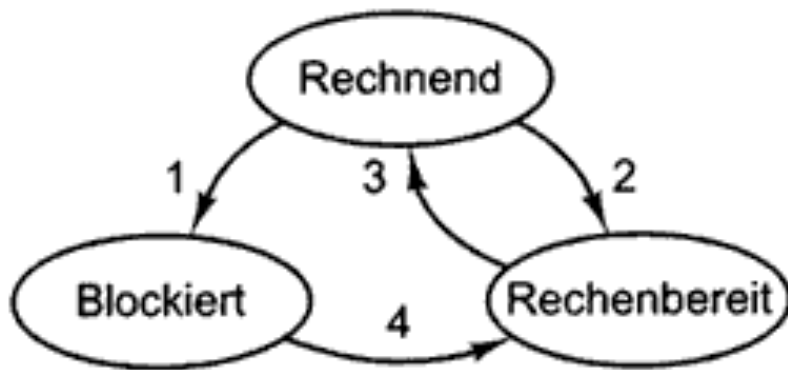
Jedem **Prozess** wird vom Betriebssystem ein **eigener Adressbereich** zugewiesen

Wird ein Prozess deaktiviert, müssen folgende Informationen gespeichert werden, um ihn später reaktivieren zu können:

- der **Programmcode** des Prozesses
- seine im **Arbeitsspeicher befindlichen Daten**
- der **Inhalt der CPU-Register** einschließlich des **Befehlszählers**
- eine **Tabelle aller geöffneten Dateien** mit ihrem **aktuellen Bearbeitungszustand**

# Prozessverwaltung

Mehrere (rechenbereite) Prozesse konkurrieren um Rechenzeit



1. Prozess blockiert wegen Eingabe
2. Scheduler wählt einen anderen Prozess
3. Scheduler wählt diesen Prozess
4. Eingabe vorhanden

Ist nur eine CPU im System vorhanden, muss entschieden werden, welcher Prozess als nächstes läuft.

Der Teil des Betriebssystems, der diese Wahl trifft, wird als **Scheduler** bezeichnet (kommt beim präemptiven Multitasking zum Einsatz)

# Speicherverwaltung

Aufgabe des Betriebssystems: Versorgung der Prozesse / Threads mit Arbeitsspeicher  
→ Memory Manager

Prozesse müssen vom OS vor gegenseitiger Beeinträchtigung durch fehlerhafte Adressierung gemeinsam benutzter Speicherbereiche geschützt werden

Wesentliche Verfahren zur Speicherverwaltung:

- **Swapping**
- **Paging**

# Speicherverwaltung

Grundproblem: Was geschieht, wenn der physikalische Arbeitsspeicher voll ist?

- Lösung I (anno pief): Programmierer teilen ihre Programme manuell in kleine Segmente (sog. Overlays) auf. Overlays wurden auf der Festplatte gespeichert und bei Bedarf dynamisch vom Betriebssystem ein- und ausgelagert.
- Lösung II: **Virtueller Speicher**
  - **Swapping**
  - **Paging**

## Kommunikation

- Basales Kommunikationsmodell
  - Sender / Empfänger
  - Adressierung
  - Nachricht
  - Übertragungsmedium
- Komplexere (und flexiblere) Kommunikation
  - Schichtenmodell
  - Verschlüsselung
  - Routing

# Kommunikation

Broadcasting

→ Analogie: Lautsprecherdurchsage am Flughafen

Multicasting

→ Übertragung an eine Teilmenge  
(der angeschlossenen Rechner)

Unicasting

→ Punkt-zu-Punkt Übertragung  
(genau zwei Kommunikations-  
partner)





# Dienste und Protokolle

**Dienst:** Gruppe von Operationen, die eine Schicht der über ihr liegenden Schicht zur Verfügung stellt.

→ Beziehen sich auf Schnittstellen zwischen den Schichten

**Protokoll:** Menge von Regeln

→ Beziehen sich auf Pakete, die zwischen gleichgestellten Einheiten auf verschiedenen Rechnern versendet werden

# Dienstprimitive

Dienstprimitive	Bedeutung
LISTEN	Blockade, auf eingehende Nachricht wartend
CONNECT	Verbindung zu gleichgestellter Einheit aufbauen
RECEIVE	Blockade, auf eingehende Nachricht wartend
SEND	Nachricht an eine gleichgestellte Einheit senden
DISCONNECT	Verbindung beenden

**Abbildung 1.17:** Fünf Dienstprimitive zur Implementierung eines einfachen verbindungsorientierten Dienstes

# ISO / OSI Referenzmodell

## OSI-7-Layer-Model (Open Systems Interconnection Reference Model)

Begriffe: Englisch - Deutsch

- |                      |  |
|----------------------|--|
| 1 Application Layer  | - Anwendungsschicht                    |
| 2 Presentation Layer | - Darstellungsschicht                  |
| 3 Session Layer      | - Sitzungs- bzw. Kommunikationsschicht |
| 4 Transport Layer    | - Transportschicht                     |
| 5 Network Layer      | - Netzwerk- bzw. Vermittlungsschicht   |
| 6 Data Link Layer    | - Sicherungsschicht                    |
| 7 Physical Layer     | - Bitübertragungsschicht               |

PC im Netzwerk  
A



W <http://www.wikipedia.org>

Der Benutzer empfängt lediglich die Antwort des Servers ("wikipedia.org"-Startseite). Im Allgemeinen bekommt er von der Schachtelung seines Seitenaufrufs durch die Ebenen seines PCs (abwärts) und vom Parsen der Antwort des Servers zurück durch die Ebenen seines PCs (aufwärts) nichts mit!

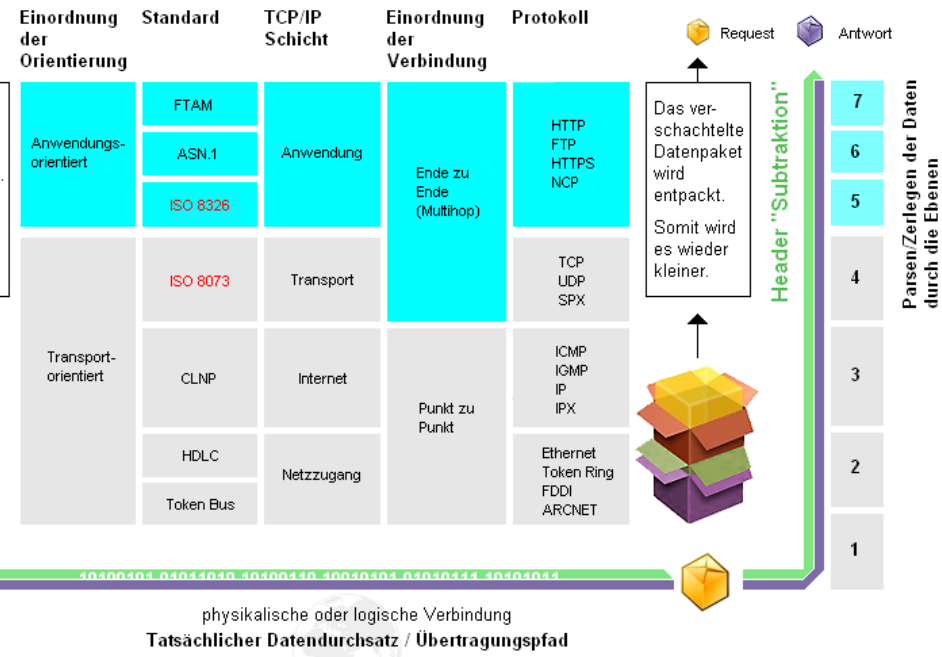
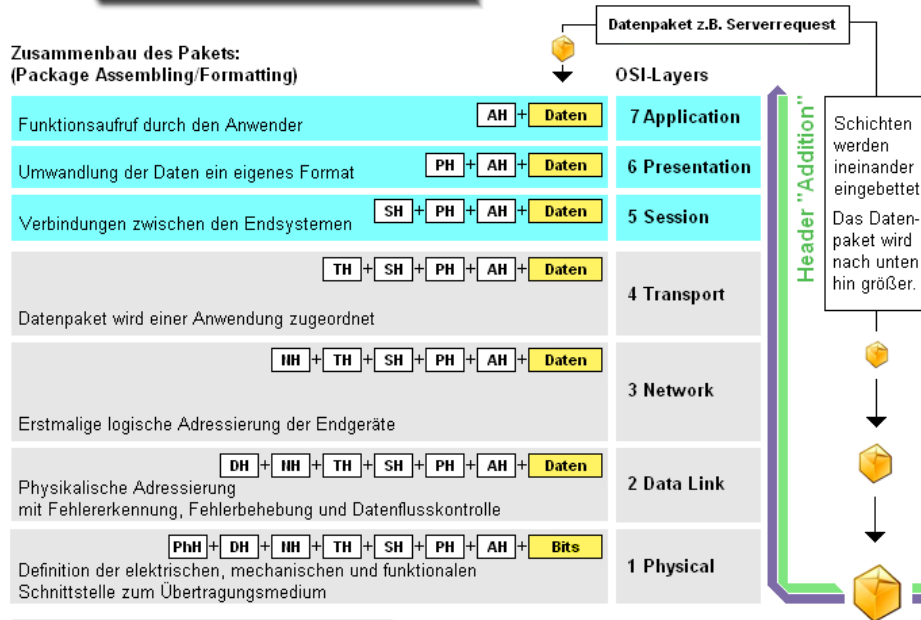


Server schickt die entsprechenden Daten über die selbe Methode zurück. (s. u.)

Server im Netzwerk  
A

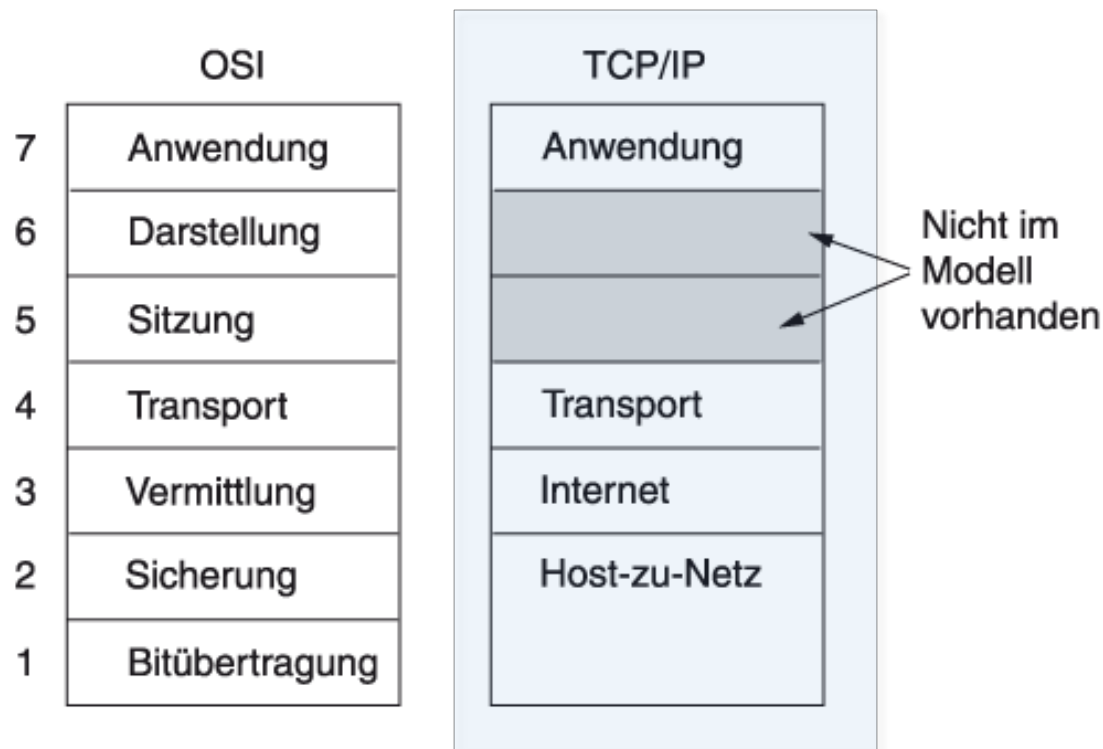


### Zusammenbau des Pakets: (Package Assembling/Formatting)



Zusammensetzung der Abkürzungen oben:  
Anfangsbuchstabe der Schicht und "H" für Header.  
z.B. Application Header = AH

# Referenzmodell II: TCP / IP



**Abbildung 1.21:**  
Das TCP/IP-Referenzmodell

# Rechnerkommunikation II

- „Warriors of the Net“
- HTTP
- HTTPS
- Anwendungsschicht
  - Email
    - SMTP
    - POP3
    - IMAP

HTTP (Hypertext Transfer Protocol)	
<b>Familie:</b>	Internetprotokollfamilie
<b>Einsatzgebiet:</b>	Datenübertragung, Hypertext u. a.
<b>Port:</b>	80/TCP
HTTP im TCP/IP-Protokollstapel:	
<b>Anwendung</b>	HTTP
<i>Transport</i>	TCP
<i>Internet</i>	IP (IPv4, IPv6)
<i>Netzzugang</i>	Ethernet, Token Bus, Token Ring, FDDI ...
<b>Standards:</b>	RFC 1945 <a href="#">↗</a> (HTTP/1.0, 1996) RFC 2616 <a href="#">↗</a> (HTTP/1.1, 1999)

TCP: Transmission Control Protocol, Verbindungsorientiertes Protokoll

IPv4: 134.95.115.23

IPv6: Hex.-Not., 8 Blöcke, je 16 Bit

Ethernet, u.a.:

- ISO/OSI Modell:
  - Schicht 1 (Physik. Schicht) und
  - Schicht 2 (Sicherungsschicht)
- TCP/IP:

Ethernet im TCP/IP-Protokollstapel:

<b>Anwendung</b>	HTTP	IMAP	SMTP	DNS	...
<i>Transport</i>	TCP		UDP		
<i>Internet</i>	IP (IPv4, IPv6)				
<b>Netzzugang</b>	Ethernet				

# HTTP: Uniform Resource Locator (URL)

Drei Standards:

- HTTP
- HTML
- URLs



IP-Adresse herausfinden?

Leeren Dauerhaft Alles HTML CSS JavaScript XHR Bilder Plug-ins Medien Schriften

URL	Status	Domain	Größe	Remote-IP
GET /page/2/	200 OK	designtagebuch.de	2,4 KB	178.16.57.240

Header Antwort HTML Cache Cookies

Antwort-Header Quelltext anzeigen

```

Cache-Control no-store, no-cache, must-revalidate,
Connection Keep-Alive
Content-Type text/html; charset=UTF-8
Date Sun, 08 May 2016 09:44:31 GMT
Expires Thu, 19 Nov 1981 08:52:00 GMT
Keep-Alive timeout=5, max=95
Link <http://www.designtagebuch.de/wp-json
Pragma no-cache
Server Apache
Transfer-Encoding chunked
X-Powered-By PHP/5.5.21-p10

```

Anfrage-Header Quelltext anzeigen

```

Accept text/html,application/xhtml+xml,application/javascript;q=0.9,*/*;q=0.8
Accept-Encoding gzip, deflate
Accept-Language de,en;q=0.7,en-US;q=0.3
Connection keep-alive
Cookie ct_cookies_test=013c20aebf8dcf33c4df91.1462700434; _gat=1; ct_timestamp=1462700434
Host www.designtagebuch.de
Referer http://www.designtagebuch.de/
User-Agent Mozilla/5.0 (Windows NT 10.0; WOW64; rv:44.0) Gecko/20100826 Firefox/44.0
X-Moz prefetch

```

## HTTP Request-Methoden

- GET – Ressourcen vom Server anfordern; die URL enthält alle benötigten Informationen, um die Ressourcen zu lokalisieren und an den Client zu senden.
- POST – Daten zur Verarbeitung an den Server senden.
- PUT – Ressource wird erstellt bzw. geändert, sofern sie bereits existiert.
- DELETE – Ressource löschen
- HEAD – Server veranlassen, Kopfinformationen der Nachricht erneut zu senden.

GET push?client=ca-pub-2964111343784013&srn=gd	204 No Co			
GET collect?v=1&_v=j41&aip=1...A-1449218-1&z=57	200 OK	google-analytics.com	33 B	64.15.112.200
GET collect?v=1&_v=j41&aip=1...-1449218-1&z=176	200 OK	google-analytics.com	35 B	64.15.112.200
GET collect?v=1&_v=j41&aip=1...A-1449218-1&z=57		google-analytics.com	0 B	
GET collect?v=1&_v=j41&aip=1...-1449218-1&z=176		google-analytics.com	0 B	
GET ads?client=ca-pub-296411...signtagebuch.de&dt	200 OK	googleads.g.doubleclick.net	26,8 KB	172.217.20.98
GET analytics.js	200 OK	google-analytics.com	10,7 KB	64.15.112.200
GET ca-pub-2964111343784013.js	200 OK	pagead2.googlesyndication.com	159 B	172.217.20.98
GET essb-fans.woff?53962973	200 OK	designtagebuch.de	8,7 KB	178.16.57.240
GET MiloWebPro-Medium.woff	200 OK	designtagebuch.de	43,1 KB	178.16.57.240





# Email: Protokolle & Co.

Absenden / Weiterleiten von Emails  
**SMTP** → Simple Mail Transfer Protocol

SMTP im TCP/IP-Protokollstapel:				
<b>Anwendung</b>	<b>SMTP</b>			
<i>Transport</i>	TCP			
<i>Internet</i>	IP (IPv4, IPv6)			
<i>Netzzugang</i>	Ethernet	Token Bus	Token Ring	FDDI ...
<b>Standard:</b>	RFC 5321 <a href="#">↗</a>			

Abholen von Emails

POP3 (Post Office Protocol Version 3)	
<b>Familie:</b>	Internetprotokollfamilie
<b>Einsatzgebiet:</b>	Abholen von E-Mail vom Provider
<b>Port:</b>	110/TCP 995/TCP (Verschlüsselt)
POP3 im TCP/IP-Protokollstapel:	
<b>Anwendung</b>	<b>POP3</b>
<i>Transport</i>	TCP
<i>Internet</i>	IP (IPv4, IPv6)
<i>Netzzugang</i>	Ethernet Token Bus Token Ring FDDI ...
<b>Standards:</b>	RFC 1939 <a href="#">↗</a> (POP3, 1996)

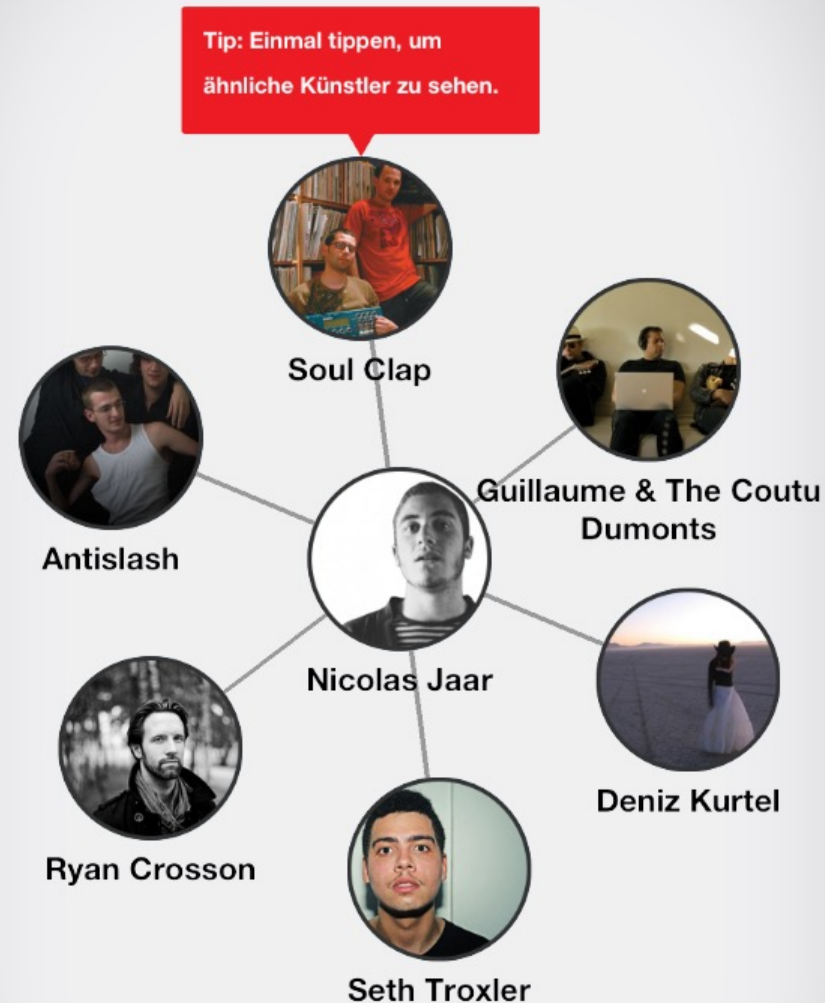
Internet Message Access Protocol	
<b>Familie:</b>	Internetprotokollfamilie
<b>Einsatzgebiet:</b>	Lesen und Verwalten von E-Mails
<b>Ports:</b>	143/TCP <sup>[1]</sup> 993/TCP (nur mit TLS)
IMAP im TCP/IP-Protokollstapel:	
<b>Anwendung</b>	<b>IMAP</b>
<i>Transport</i>	TCP
<i>Internet</i>	IP (IPv4, IPv6)
<i>Netzzugang</i>	Ethernet Token Bus Token Ring FDDI ...
<b>Standard:</b>	RFC 3501 <a href="#">↗</a>

# Semantic Web

# World Wide Web

„What’s wrong with the web?“ – die Grenzen des WWW

- „Finde Musik, die ich mögen könnte“
  - Knackpunkt:  
Hintergrundwissen –  
Welche Musik mag ich  
derzeit? (Musikgeschmack  
verändert sich mitunter)



Oh weh, was meint  
sie / er damit  
bloß???

Knackpunkt: Den Computermechanismen  
mangelt's an Wissen!

„knowledge gap“:

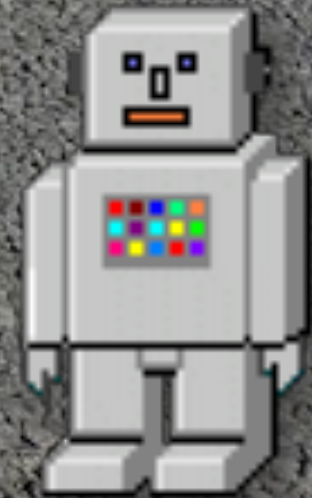
- Probleme im Verständnis natürlicher Sprache
- Interpretation des Inhaltes von Bildern oder anderen multimedialen Dingen
- Computer verfügt nicht über Hintergrundwissen über das der Benutzer / die Benutzerin verfügt
- Computer verfügt nicht über Hintergrundwissen über die Benutzerin / den Benutzer



Ach so ist das gemeint!  
...hätte sie / er das nicht  
gleich sagen können?!?

ToDo: Wissenslücke  
zwischen Benutzer und  
Computer mindern

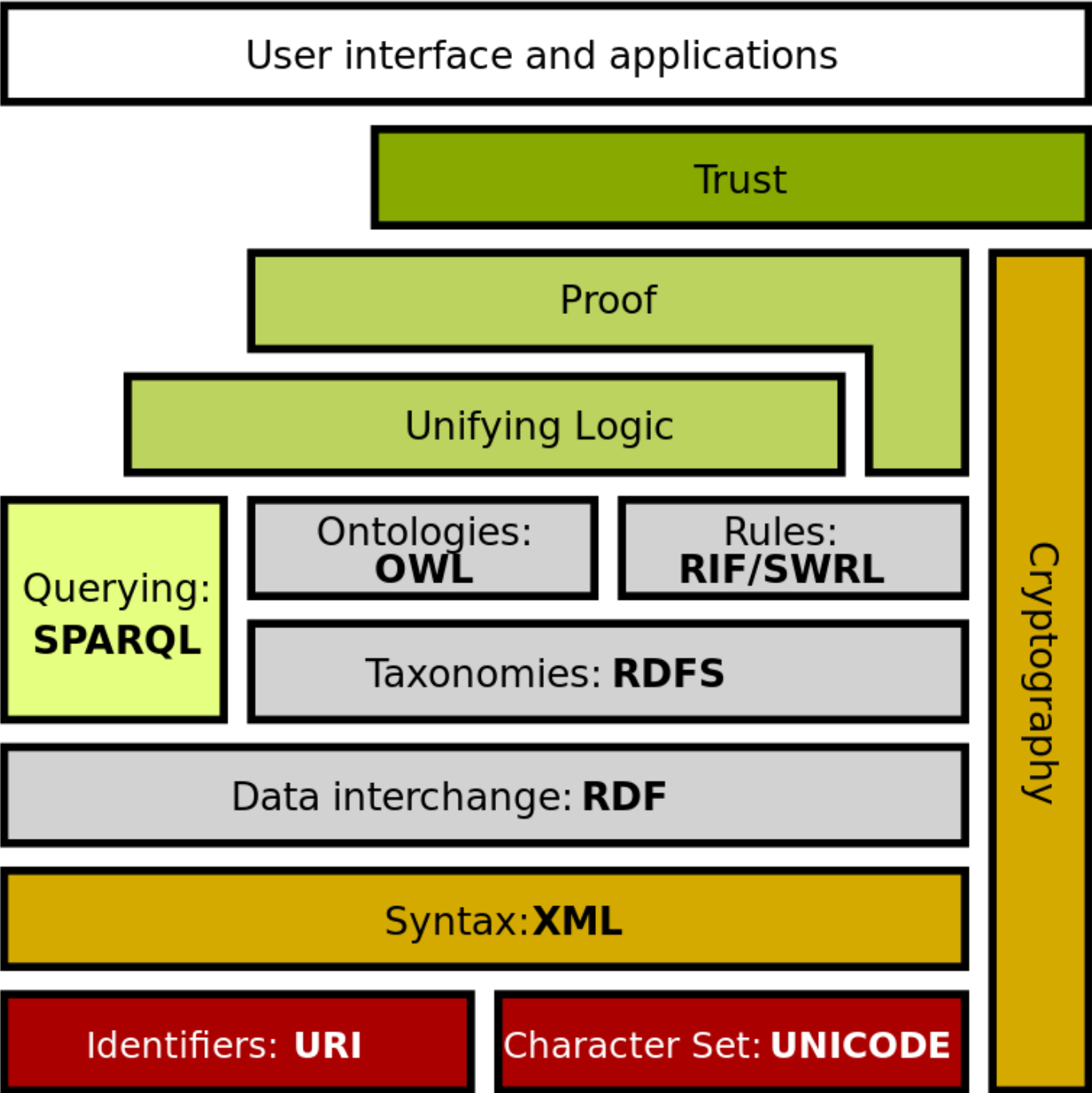
- Bereitstellung von Wissen in einer Art und Weise, in der es von Computern verarbeitet werden und für weiteres Schließen verwendet werden kann
- Z.B.: Bereitstellung von (semantischen) Metainformationen, die die Inhalte der Website beschreiben (description, keywords, etc.)



# Semantic Web

- Konzept des Semantic Web formuliert 1996 von Tim Berners-Lee
- Kerntechnologien (logikbasierte Sprachen zur Representation von Wissen und (automatisiertem) Schließen) entwickelt im Forschungsfeld der Künstlichen Intelligenz.
- Standards: W3C
- Ursprüngliche Intention: Annotation – Anreicherung der Inhalte im WWW durch Metadaten



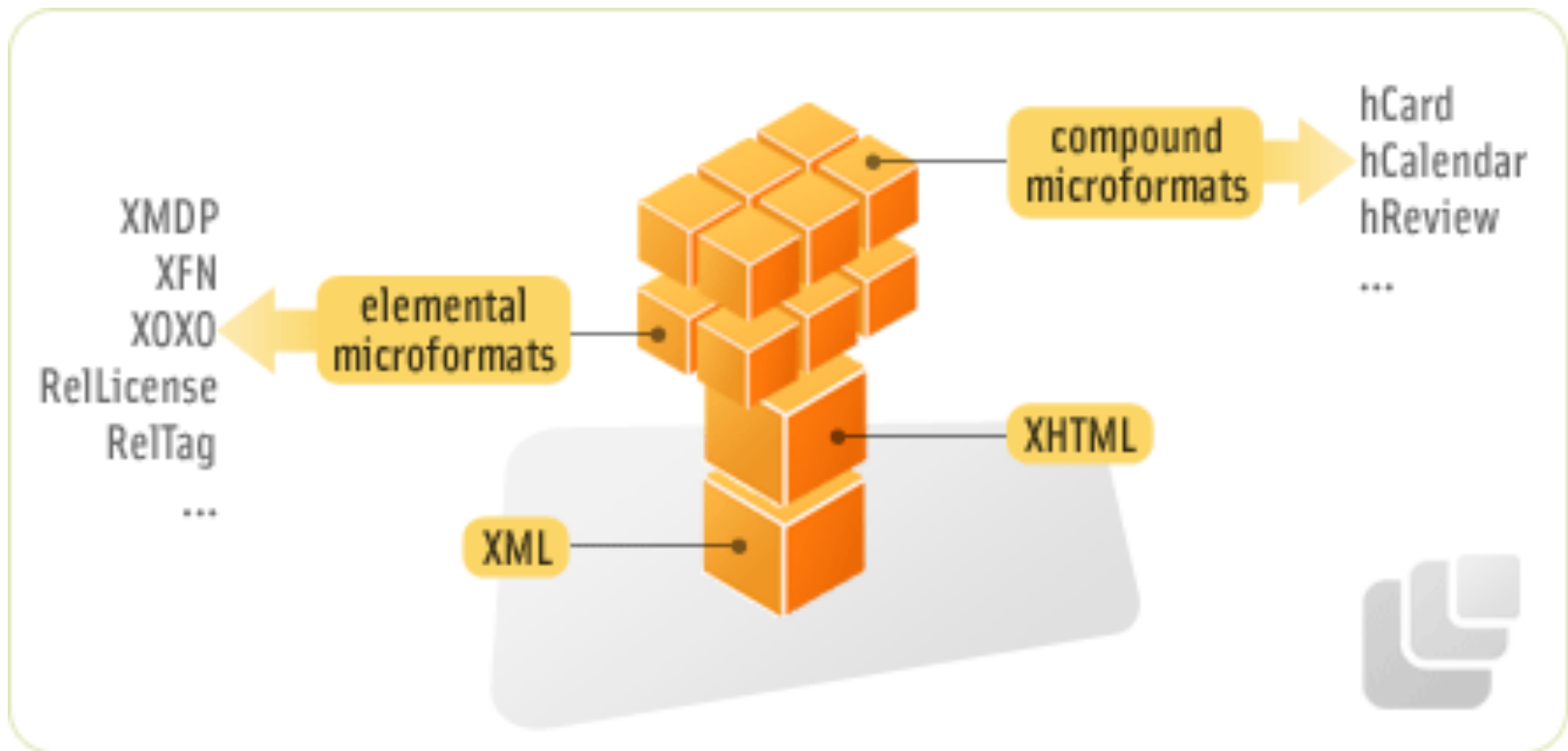




# Mikroformate



„Designed for humans first and machines second, microformats are a set of simple, open data formats built upon existing and widely adopted standards.“  
(<http://microformats.org/>)



# RDF Grundkonzepte

## Natürlichsprachige Aussage:

Die Webseite „<http://www.example.org>“ hat einen Urheber namens Jan Wieners.

Die Aussage besteht aus dem **Gegenstand der Aussage**,

einer **Eigenschaft des Gegenstandes** und einem **Wert für diese Eigenschaft**.

In RDF Terminologie wird der Gegenstand der Aussage als **Subjekt (subject)**, die

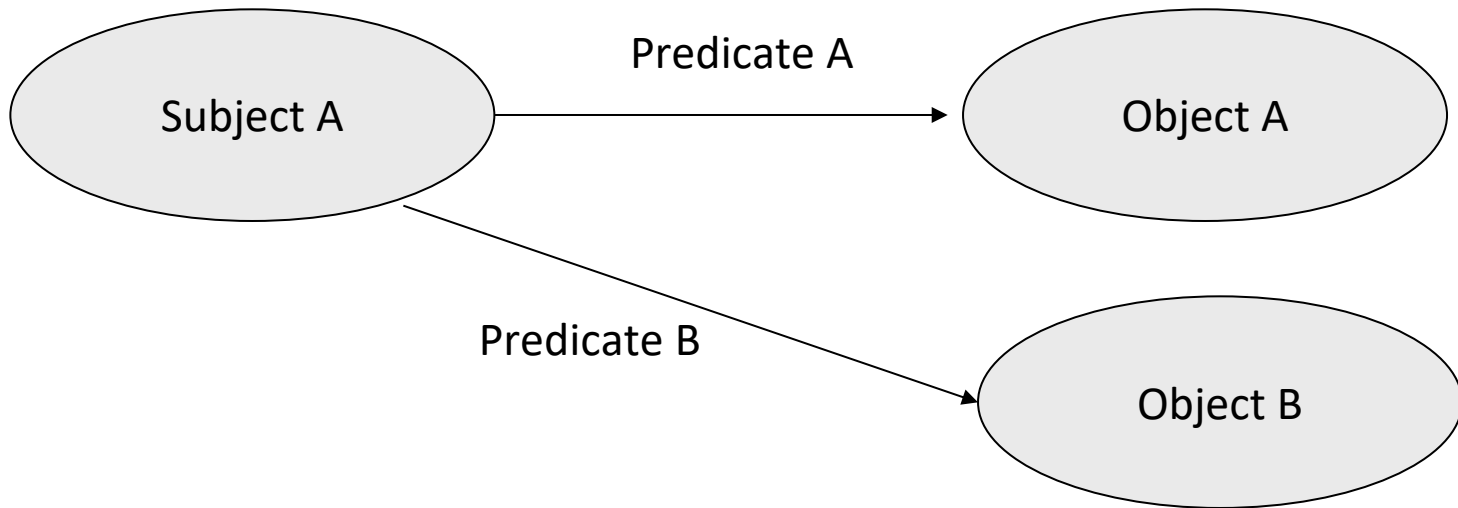
Eigenschaft als **Prädikat (predicate)** und Wert der Eigenschaft als **Objekt (object)**

bezeichnet.

→ Subjekt, Prädikat und Objekt bilden ein **Tripel**.

# RDF (Graphen)modell

Ein oder mehrere Tripel bilden einen **RDF-Graphen**:



# Uniform Resource Identifier (URI)

Intention:

- Globaler, eindeutiger Bezeichner für Entitäten
- Unterklassen: URL, URN

Beispiele:

- [http://de.wikipedia.org/wiki/Uniform\\_Resource\\_Identifier](http://de.wikipedia.org/wiki/Uniform_Resource_Identifier)
- `urn:isbn:4-7980-1224-6`

# RDF

```
<?xml version="1.0"?>
```

Ressource (resource)  
→ Worum geht's?

```
<rdf:RDF
```

```
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:hki="https://dh.phil-fak.uni-koeln.de/">
```

```
<rdf:Description rdf:about="https://lehre.idh.uni-koeln.de/">
```

```
<idh:author>Jan Wieners</idh:author>
```

```
<hki:homepage>https://jan-wieners.de/</hki:homepage>
```

```
</rdf:Description>
```

```
</rdf:RDF>
```

Eigenschaft (property)

Eigenschaftswert (property value)

# Abfrage von Informationen

## SPARQL – SPARQL Protocol And RDF Query Language

Ein Beispiel (vgl. <http://de.wikipedia.org/wiki/SPARQL>):

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?name
       ?email
WHERE
{
  ?person a          foaf:Person .
  ?person foaf:name  ?name .
  ?person foaf:mbox  ?email .
}
```

# Algorithmen der Bildverarbeitung

## Grundbegriffe:

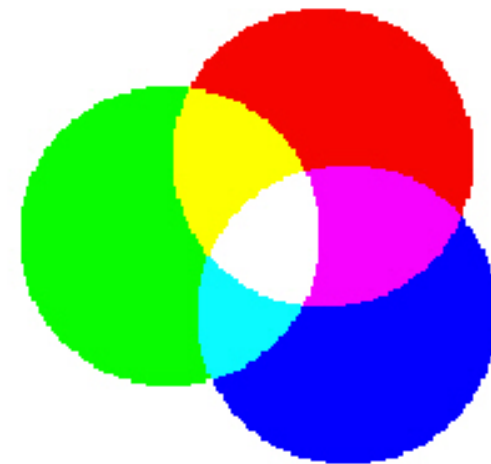
- Farbmischung: Additiv, Subtraktiv
- Raster- vs. Vektorgrafik
- Pixel
- Auflösung
- Farbtiefe

## Kompressionsverfahren

- Nicht verlustbehaftet
  - Run Length Encoding (RLE)
  - Wörterbuch-Algorithmen, z.B. Lempel-Ziv-Welsh (LZW)
  - Huffman-Codierung
- (Verlustbehaftet)

# Additive Farbmischung

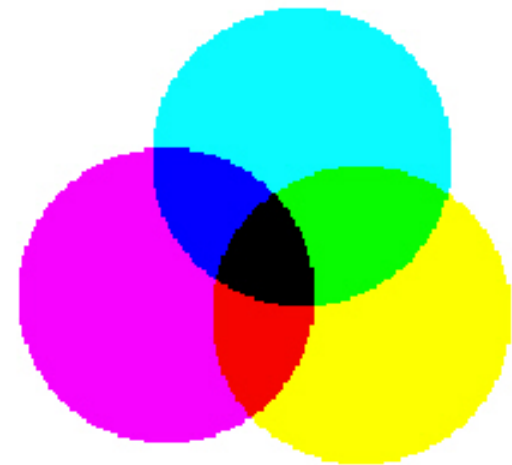
Primärfarbe			Grundfarbe
Blau	Grün	Orangerot	
⊗	–	–	Blau
–	⊗	–	Grün
–	–	⊗	Orangerot
⊗	⊗	–	Cyan
⊗	–	⊗	Magenta
–	⊗	⊗	Gelb
⊗	⊗	⊗	Weiß
–	–	–	Schwarz

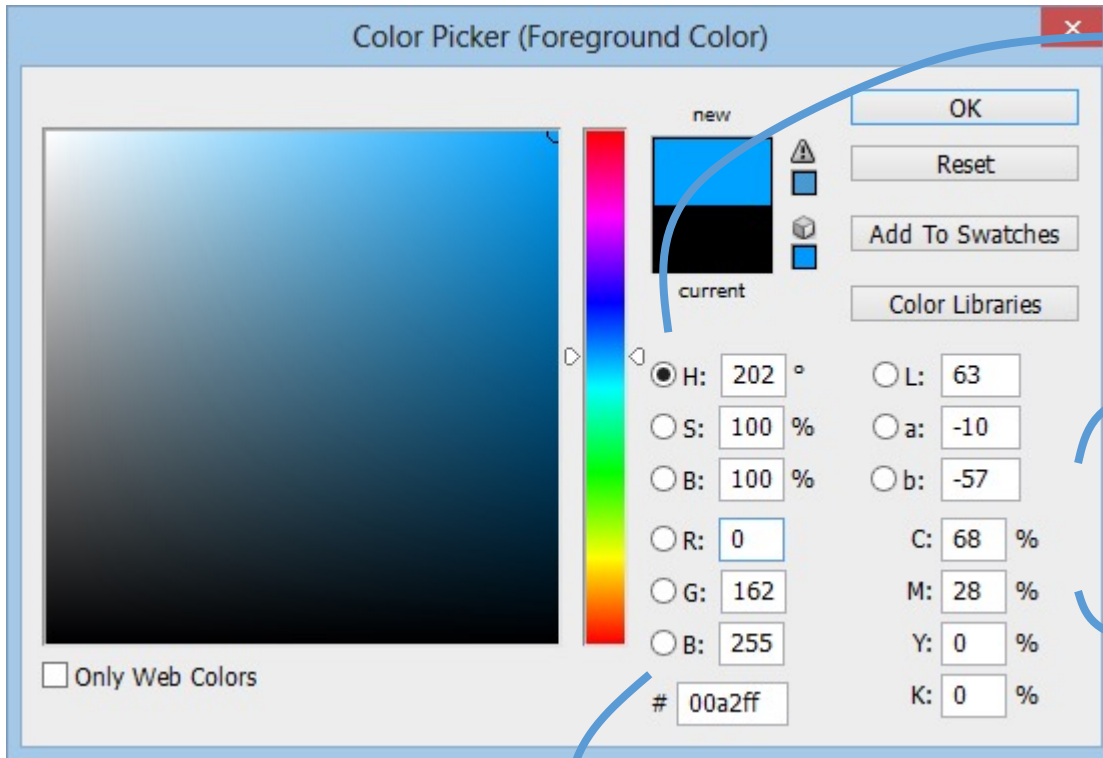




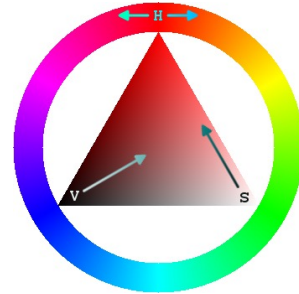
# Subtraktive Farbmischung

Primärfarbe			Grundfarbe
Cyan	Magenta	Gelb	
⊗	—	—	Cyan
—	⊗	—	Magenta
—	—	⊗	Gelb
⊗	⊗	—	Blau
⊗	—	⊗	Grün
—	⊗	⊗	Orangerot
⊗	⊗	⊗	Schwarz
—	—	—	Weiß





**HSV:** Farbwinkel, Sättigung, Hellwert bzw. absolute Helligkeit (B), Brightness



**L\*a\*b\*:** Menschl. Wahrnehmung



**RGB:** Wenn jede der drei Primärfarben mit einer Auflösung von 256 Werten dargestellt werden kann, dann erhalten wir  $256^3 = 16,7$  Mio. verschiedene Farbtöne.

# Raster- vs. Vektorgrafik





Format	Kompression	Unterstützte Bittiefen	Vor- und Nachteile	Größe eines 8-Bit Graustufenbildes
BMP	RLE	1, 4, 8 und 24	+ leichte Erzeugung - Größe	256 KB ( <u>lossless</u> )
GIF (Graphics Interchange Format)	LZW	1 bis 8 (aus einer 24 Bit Tabelle)	+ Steuerblock + Schachtelung + Interlaced / Transparent - Lizenzpflicht seit 1994	240 KB (lossless)
PNG (Portable Network Graphics Format)	LZW-Variante	8, 16, 24, 48	+ Lizenzfrei + auch 24 Bit + Interlaced / Transparent	219 KB (lossless)
JPEG	DCT und Huffman	24	+ für Fotografien + Größe + Progressive Format - schlecht für "digitale" Bilder	21 KB ( <u>lossy</u> )
TIF (Tagged Image Format)	LZW-Variante	1 bis 48	+ Plattformunabhängig + viele Metainformationen ablegbar - viele Unterformate	264 KB (lossless)

# Ansätze zur Datenkompression

- RLE – Run Length Encoding (Lauf­längen­kodierung)
- Wörterbuchbasierte Kompressionsmethoden
  - Lempel-Ziv-Welch (LZW)
- Statistische Kompressionsmethoden
  - Huffman-Algorithmus

# Optische Zeichenerkennung (OCR)

- Eine Herausforderung um OCR und Merkmalsextraktion
- Vorverarbeitung: Verbesserung des Quellmaterials
  - Punktoperationen
    - Umwandlung RGB → Graustufenbild
    - Histogramm
    - Binarisierung
  - Filter
    - Mittelwertfilter
    - Medianfilter

homo. **E**rite filii audite  
morem dñi docebo vos.  
homo qui vult vitam: di  
hies videre bonos. **R**olibe  
in tuā a malo. **&** labia tua  
quantum dolum. **R**uerite  
o **&** fac bonum: inquire pa  
psequere eum. **E**culi dñi  
stos. **&** aures eius in preces  
n. **V**ultus autem sup facien  
ala: ut pda de terra meo  
eorum. **L**anauerunt iusti  
s exaudiuit eos. **&** ex omni  
bulacionibus eoz libera  
eos. **I**urta est dñs his qui  
mlata sunt corda: **&** hanc  
bū saluabit. **A**ulte tribu  
nes eorum: **&** de his omnib  
caute eos dñs. **C**ustodit  
omnia ossa eoz: **n**unquam ex  
non conterent. **N**on peccato  
a pessima: **&** qui oderunt ius  
a delinquent. **E**demer dñs

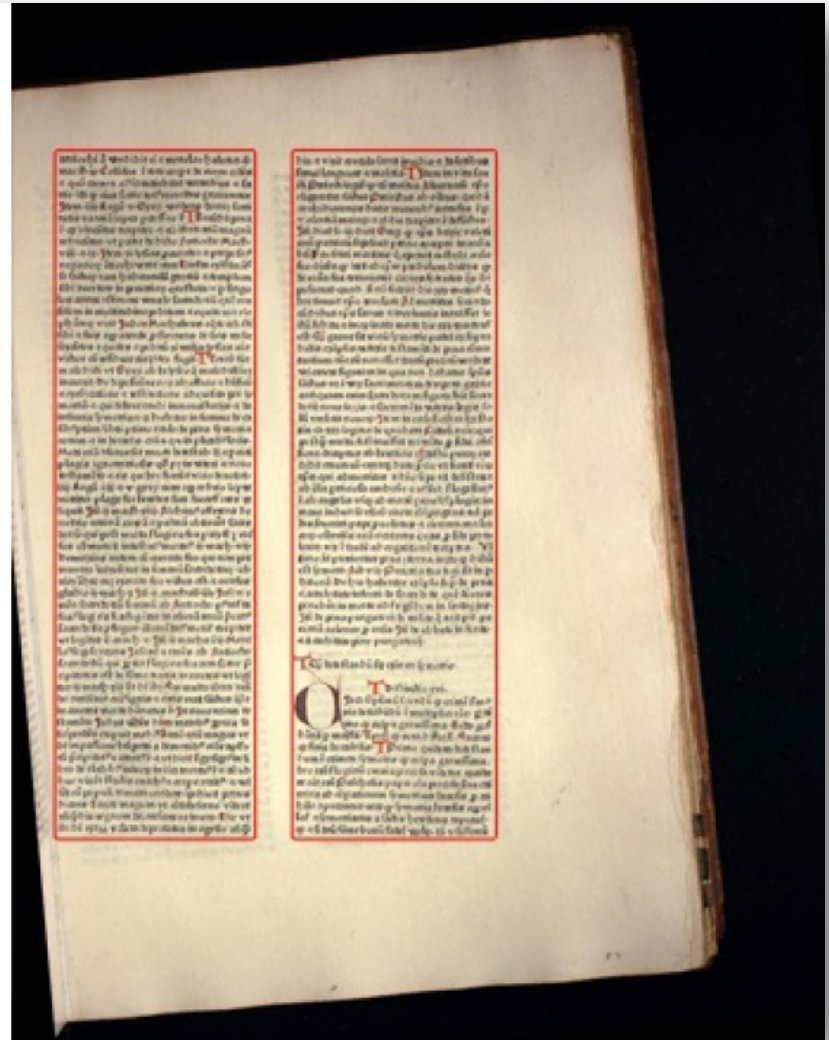
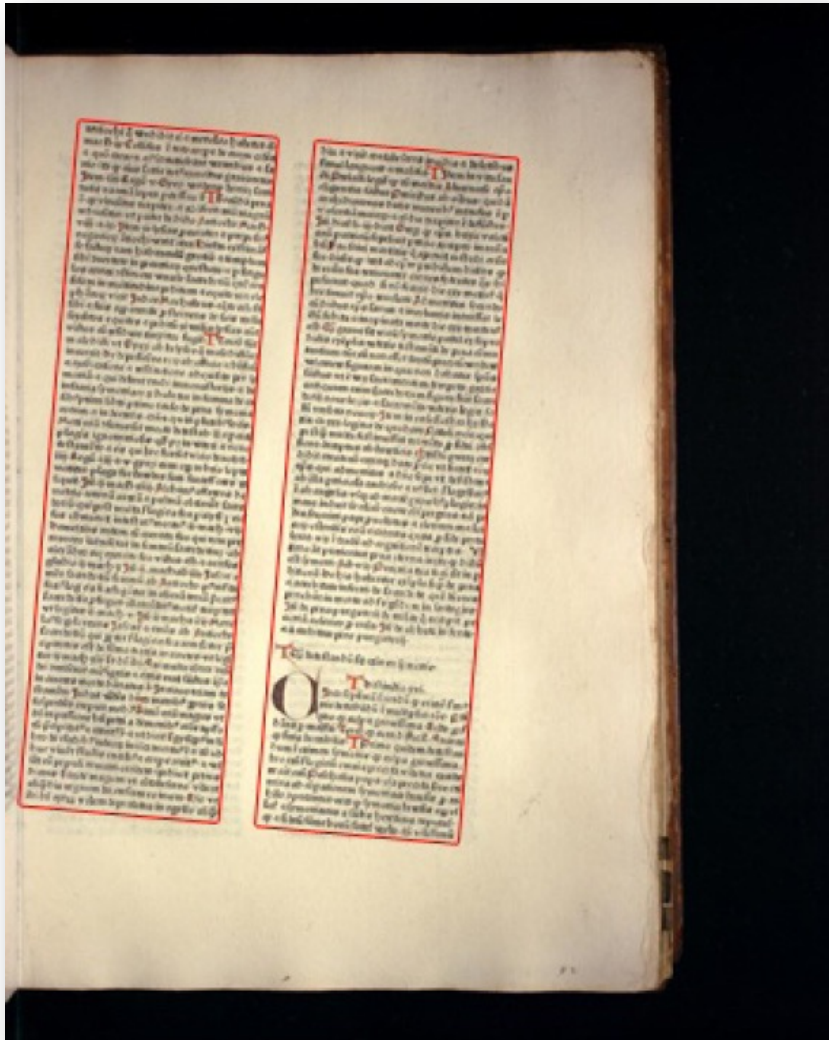
99. 3. 35. \*

19. 7. 1 gram.

mensurabilem posuisti dies meos.  
substantia mea tanquam nichil  
luna ante te. **V**erum tamen  
uersa uanitas: omnis homo u  
uens. **E**ritum tamen in u  
ne pte: **n** homo: sed **&** frustis  
conturbatur. **H**esauriat **&**  
nonat: cui congregabit ea. **T**  
que est expectatio mea non me  
**&** substantia mea apud te est. **E**  
omnib: iniquis: **h**is non  
mei obprobrium in seipsum de  
mei. **I**urta est non querit  
meum: **o**mni tunc est a me  
plagas eoz. **F**erunt dñs  
tue **&** id est: **n** increpare  
pore iniquitatem: **n** corru  
herum est: **t**abescit fact  
licet dñs in manu eius. **T**  
cū tamen uane uoluntate  
nā homo. **A**ul: nationes  
nā dñs: **d**ignationem  
an dñs: **q**ue non as



# Deskew / „Geraderücken“



**D**e voluntate dei . xxxij

**D**e iusticia dei . xxxij

**D**e misericordia dei . xxx

D e v o l u n t a t e d e i . x x x j

D e i u s t i c i a d e i . x x x j

D e m i s e r i c o r d i a d e i . x x x

# Maschinelles Sehen

Von der Rastergrafik zur Merkmalsextraktion –  
Algorithmen & Co.

- Vorbereitung bzw. Vorverarbeitung, u.a.:
  - Punkt-Operationen
  - Filter
  - Geometrische Operationen
    - Drehung, Streckung, Verschiebung des Bildes
- (Objektidentifikation)
- (Merkmalsextraktion)
- (Clustering, z.B. Kohonen-Karte /  
Selbstorganisierende Karte)

## Das IHS-Farbsystem (auch HSI, HSV Farbsystem):

- Leuchtstärke (Intensity): Maßeinheit der Helligkeit, resultiert aus dem Durchschnitt der Farbwerte

$$I = \frac{R + G + B}{3}$$

- Sättigung (Saturation): Beschreibt die Farbreinheit

$$S = 1 - \frac{\min(R, G, B)}{I}$$

- Farbton (Hue): Proportional zur durchschnittlichen Wellenlänge der Farbe; basiert auf Abbildung der Farben in einem Polarkoordinatensystem:

$$H = \cos^{-1}\left(\frac{(R - G) + (R - B)}{2} \cdot \frac{1}{((R - G)^2 + (R - B)(G - B))^{1/2}}\right)$$

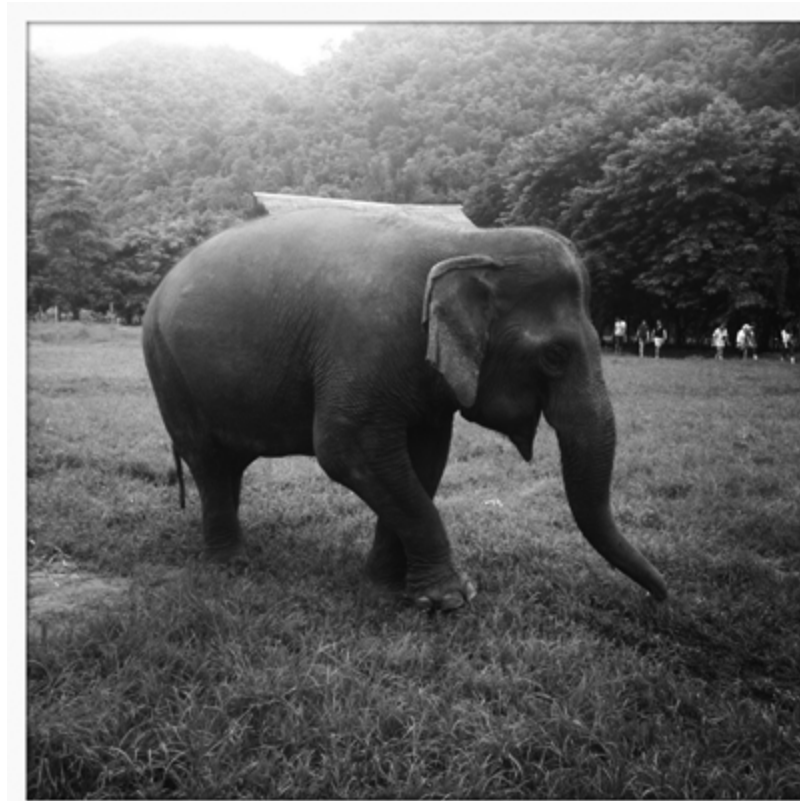
## Relevant für Umwandlung in Graustufenbild: Leuchtstärke (Intensity)

### Algorithmus:

- Betrachte jedes Pixel  $P$  der Rastergrafik
- Weise jedem Farbkanal des betrachteten Pixels  $P$  die Summe der Werte der einzelnen Farbkanäle R, G und B zu und dividiere anschließend durch die Anzahl der drei:

$$P = rgb\left(\frac{P_R + P_G + P_B}{3}, \frac{P_R + P_G + P_B}{3}, \frac{P_R + P_G + P_B}{3}\right)$$

# Binarisierung



Grundfrage dieses (und der vergangenen Verfahren): Wie lässt sich Bildinformation verwerfen, ohne die Bildcharakteristika (i.e. die Form des Elefanten) zu zerstören?

# Binarisierung



Grundfrage dieses (und der vergangenen Verfahren): Wie lässt sich Bildinformation verwerfen, ohne die Bildcharakteristika (i.e. die Form des Elefanten) zu zerstören?

# Filter

- **Lineare Filter (LSI-Filter, linear shift-invariant filters):**  
Jedes Pixel im Verarbeitungsfenster wird mit einem vordefinierten Wert aus einer Faltungs- bzw. Filtermatrix multipliziert.
- **Nichtlineare Filter: U.a. heuristische Ansätze**  
Heuristik (vgl. Prechtel / Burkard): „Lehre bzw. Theorie der Verfahren zum Finden von Neuem und Problemlösen.“



# KI: eine lange Geschichte

- Interesse an Wellen
- Angst vor dem Rechner – und Liebe
  - Roboter
  - der militärisch-industrielle Komplex
- Overselling
  - Der Verkauf von Produkten durch aufgeblähte und übertriebene Behauptungen

# GPS: General Problem Solver

- Newell und Simon 1957
- Ein Programm für die Lösung aller Probleme
  - sie müssen jedoch erst formuliert werden
- Wichtig für die Entwicklung der Programmierung
  - Separation zwischen den Regeln und der allgemeinen Problemlösungsmaschine
- jedoch nicht besonders intelligent

# Eliza

- Joseph Weizenbaum, 1966
- MAD-Slip auf einem IBM 7094
- Gespräch mit einer Maschine
- Mustererkennung
- Regelbasierte Transformationen

# Expertensysteme

- Mycin 1974
  - medizinische Diagnose
- Protokollanalyse
  - ≈ User Stories
- Regelbasiert
- In Lisp-Programm definiert

Rule 52:

If

- 1) THE SITE OF THE CULTURE IS BLOOD
- 2) THE GRAM OF THE ORGANISM IS NEG
- 3) THE MORPHOLOGY OF THE ORGANISM IS ROD
- 4) THE BURN OF THE PATIENT IS SERIOUS

Then there is weakly suggestive evidence (0.4) that

- 1) THE IDENTITY OF THE ORGANISM IS PSEUDOMONAS

```
(defrule 52
```

```
  if (site culture is blood)
```

```
    (gram organism is neg)
```

```
    (morphology organism is rod)
```

```
    (burn patient is serious)
```

```
  then .4
```

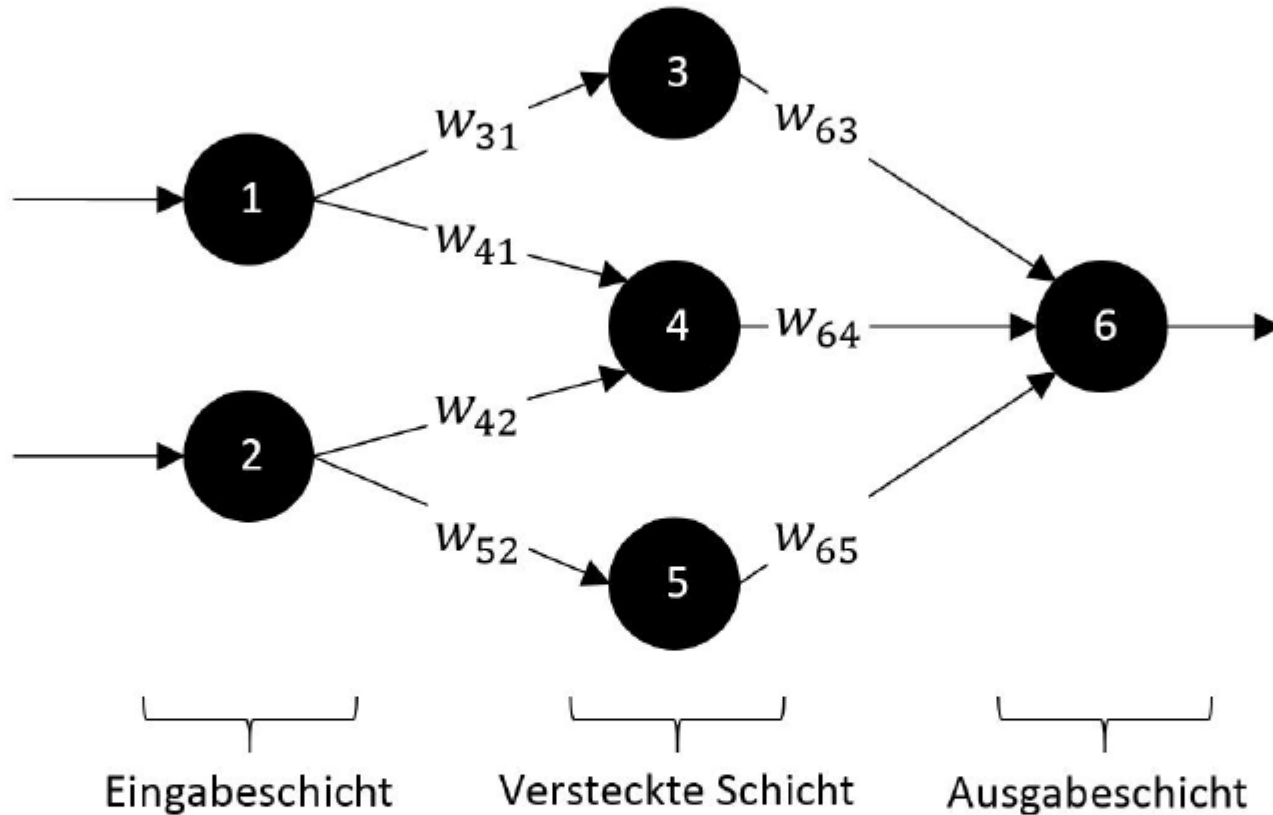
```
    (identity organism is pseudomonas))
```

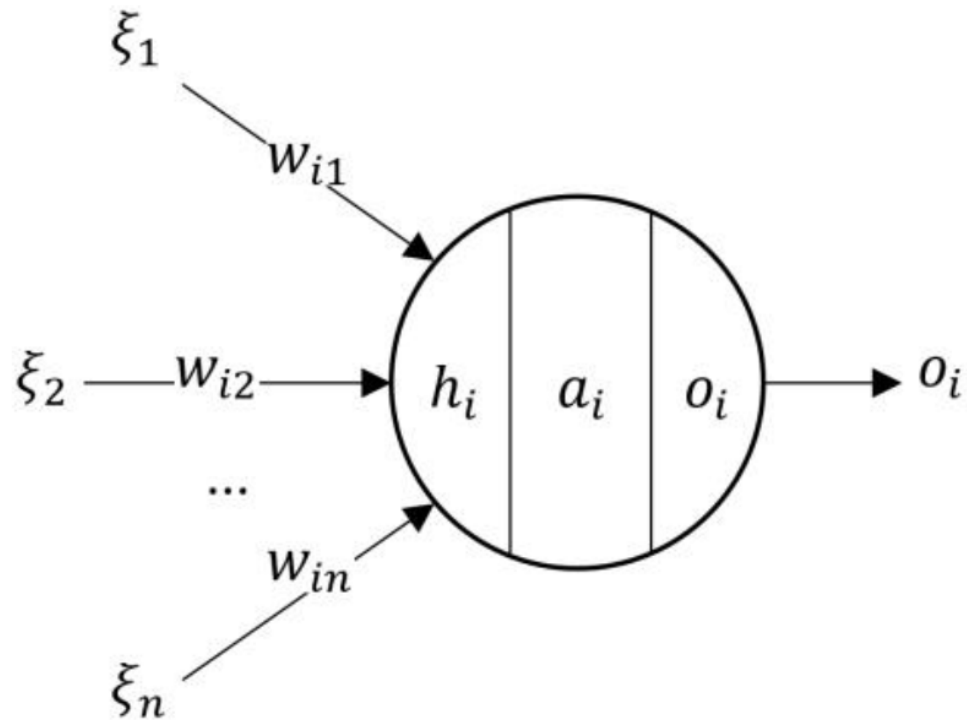
# Sprachverarbeitung

(Natural language processing, NLP)

- Künstliche bzw. natürliche Sprache
  - Ambivalenz
- Sprache für Parsen
- Maschinenübersetzung
- Phrasenstrukturgrammatiken
- Maschinelles Lernen
- Statistikbasiert

# Ein künstliches neuronales Netz





Einfaches Modell eines künstlichen Neurons  $i$  mit:

- $h_i$ : Summe der gewichteten Eingabeimpulse  $w_{i1} \dots w_{in}$ , die das Neuron von verknüpften externen Einheiten oder von Sensoren ( $\xi_1 \dots \xi_n$ ) erhält
- $a_i$ : Aktivierungsmaß
- $g(h_i)$ : Aktivierungsfunktion
- $o_i$ : Ausgabe des Neurons

## Endlicher Markov-Entscheidungsprozess

→ Verlauf von Entscheidungen, bei dem die Belohnung des Agenten einzig von dem Zustand der Umwelt und der Aktion des Agenten abhängt. Definiert als Tupel  $(S, A, P, R, \gamma)$ :

- $S = \{s^1, s^2, \dots, s^n\}$ : Zustandsraum, d.h. die möglichen Zustände der Umwelt;  $s_t \in S$  repräsentiert den Zustand der Agentenumwelt zum Zeitpunkt  $t$ .
- $A$ : Menge der Entscheidungen bzw. Zugmöglichkeiten, die dem Agenten zu jedem Umweltzustand  $A(s)$  zur Verfügung stehen.





## Endlicher Markov-Entscheidungsprozess

$(S, A, P, R, \gamma)$ :

- $a_t \in A(s_t)$ : von dem Agenten ausgeführte Aktion zum Zeitpunkt  $t$ .
- Übergangsfunktion  $P(s, a, s')$ : führt die Umwelt in ihrem Zustand  $s$  aufgrund der Handlung  $a$  des Agenten zu einem neuen möglichen Nachfolgezustand  $s'$ .
- $R(s, a, s')$ : Belohnungsfunktion, Feedback, das der Agent aufgrund seiner Handlung erhält.
- Diskontierungsfaktor  $\gamma$  mit  $0 \leq \gamma \leq 1$  lässt sich dazu verwenden, um das Lernverhalten des Agenten feinzustimmen. Dient dazu, künftige Belohnungen gegenüber zeitnahen Belohnungen abzuschwächen.



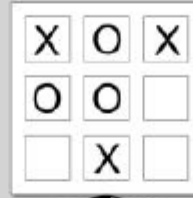
# Spiele

- Spielen als suchen
- Vom Sieger zum Sieger
  - Othello (1980er)
  - Schack (1997)
  - Go (2016)
- Können Rechner alle Spiele spielen?

MAX (X)

$$\max(0, -1, -1) = 0$$

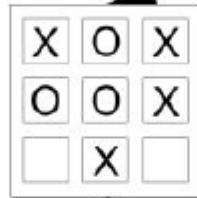
$x_0$



MIN (O)

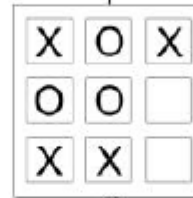
$$\min(+1, 0) = 0$$

$x_1$



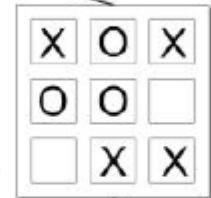
$$\min(-1, 0) = -1$$

$x_2$



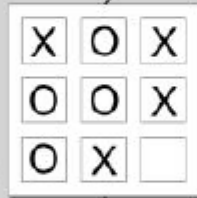
$$\min(-1, +1) = -1$$

$x_3$



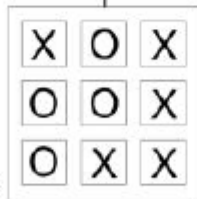
MAX (X)

$x_4$



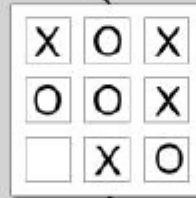
+1

$x_{10}$



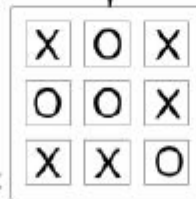
+1

$x_5$



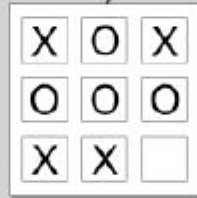
0

$x_{11}$



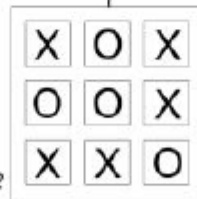
0

$x_6$



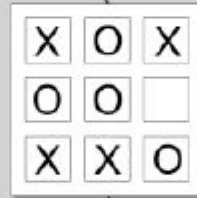
-1

$x_{12}$



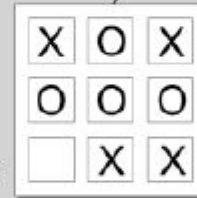
0

$x_7$



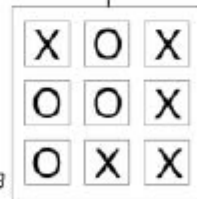
0

$x_8$



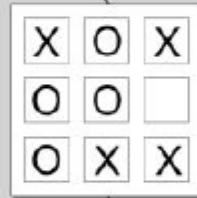
-1

$x_{13}$



+1

$x_9$



+1

# Ethik

- Angst vor dem Rechner
- Überwachung
- Bots und Destabilisierung
- Filterblasen
- Lernbasis
- Wer macht Entscheidungen?
- Was sind die Grundlagen für Entscheidungen?
- Autonomie
  - Autos
- Verantwortung
- Relevant für uns? Für Sie?