

Sprachressourcen teilen – Rechtliche Fragen

Autor: Christian Galinski (Infoterm)

Wien, 25. September 2018

Teilen – verteilen – übernehmen – weaternutzen



- Sprachressourcen? → Mikrocontent
 - (semantisch) strukturierter Content – auch nicht-sprachlicher Art
 - Texte, die solche Contenteinheiten enthalten
- Datenaustausch zwischen Personen/über Plattformen → neue LTT / IKT:
 - Maschinelle Übersetzung (MÜ)
 - Übersetzungsspeicher (TM)
 - Dialogsysteme
 - NEU: Produkthaftung, DSGVO, Barrierefreier Content, ...
- Welche Rechtsprobleme & Risiken?
 - UhG und benachbarte / verwandte Gesetze
- Lösungsansätze: → DRM in Verbindung mit Normung

Was ist Content?



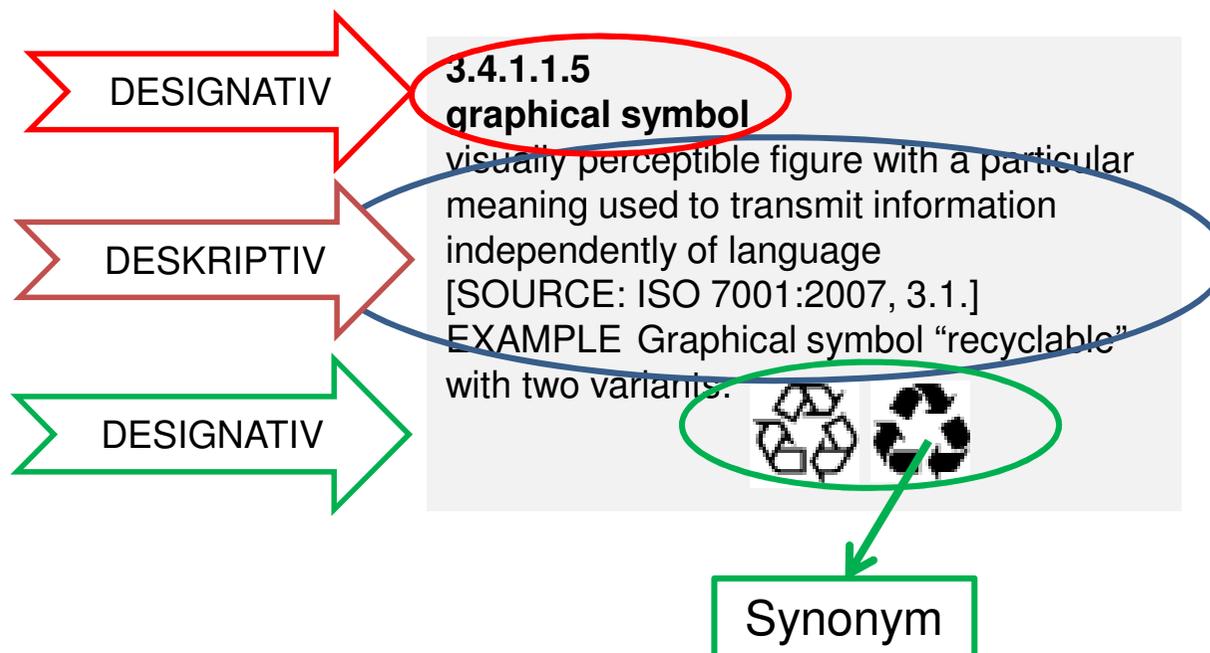
- **Potenziell: jede Art von Content mehrsprachig und **multimedial**:**
 - Unstrukturierter Content: Texte, Musik, Videos, ...
 - Strukturierter Content = **Mikrocontent**: terminologische und lexikographische Daten, Fotos, Videoclips, Blogbeiträge, visuelle nichtsprachliche Darstellungen, Audiodarstellungen, ...
→ kann alles heute automatische auf/abgerufen werden
 - Übersetzungsspeicher? (TM – translation memory)
- **IKT: ständig neue Tools, Hybride, Kombinationen + Kooperation**
 - Alles kann „gelesen“/ausgewertet werden: PDF, Bilder, usw.
- **UHG-Schutz unterschiedlich für verschiedene Contentarten**
→ höherer Granularitätsgrad der Ressourcen-Beschreibung erforderlich
- ➔ **Kennzeichnung von „geschützten“ digitalen Objekten**
→ in TDB, Texten, eigentlich jeder Art von Content



- nationale Gesetzgebungen weitgehend angeglichen,
- **ABER:** Rechtsprechung nicht einheitlich
- **Datenschutzgrundverordnung (DSGVO):**
 - wenn MÜ z.B. auch personenrelevante Daten auswertet
 - automatische Auswertung: Daten wandern in TM
- **UN-BRK → EU-Direktiven → nationales Recht:**
 - Barrierefreiheit aller Systeme der öffentlichen Verwaltung
 - Barrierefreiheit aller mobilen Zugänge
- **Schutzrechte für geistiges Eigentum: neue EU-Direktive**
 - **Kennzeichnung, Filter, usw.**

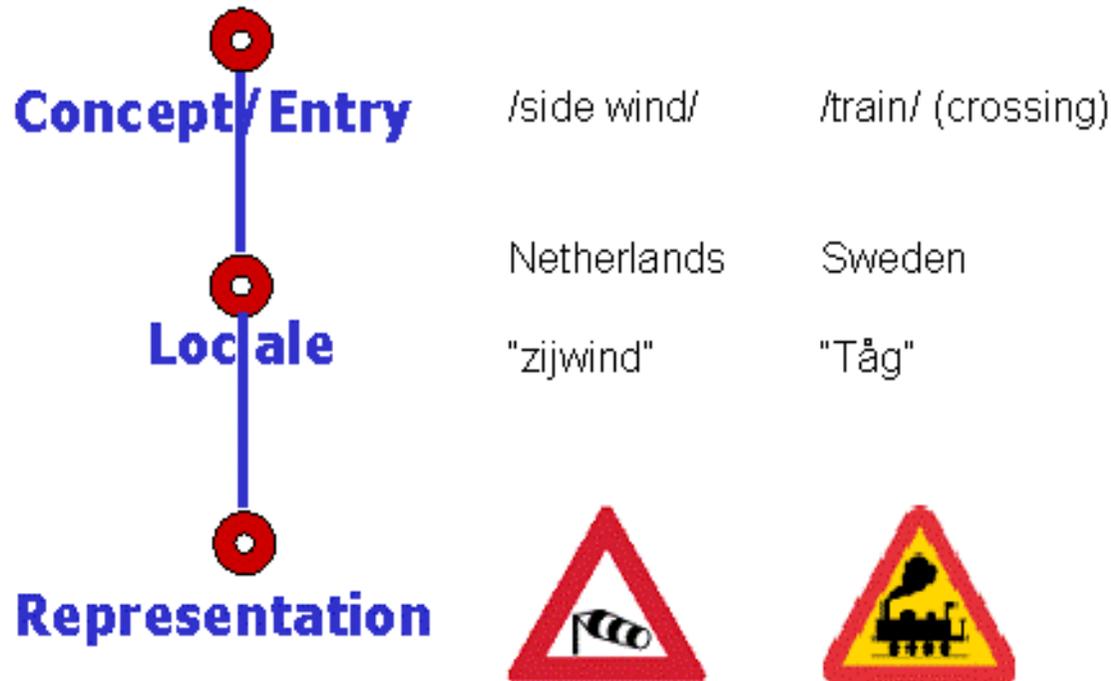
Entsprechend terminologischem Datenmodell

[aus: ISO 10241-1:2011]



Eintrags-Identifikation: ISO 10241-1:2011, 3.4.1.1.5 – fehlt (=implizit) © ISO + Nutzungsrechte

Locale / Varianten



- „Morpheme“
- „Synonyme“
- „Homonyme“
- „Homographen“
- „Abkürzungen“
- „Locale“
- „Phraseologie“
- ...

Ähnliche Phänomene wie in der Sprache

Helicopter-Einweisungssignale



DESKRIPTIV



DESIGNATIV

LIFT OFF

LAND

MOVE UPWARD



MOVE REARWARD

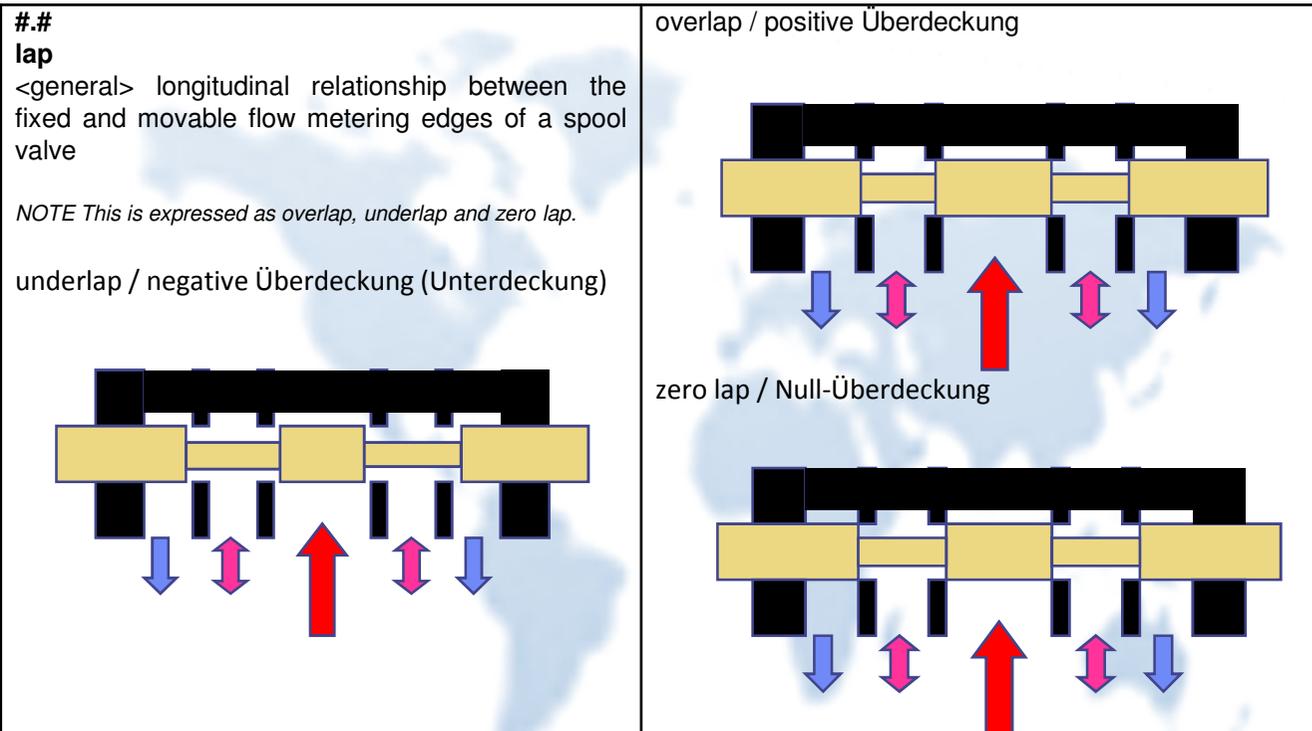
HOLD-HOVER

RELEASE SLING LOAD

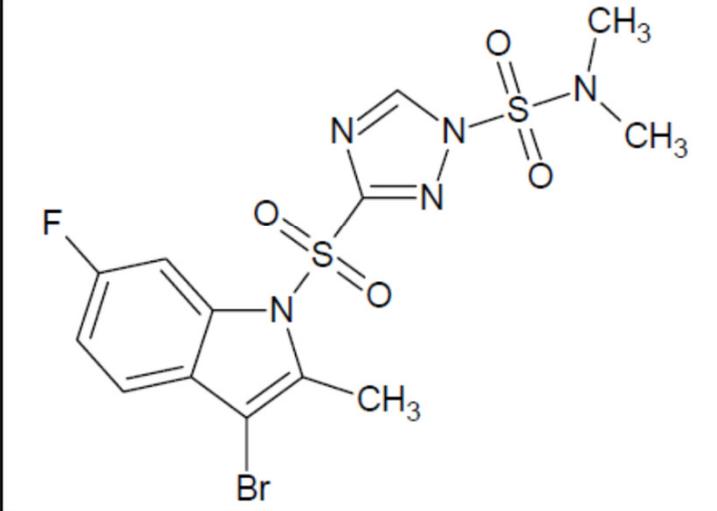
https://en.wikipedia.org/wiki/Aircraft_marshalling – NOT from ICAO source!

Beispiel für non-verbale deskriptive / definierende Repräsentation

*Info
term*



Chemie: >60 mio Substanzen

E: Common name F: Nom commun	Chemical name Nom chimique E: IUPAC F: UICPA C: CAS	Structure Structure		Use Application
		Molecular formula Formule brute	CAS Registry Number® Numéro d'enregistrement CAS®	
IUPAC International Chemical Identifier (InChI™)				
E amisulbrom F amisulbrom, m	<p>3-(3-bromo-6-fluoro-2-methylindol-1-ylsulfonyl)-<i>N,N</i>-dimethyl-1<i>H</i>-1,2,4-triazole-1-sulfonamide</p> <p>3-(3-bromo-6-fluoro-2-méthylindol-1-ylsulfonyl)-<i>N,N</i>-diméthyl-1<i>H</i>-1,2,4-triazole-1-sulfonamide</p> <p>3-[(3-bromo-6-fluoro-2-methyl-1<i>H</i>-indol-1-yl)sulfonyl]-<i>N,N</i>-dimethyl-1<i>H</i>-1,2,4-triazole-1-sulfonamide</p>			F
		C ₁₃ H ₁₃ BrFN ₅ O ₄ S ₂	348635-87-0	
InChI=1/C13H13BrFN5O4S2/c1-8-12(14)10-5-4-9(15)6-11(10)20(8)25(21,22)13-16-7-19(17-13)26(23,24)18(2)3/h4-7H,1-3H3				

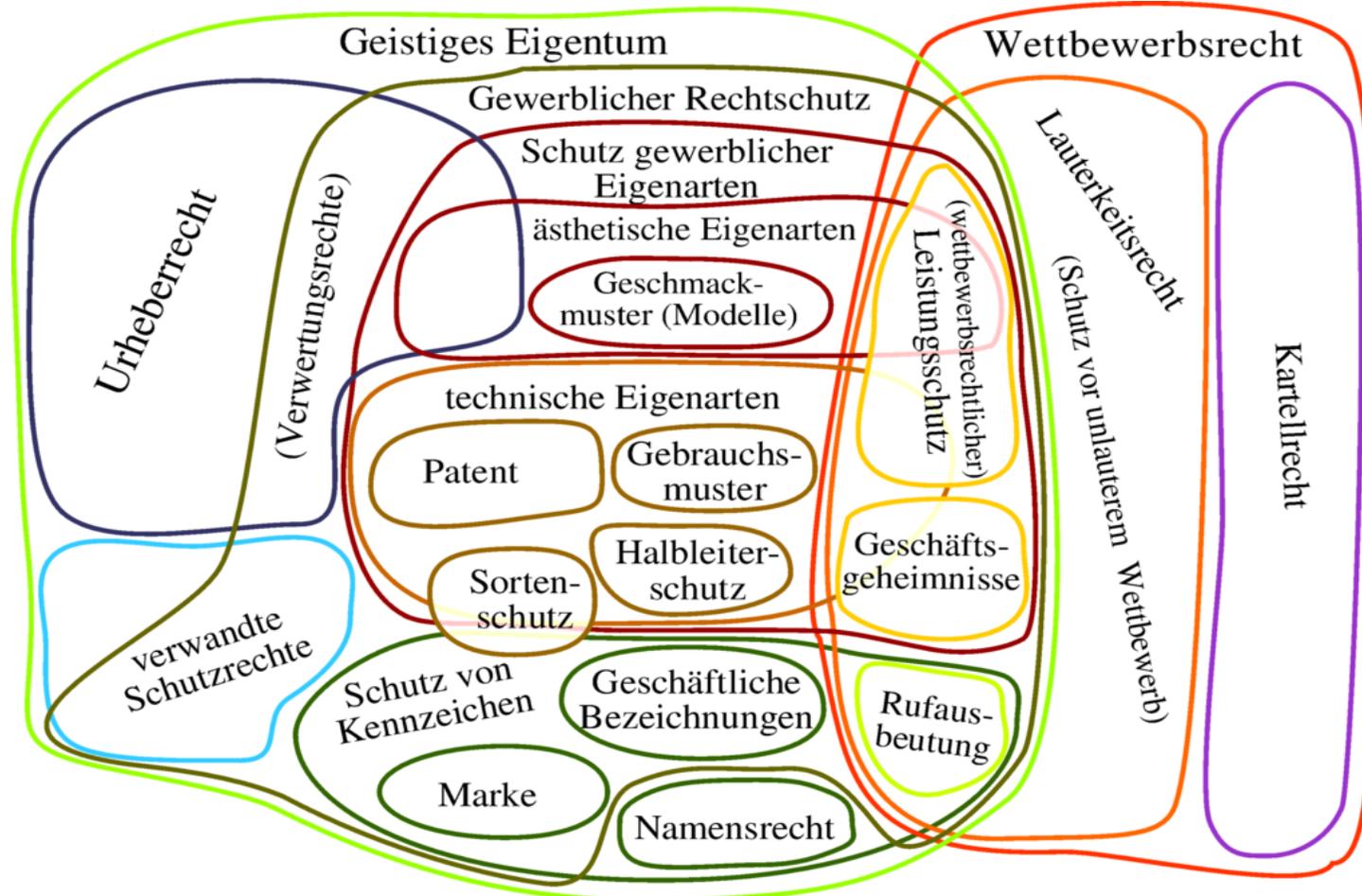
Unicode & Fonts

hand-single-finger

Nº	Code	Browser	Appl	Goog	Twtr	One	FB	Sams	Wind	GMail	SB	DCM	KDDI	CLDR Short Name
158	U+1F448											—		backhand index pointing left
159	U+1F449											—		backhand index pointing right
160	U+1F446											—		backhand index pointing up
161	U+1F595									—	—	—	—	middle finger
162	U+1F447											—		backhand index pointing down
163	U+261D											—		index pointing up

1F3CA ; Emoji_Modifier_Base # 6.0 [1] (🏊) person swimming
 1F442..1F443 ; Emoji_Modifier_Base # 6.0 [2] (👂..👃) ear..nose
 1F446 ; Emoji_Modifier_Base # 6.0 [11] (👉) **backhand index pointing up**
„Fonts“ wie bei Alphabeten

Verwandte und benachbarte Rechte



http://de.wikipedia.org/wiki/Geistiges_Eigentum - **Geistiges Eigentum und Wettbewerbsrecht**

Geistiges Eigentum und Wettbewerbsrecht



I, the copyright holder of this work, hereby publish it under the following licenses

Permission is granted to copy, distribute and/or modify this document under the terms of the [GNU Free Documentation License](#), Version 1.2 or any later version published by the [Free Software Foundation](#); with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled [GNU Free Documentation License](#).

This library is [free software](#); you can redistribute it and/or modify it under the terms of the [GNU Lesser General Public License](#) as published by the [Free Software Foundation](#); either version 2.1 of the License, or (at your option) any later version. This library is distributed in the hope that it will be useful, but **without any warranty**; without even the implied warranty of **merchantability** or **fitness for a particular purpose**. See [version 2.1](#) and [version 3 of the GNU Lesser General Public License](#) for more details.

Geistiges Eigentum und Wettbewerbsrecht.svg

http://de.wikipedia.org/wiki/Geistiges_Eigentum

Verwandte Schutzrechte

- Schutz wissenschaftlicher Ausgaben
- Schutz nachgelassener Werke
- **Schutz von Übersetzungen**
- Schutz der Lichtbilder
- Schutz des ausübenden Künstlers
- Schutz des Veranstalters, Tonträgerherstellers, Sendeunternehmens, Presseverlegers etc.
- **Sonderregelungen für Computerprogramme**
- **Schutz des Datenbankherstellers**

Benachbarte Schutzrechte

Gewerblicher Rechtsschutz:

- Patentrecht
- Gebrauchsmusterrecht
- Geschmacksmusterrecht
- **Markenrecht**

Markenrecht

- Zweck: Wettbewerbsschutz, Vermarktung
- Schutzgegenstand: Zeichen
 - Wörter einschl. Namen
 - Buchstaben, Zahlen
 - Abbildungen
 - Hörzeichen
 - dreidimensionale Gestaltungen
 - Farben
- Voraussetzung: Unterscheidungskraft

Verwertungsrechte

In Österreich gibt es 5 Rechte:

- Vervielfältigungsrecht
 - Verbreitungsrecht
 - (drahtlose und drahtgebundene) Sendung
 - Aufführungsrecht, Vortragsrecht, Ausstellungs- und Vorführrecht
 - öffentliche Zugänglichmachung
- nur mit Einwilligung des Urhebers**

Rechtsfälle /-beispiele

1. Europeana: früher „naiv“, heute 14 Kennzeichnungen
<https://pro.europeana.eu/page/available-rights-statements>
2. TU Köln: 1 Foto in „Webterm“ → alle Bilder unterdrückt
3. Dissertantin (UW): Neuentwicklung eines Diagramms
4. Dialoge, zB für Behindertenkommunikation (fiktiv)
5. IKT-Tools zum automat. Erstellen von Wikipedia-Stubs
6. Markenstreit: **CATS®** vs. **CATS™**
7. CIT: Content Identification Techn. → weniger Ausnahm.
8. Facebook „share“ → 1000 EUR Strafe wegen Foto

Software zum Aufspüren von RS-Verletzungen



- **Welche Barrieren?**
 - **Klassifikationen** (verschiedene Klassifikationssysteme / Klassifikationssysteme unter Copyright, u.a.)
 - **Datenmodelle** → Normung (TMF, LMF, TBX, usw.)
- **Copyright** (weniger bekannt)
 - vor allem: nichtsprachliche Darstellungen
- **~1996 Datenbankschutz und -herstellerschutz**
 - Schutzrechte für Datenbanken oder substantielle Teile hiervon
- **Risiken für Datenabgebende, Datenempfänger, hw/sw-Eigentümer/-Dienstleister, andere Dienstleister (ISPs & LSPs)**

Was ist geschützt?

Gute Ideen, Einfälle, Lehren, Methoden, Systeme, Informationen oder Wissen, viel Arbeit? → **NEIN**
→ ein **“Werk”** – abhängig von **“Schöpfungshöhe”**
(*Kreativität vs. “sweat of the brow”?*)

Ein **“Werk”** ist die **wahrnehmbare Formgestaltung** einer oder mehrerer Ideen, von Wissen oder Traditionen...

→ **“Werke”** in Literatur, Wissenschaft und Kunst:
ursprünglich vor allem Texte
(ab einer bestimmten **“Schöpfungshöhe”**)

→ heute gibt es viele und ständig neue **“wahrnehmbare Formgestaltungen”** (Multimedia) von Ideen... - und deren Verbreitung - **dank IKT!**

Was sind „Werke“? 1

- **UhG: “...eigentümliche geistige Schöpfungen auf den Gebieten der Literatur, der Tonkunst, der bildenden Künste und der Filmkunst”**
- **Von seinem Urheber (Schöpfer) in besonderer Weise geprägt – *muss es sich von Werken anderer Personen unterscheiden***
- **Persönlichkeit des Urhebers hat besonderen Einfluss auf das Ergebnis**
- **Kann einen Gebrauchszweck haben (→ nicht nur “echte” Kunst)**
- **Verwendete Technik ist irrelevant**

Was sind „Werke“? 2

- **Werke der Literatur:** (literarische) Sprachwerke, Computerprogramme, Bühnenwerke, Werke wissenschaftlicher oder belehrender Art
- **Werke der bildenden Künste:** Bild, Gemälde, Skulptur, u.ä. (ohne Anforderung an Dauerhaftigkeit), Lichtbildwerk (vs. einfaches Lichtbild)
- **Werke der Filmkunst:** Filmwerk (unabhängig vom Verfahren und Aufzeichnung – NICHT: Laufbilder; auf Tonspur evtl. eigenes Werk) – Computerspiele, Flashes, Animationen, ggf. auch Bildschirmschoner
- **Werke der Tonkunst**
- Sammelwerke, Bearbeitungen, Werkteile ...

Was sind „Werke“? 3

- Algorithmen in Informatik **nicht** geschützt – nur deren Implementierungen
- “zufällig” entstehende oder transiente Werke
 - Musik-Improvisationen
 - Chats
 - Vorträge
- Werke und Teile von Werken (Werkteile)

→ Beurteilungsprobleme:

→ “Schöpfungshöhe” für jeden Fall individuell zu definieren

- Rohfassung eines Romans?
- Detaillierte Spezifikationen der Klassen eines Programms?
- Welche Teile einer Datenbank?
- Kollektiv geschaffene Werke?

Geschützt in TDB?

- Kompositionen (✓)
- Musik- und Tonaufnahmen (✓)
- *(Reproduktionen von)* Bildern (✓), Skulpturen (✓) ...
- *(Reproduktionen von)* Bühnenaufführungen, Ballett, ... (✓)
- *(Reproduktionen von)* Werke der Architektur (✓)
- Rundfunkübertragungen (✓)
- Texte ✓ *(aber nicht Definitionen)*
- (bestimmte) Formeln ✓ *(auch Patentierung)*
- Eigennamen ✓ *(insbes. Marken, auch Logos)*
- Filme ✓ *(Filmwerke, nicht Laufbilder; ggf. Webseiten)*
- Fotos ✓ *(Bildwerke, nicht einfache Lichtbilder) & bestimmte graphische Darstellungen ✓*
- Computerprogramme ✓ *(evtl., auch Patentierung)*

Was ist ein Urheber?

- **Ursprünglicher Urheber eines Werkes (in EU/EWR) ist **ausschließlich** diejenige physische Person, welche das Werk tatsächlich geschaffen hat (inkl. anonyme Urheber)**
- **Miturheber können sein:**
 - **Echte Urheber, wenn die Teile des Werkes unabhängig („teilbar“) sind, jeweils für ihren Teil**
 - **Miturheber zu gleichen Teilen**
(=zur gleichen Hand, auch wenn Anteile verschieden)
 - **Zuarbeiter? →KEIN Schutz**
 - **Ersteller von Sammelwerken? →bedingt**
- **Bearbeiter, Übersetzer usw. schaffen neues Urheberrecht – nicht bei rein technischer Bearbeitung**
→u.U. Zustimmung des ursprünglichen Urhebers erforderlich

- Datenbank – Datenbankwerk?
- **„sui generis“** Investitionsschutzrecht des Datenbankherstellers
 - Rechtsgrundlage: EU-Datenbankrichtlinie 1996
 - Schutz sui generis gilt unabhängig vom Urheberrechtsschutz eines Datenbankwerks
 - ausschließliches Recht der Vervielfältigung, Verbreitung und öffentlichen Wiedergabe der gesamten DB sowie nach Art und Umfang wesentlicher Teile
 - Verlängerung bei Aktualisierung

Translation Memory (TM)

- Das Urheberrecht an **Sprachwerken** ist **unabhängig** von den Rechten an **Datenbanken/ Datenbankwerken** zu beurteilen
- TM als Datenbank: Sammlung unabhängiger Elemente UND **wesentliche Investition**
- TM als Datenbankwerk: Sammlung unabhängiger Elemente UND **persönliche geistige Schöpfung**
(Auswahl **oder** Anordnung des Inhalts der Datenbank)

Astrid Cruse, Starnberg, 2014

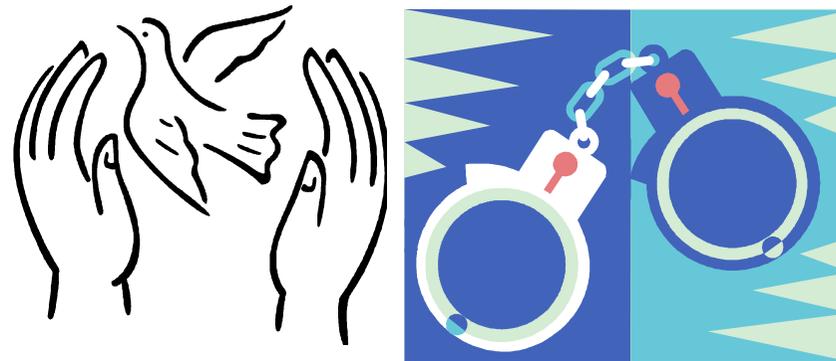
Rechtsverletzung durch Aufnahme in TM

Es gibt **kein ausdrückliches Recht**, das den Urheber eines geschützten Werks vor der Aufnahme in eine Datenbank schützt.

Aber andere Rechte können verletzt werden:
Vervielfältigungsrecht, Bearbeitungsrecht etc.

Astrid Cruse, Starnberg, 2014

Grundsatz der Vertragsfreiheit



→ **Vertragsrecht**

Astrid Cruse, Starnberg, 2014

Auswirkungen der IKT 1

- Vervielfältigungen von urheberrechtlich geschütztem Material in Originalqualität
- technische Sperren? problematisch
 - „blind“ gegen die gesetzlich vorgesehenen Differenzierungen und Freiräume
 - Freiräume könnten einseitig durch Rechteinhaber eliminiert werden
- Konvergenz Internet – Telekommunikation – Multimedia

Auswirkungen der IKT 2

- IKT und **Urheber** → Urhebercodierung?
 - IKT und **Investoren**
 - IKT und **Nutzer** → Nutzercodierung?
 - IKT und „**Mittler**“:
 - Verwertungsgesellschaften:
Nanopayment, Micro-credits
 - Verteilungsplattformen:
Monopole? Open Source? Eigene Produkte und Dienstleistungen?
- Mittler treten potenziell in Konkurrenz zu Urheber / Investor / Nutzer
- Rechtsunsicherheit bei Investoren / Nutzern
- Offene Welt ↔ geschützte Welt?
Beide Welten parallel?



Makroebene: Identifikation der Quelle

- ISBN, ISSN, ISAN, ISTC, ISWC, ...
- DOI (ISO 26324:2012), ISLI (ISO/DIS 17316), database elements (ISO/TS 29002-5:2009), ...
- Codierung von Lizenzkonditionen, z.B. CC-BY-SA

+ Technologie (über alle Endgeräte): Nanopayment u.a.

Mikroebene: DRM generisch einbauen in Datenmodelle

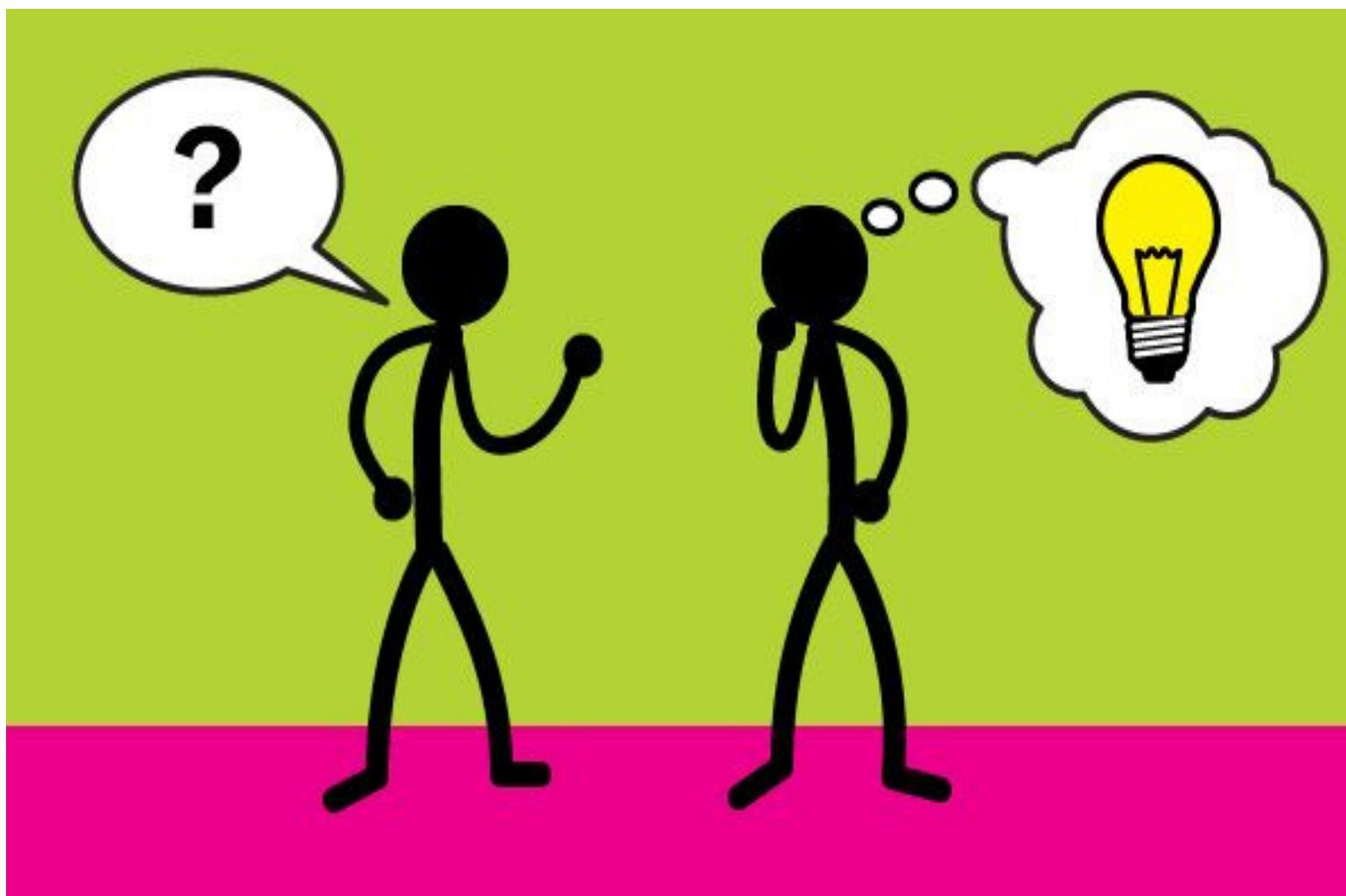
- Metamodelle für alle Arten von strukturiertem Content: TMF, LMF, ... OntoIOp
- Metadaten/Datenkategorien für die DRM-relevanten Codierungen
- eindeutige Codierung von Nutzungs-/Lizenzkonditionen, z.B. CC-BY-SA
→ ISO/TC 46?
- eindeutige Codierung der Entnahmestellen und deren nutzerfreundliche Umwandlung in bibliographische Angaben → ISO/TC 46 & ISO/TC 37?

1. **Eineindeutige Identifikation der Quelle** (bzw. C.-einheit)
2. **Angabe der Entnahmestelle des geschützten Elements**
3. **Angabe von Zeit/Zeitraum des Zugriffs/der Nutzung**
4. **Angabe ob urheberrechtlich geschützt oder nicht**
5. **Angabe der Art der Lizenz (codiert)**
6. **Angabe der Nutzungsbedingungen (codiert)**

(1) resource ID	(2) slot	(3) time stamp	(4) IPR y/n	(5) licence	(6) use conditions
-----------------	----------	----------------	-------------	-------------	--------------------

→ möglichst automatisch (=nutzerfreundlich): Hilfen bei der Eingabe, Warnungen bei der Übernahme/-gabe, usw.

Vielen Dank! Gibt es Fragen?





Dr. Christian Galinski
Internationales Informationszentrum für
Terminologie – Infoterm
1060 Wien, Gumpendorferstraße 65/1
christian.galinski@chello.at

References:

<http://www.copyright.com/Services/copyrightoncampus/content/digital.html>

<http://en.wikipedia.org/wiki/Copyright>

http://en.wikipedia.org/wiki/Universal_Copyright_Convention

http://en.wikipedia.org/wiki/Creative_Commons

GALINSKI, Christian and WRIGHT, Sue Ellen. Copyright and terminology. In: Sue Ellen WRIGHT and Gerhard BUDIN, comps. Handbook of terminology management. Vol. 1: Basic aspects of terminology management. Amsterdam/Philadelphia: John Benjamins, 1997, pp. 281-302

GALINSKI, Christian and GOEBEL, Jürgen W. Guide to Terminology Agreements. Vienna: International Network for Terminology (TermNet), 1996, 7+35 p.

GALINSKI, Christian and Inke Raupach and Klaus-Dirk Schmitz. Standard document templates and procedures. Deliverable D2.2 of the EU-Project EUROTERMBANK (Collection of Pan-European Terminology Resources through Cooperation of Terminology Institutions), 2005 (published 2006)

CRUSE, A.; Furlinger, M.; Galinski, Ch.; Raupach, I.; Schmitz, K.-D. Urheberrecht an Terminologie [Copyright for terminological data]. In: Drewer, P.; Pulitano, D.; Schmitz, K.-D. (eds.). Terminologiearbeit. Best Practices 2.0 [Terminology work. Best practices 2.0]. Köln: Deutscher Terminologie-Tag e.V., 2014. Modul 7, 16 p.