

Zur Urheberrechtlichen Zulässigkeit von Screen Scraping

Institute for Information Processing and
Microprocessor Technology (FIM)
Johannes Kepler University Linz, Austria

E-Mail: sonntag@fim.uni-linz.ac.at
<http://www.fim.uni-linz.ac.at/staff/sonntag.htm>



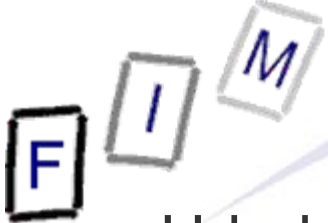
Was ist “Screen Scraping”?

- Auslesen von Daten „vom Bildschirm“, heute praktisch fast ausschließlich aus an Menschen adressierten Webseiten
 - Automatisierter Abruf der Webseite
 - Interne Weiterverarbeitung: Filtern, erweitern, umbauen, ...
 - Optional: Speicherung für später, zB in einer Datenbank
 - Weiterleitung der extrahierten Daten an Endkunden
 - » Meist wieder über das Internet
- Typische Beispiele: Preis- (oder: Meta-, ...) Suchmaschinen
 - Greifen auf die Webseiten von Anbietern zu, extrahieren dort Produktbezeichnung/-details und Preis
 - Ergebnisse mehrerer Anbieter werden zusammengefasst
 - Darstellung für den Endkunden
 - Variabel: Permanent gespeichert oder „live“ auf Anfragen hin



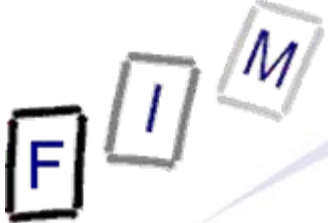
Urteilsübersicht Urheberrecht

- Je zwei Urteile aus Deutschland:
 - **Flugbuchungs-Website**
 - » Vergleich von Flugtickets bei verschiedenen Anbietern und indirekte Buchung bei diesen über die Vergleichs-Website
 - **Verkauf von Software für automatisierte Website-Abfrage**
 - » Einmal Online-Automobilbörse, einmal Flugbuchung
- Alle Fälle wurden auf UWG gestützt, drei zusätzlich auf Urheberrecht, und zwar auf „geschützte Datenbank“
- Drei der vier Fälle wurden als zulässig angesehen
 - **Eine Flugbuchungssoftware widersprach den Quellenanbieter-AGBs, da die Tickets weiterverkauft wurden**
 - » Im zweiten Fall war die Webseite lediglich eine technische Zwischenstation und Endkunden kauften die Tickets direkt!
- Weitere Urteile betreffend die Übernahme fremder Webinhalte in eigene Webseiten → Keine eindeutige Linie



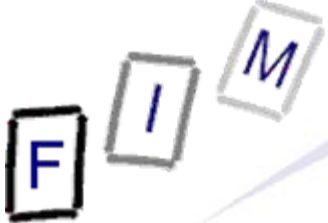
Urheberrechtliche Problematik

- Urheberrechtlicher Schutz besteht nur für Werke, aber nicht für bloße Fakten (wie zB Produkte und Preise)
 - Die bloße Übernahme von Fakten ist daher zumindest urheberrechtlich unbedenklich
- Daher der Ansatz, eine geschützte DB zu behaupten, die „ungebührlich“ ausgenutzt wird
 - Ob eine solche vorliegt, ist oft ein Streitpunkt!
- Deutlich klarer ist die Lage bei Werken, doch auch dort ist zu differenzieren:
 - In die bloße Abfrage (=Vervielfältigung),
 - eine etwaige Speicherung auf Vorrat, sowie die
 - Wieder-Darstellung im Internet
- Darüber hinaus stellt sich die Frage, **wer** dies durchführt
 - Der Anbieter des screen scraping?
 - Der Endnutzer mittels der Website als Werkzeug?



Datenbank

Nur Fakten, keine Werke



Entnahme von Daten aus geschützter DB

- Für ein Verbot müssten wesentliche Teile der DB entnommen werden (+ Vervielfältigung/Veröffentlichung)
 - Ausnahme: Zustimmung des Berechtigten
 - » Diese liegt meistens nicht vor (siehe aber Preissuchmaschinen!)
- Weiters muss es sich nach Art oder Umfang um wesentliche Teile der Datenbank handeln
 - Gesamte DB auslesen und speichern → Verboten
 - » Spider + screen scraping → Illegal bei geschützten DBs!
 - Live-Weiterleitung von Anfragen:
 - » Jeder Benutzer greift nur auf einen kleinen Teil zu
 - » Selbst bei vielen Benutzern sind diese separat zu betrachten
 - Lauter „Einzeltaten“; Jeder Benutzer handelt aus eigenem Antrieb!
 - » Ebenso bei dauernder Wiederholung der gleichen Anfrage
 - Überprüfung auf neu hinzugekommene Daten in der DB
 - » Haftung nur, wenn ein Benutzer alleine dies erreichen würde!
 - Begrenzung der Suchmöglichkeiten erforderlich



Entnahme von Daten aus geschützter DB

- Weiteres Verbot:
 - Entnahme unwesentlicher Teile, dies jedoch
 - wiederholt und systematisch, sowie
 - unzumutbare Beeinträchtigung der Herstellerinteressen, oder
 - Behinderung der normalen Verwertung der DB
- Beurteilung wie oben: Individuelle Benutzer können nicht zusammengerechnet werden
 - Auch hier ist Akzessorietät immer nur separat möglich
 - Für einzelne Benutzer aber hier viel leichter möglich!
- Verletzung daher unwahrscheinlich, aber potentiell möglich
 - Reine Zusatzdienste → Unbedenklich
 - Konkurrenzierung oder negativer Einfluss → Gefährlich!
 - » Eine gewisse Beeinträchtigung ist hinzunehmen
 - Behinderung der Verwertung der DB: Wie weit dies geht ist unklar, aber wohl eher eng (Verkauf von Zusatzprodukten?)



Abruf der Daten

Ab hier Werke, nicht mehr Datenbanken



Abruf = Vervielfältigung

- Werkschutz entsteht sehr schnell – auch kurze Texte, zB RSS Feed-Einträge, können Werke sein
 - Bilder, Videos, Software, ... → Fast immer ein Werk
- Automatisierter Abruf (sowie Speicherung ...) ist daher (potentiell) eine Verletzung der Ausschließlichkeitsrechte
- Gesetzliche Ausnahme?
 - Kein Caching: Filterung/Speicherung/Eigene wirtschaftliche Bedeutung
 - Kein Zitat: Gesamtes Werk/Einbettung (Beleg, ...)
- Zustimmung/Einwilligung?
 - Für wen werden HTML-Seiten produziert?
 - » Für Menschen, nicht für Maschinen!
 - Aber: Jeder Nutzer setzt eine Maschine (=Browser) ein!
 - » Zweck des Abrufs = Darstellung → Blinde dürften nicht surfen?



Einwilligung zur Vervielfältigung: Google Bildersuche

- „Anlocken“ bzw „Unterlassung der Verhinderung“ sowie „allgemeine Bekanntheit“ → Einwilligung
- Screen scraping:
 - Anlocken: Gleich wie bei Bildersuche!
 - » Allgemeine Suchmaschinenoptimierung
 - » Keine Optimierung für Abruf der Bilder (bzw. hier: DB/Werke!)
 - Unterlassung der Verhinderung: Ähnlich
 - » Kann exakt gleich wie Google erfolgen
 - » Screen scraping oft exakt wie Endbenutzer → Ev. anders
 - Allgemeine Bekanntheit: Je nach Größe
 - » Fluglinien kennen zB die Flug-Vergleichs-Websites sicher!
- Urteil ist aber stark diskussionswürdig!
- Zulässigkeit eher gegeben:
 - Persönliche Nutzung, anschauen, ausdrucken: Ja
 - Automatisierter Abruf: In Ausnahmen (Zweck) ev. nicht mehr



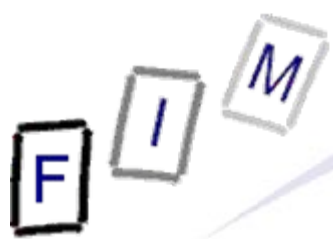
Wieder-Veröffentlichung der Daten

Zurverfügungstellung bzw Weiterleitung



Zurverfügungstellung

- Eine Einwilligung zur Zurverfügungstellung ist bei bloßem ins-Internet-stellen nicht anzunehmen (aber: BGH!)
 - „Caching“ von Webseiten durch ins Netz stellen einer Kopie wurde ansonsten bisher eindeutig als unzulässig angesehen
- Ähnlicher Fall: RSS-Feed wird in andere Website integriert
 - Bei RSS kann man eine Einwilligung zu automat. Abruf jedenfalls annehmen: Kein Mensch wird diese direkt lesen
 - Davon strikt zu trennen ist aber die Wieder-Veröffentlichung!
 - » Wurde in eine Website eingebunden → Anders als RSS-Reader
- Bei Werken ist screen scraping daher nur zulässig, insofern es von jedem Benutzer individuell durchgeführt wird
 - Soweit dieser Benutzer die Daten überhaupt abrufen darf!
 - Dann liegt keine Zurverfügungstellung mehr vor
 - Jedoch kein erneutes posten ins WWW erlaubt, also keine „häufigsten/beliebtsten/... Ergebnisse“ etc.



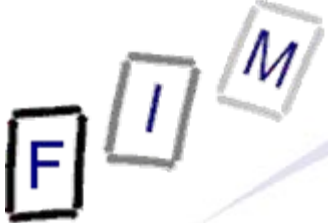
Individueller Abruf = Privatkopie?

Screen scraping als technisches Hilfsmittel



Rechtfertigung als Privatkopie?

- Digitale Vervielfältigung der Werke → Privatkopie?
 - Funktioniert sicher nicht bei Speicherung, sondern höchstens bei Live-Weiterleitung der Daten
 - » Screen scraping: Archiv potentiell erlaubt, siehe bei Werken aber § 42 Abs 5 UrhG → Nur reine Live-Durchleitung möglich
 - Keine Zugänglichmachung für Öffentlichkeit
 - Problem bei geschützten DB und bestimmten Werkarten (DB-Werk, Computerprogramme) → Keine Privatkopie erlaubt!
- Eingesetzte Werkzeuge können auch sehr komplex sein sowie von Dritten betrieben werden
 - Shift.tv: Online Videorekorder können Werkzeuge sein
 - Vervielfältigung muss aber „techn.-mech. Vorgang“ sein
 - » CB-Infobank I: Suchen passender Artikel und kopieren → Nein!
 - » Kunde müsste Anweisung geben, was genau zu kopieren ist
 - Daher: Keine eigenen Kriterien wie „beste“, „beliebteste“, ...
 - » Nur objektive Kriterien erlaubt: Preis, Verfügbarkeit, ...



- Screen scraping ist urheberrechtlich im allgemeinen zulässig bei **bloßen Fakten** bzw **Live-Durchleitung**
- Verboten oder zumindest gefährlich sind jedoch
 - Zulassen von Gesamtabfragen: Mindestlängen für Suchworte, Ergebnisanzahl-Beschränkungen vorsehen
 - Automatisierte Variation von Suchanfragen bzw indirekte Feststellung (alles abrufen und auf „Eignung“ untersuchen)
 - » **Selbständiges abrufen der gesamten Datenbank**
 - Bearbeitung der Ergebnisse nach eigenen Kriterien
 - » **Nur vom Endkunden vorgegebene sind problemlos**
 - Ergebnis-Speicherung für späteren Abruf oder auf Vorrat
 - » **Einwilligung erforderlich**
 - Keine Umgehung technischer Schutzmaßnahmen
 - » **Unabhängig von ihrer Wirksamkeit**

F I M

Fragen?

Vielen Dank für Ihre Aufmerksamkeit!