Legal session: copyright status of statistical data, privacy issues

JISC Usage Statistics Workshop

Statistics as Facts

- Copyright protects expression, not fact.
- Facts per se have no protection under US / EU law.
- Web log example: this person from this url accessed this web page at this time.
- Can anyone claim that an individual entry in a web log has the originality necessary for protection?

Statistics as processed data

- Processing web log entries requires software.
- Does the output of an analysis program have enough originality to qualify for protection?
- If the analysis requires judgments about statistics, does the degree of originality increase?

Statistics as database data

- Does a web log represent a database?
- how much organization / management does data require to become a database?
- Under US law, a database has no legal protection
- Under EU law, a database has limited protection

Statistics as private data

- Raw log files can hold private data
- How much can be stripped away to protect privacy without losing information?
  - This depends on the granularity of the analysis
  - It also depends on the context.

Copyright enforcement: law

- Copyright enforcement requires rightholder action.
- Police do not enforce copyright
- If a rights holder does not care about infringement, no action occurs.

Copyright enforcement: society

- To a significant extent, copyright enforcement is a socially defined issue, not a legal one.
- US courts recognized significant forms of infringement as “fair” before 17 USC 107 came into the law code.
- Statistics fit this model as well.

Statistics and information economics 1

- The key factor in determining whether a use is fair has become the effect on the value of a work.
- As long as log files and other web-based statistics are not perceived as having economic value, the question of whether they have protection likely will be moot.

Statistics and information economics 2

- In other words, the more value we find in something like the statistics from a log-file analysis, the more likely we make it that people will dispute the right to use the data.

DFN Deutsches Forschungsnetz

Logfiles and Statistical Data under Data Privacy Law

Legal Requirements for Server-Logfiles and Repositories

Hannes Obex

University of Muenster

Institute for Information-, Telecommunication- and Media-Law

Outline

I. Introduction into US Data Privacy Law

II. Introduction into UK Data Privacy Law

III. Introduction into German Data Privacy Law

IV. Specific Regulations for Electronic Communication in Germany

V. Webserver-Logfiles and Repositories

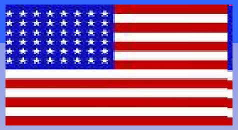
VI. Publication of Statistical Data

VII. Conclusion

I. Introduction into US Data Privacy Law

- Data protection not a written fundamental right in Federal Constitution (Bill of Rights)
- U.S. Supreme Court recognises right to privacy → Data Protection
- Binding only for government agencies, not for private persons or entities

I. Introduction into US Data Privacy Law

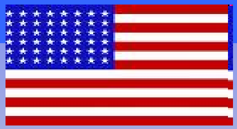


- States have own constitution
- Some explicitly recognise privacy protection
- But protection equal to Federal Constitution (not applicable for private persons)

I. Introduction into US Data Privacy Law

- Federal Privacy Protection Act (1974)
  - E.g. prior consent and information of person concerned
  - Only binding for public authorities
- Telecommunications Act (1996)
  - Applicable for telecommunication carriers
  - Telecommunication services have to be in return for payment
  - Mainly for access providers

I. Introduction into US Data Privacy Law



- Self Regulation
  - Idea: consumer and market will reward privacy protection
  - Declaration of privacy statements (e.g. of Online-Privacy Alliance)
  - Enforcement and surveillance by seal programs (Trustmarks)
  - Requirements: Notice and Choice

    (“good notices of bad practices”)

I. Introduction into US Data Privacy Law

- Safe Harbour
  - Commission of European Union: Level of data protection in the US not sufficient
  - No data transfer from EU to US
  - Exception: Consent or Safe Harbour
  - Higher protection level than US self regulation, but less than EU
  - 1506 companies (01/07/2008)

I. Introduction into US Data Privacy Law

- CONCLUSION
  - Practically no privacy protection legislation for the private electronic communication sector
  - Self regulation and seal programs
  - Safe Harbour (permitting data transfer with EU)
  - Data protection level lower than in the EU

II. Introduction into UK Data Privacy Law

- Influenced by European legislation
- Implementation of EU directives at national level

→Legal situation and basic principles similar to Germany (III.)

- Data transfer to/from Germany permitted

III. Introduction into German Data Privacy Law

- German Federal Constitutional Court (1983):
  - „Right to decide, when and to whom disclose personal data“
  - „there is no irrelevant data“
  - Usage of personal data needs specific legal regulation
  - Principle of notification and choice
  - Principle of data minimisation
  - Principle of purpose limitation
  - Principle of proportionality

III. Introduction into German Data Privacy Law

- Results of this decision
  - Data Privacy is guaranteed by the constitution (right of informational self-determination)
  - Every government agency is directly bound
  - Legislator has to provide protection against intrusion by private persons or entities
  - Crucial statements incorporated into Data Privacy Laws (e.g. Federal Data Protection Act, State Law, specific legislation)

III. Introduction into German Data Privacy Law

- European Law
  - Strong influence on German legislation by implementation at national level
  - Several Directives concerning Data Privacy
  - Data transfer within the EU is permitted
  - Data transfer to other countries forbidden, unless adequate level of data protection (e.g. Safe Harbour)

IV. Specific Regulations for Electronic Communication

- Telecommunications Act (TKG)
  - Section 91 et seqq.
  - Applicable for telecommunications service providers
  - TKG only for technical aspects (not for content), therefore only for access providers

IV. Specific Regulations for Electronic Communication

- German Telemedia Act (TMG)
  - Section 11 et seqq.
  - Applicable for access, host and content provider
  - Content-related services
  - Applicable for IP-Logfiles, user profiles, cookies, etc.

IV. Specific Regulations for Electronic Communication

- Basic Principles
  - – Data collection and usage forbidden, unless:
    - Specific statutory authorisation
    - Consent by user
  - – Clear and comprehensive information
  - – Right to object certain data processing (and to still use the service)
  - – Limitation of purpose

IV. Specific Regulations for Electronic Communication

- Personal Data
  - Data which refers to personal or factual circumstances of a natural person
    - Personal: age, gender, name
    - Factual: property, income, bank balance
  - Person has to be identified or identifiable for the data processor

V. Webserver-Logfiles and Repositories

- Logfiles and user profiles
- Control and improvement of online services
- Crucial question: Logfiles = personal data?
  - Static IP-number: identification possible (e.g. whois-query)
  - Dynamic IP-Number: access provider can identify user
  - Can content oder host provider identify users?

V. Webserver-Logfiles and Repositories

- Decision AG/LG Berlin
  - Person identifiable, if there is **any** possibility of content provider to get the information
  - Not only legal possibilities (especially protection against illegal measures)
  - Content provider might (illegally) receive user information from access provider
  - Therefore: Dynamic IP-Number = Personal Data

V. Webserver-Logfiles and Repositories

- Not prevailing case law or legal opinion!
  - Very disputed decision
  - Legal experts: only adequate and legal possibilities taken into consideration
  - Recent decision by lower courts
  - Dynamic IP-Number by itself is may not be personal data for content provider

    (unless content provider receives identification from user, e.g. by registration/login)

V. Webserver-Logfiles and Repositories

- Authorisation to process personal data
  - Section 15 TMG:Usage Data, especially:
    - Identifying Features (IP-Number)
    - Information about time and extent of use
  - Sec. 15 paragraph 1 sentence 1: data essential for providing the telemedia service and for billing

    → Not allowed, if service is free

    → Not essential for delivering the service

V. Webserver-Logfiles and Repositories

- Sec. 15 paragraph 3 TMG:
  - Purpose of advertising, market research and demand-responsive design of telemedia
  - User profiling allowed
  - Data has to be pseudonymised

V. Webserver-Logfiles and Repositories

- Anonymous data: data processor cannot (or only with unproportional effort) identify the user
- Pseudonymous data: data is encoded, so that information originating from one user can be related to other data from this user

  → decoding theoretically possible

  → measures to prevent decoding (Sec. 13 paragraph 4 no. 6)

V. Webserver-Logfiles and Repositories

- User has right to object to the profiling
- User has to be informed about data processing and his rights **at the beginning** of utilisation of telemedia service

V. Webserver-Logfiles and Repositories

- Cookies
  - Can contain personal data (IP, name, login etc.)
  - If later identification is enabled and collection or processing of personal data is prepared, user has to be informed (Sec. 13 paragraph 1 sentence 2)
  - Free use of cookies, if cookie uses identifier without connection to user‘s real identity (random code)

V. Webserver-Logfiles and Repositories

- Data transfer
  - Allowed for purpose of market research (Sec. 15 paragraph 5 sentence 3)
  - Data has to be anonymised (no identification possible)
  - Regulation conclusive for telecommunication media
  - If data transfer serves other purpose: consent required

V. Webserver-Logfiles and Repositories

- Consent
  - „informed consent“, i.e. user has to be informed about the type of data processed, the extent of the processing, the purpose etc.
  - Information has to be given prior to use of service / data processing
  - „free consent“, i.e. user must have the choice to use the service without agreeing to the data processing Consent must be given knowingly and clearly
  - It must be protocolled
  - Content must be accessable for user
  - User can revoke consent at all times

V. Webserver-Logfiles and Repositories

- CONCLUSION
  - If Logfiles contain personal data: TMG applicable
  - Statutory authorisation for creation of logfiles on the basis of pseudonymous data for the purpose of advertising, market research and demand-responsive design of services
  - Data transfer permitted fur purpose of market research, when data is anonymised
  - User has to be informed about data processing and his right to object to the creation of a profile
  - For every other data processing or transmission: user consent required

VI. Statistical Data of Documents

- Numbers of visitors of a certain documents
- „Hits“ (counter)
- Personal data, if IP-Number, user name etc. is logged
- If mere number of clicks: no person identifiable, therefore no personal data

VI. Statistical Data of Documents

- Rights of the author
  - Author might not approve if his number of hits is low compared with other users
  - Number of hits is no data concerning author (only shows users‘ behaviour)
  - Number of hits is a factual statement
  - As long as statistic is true: no defamation
  - RESULT: Statistical data can be published

VI. Conclusion

- Logfiles can contain personal data (controversal: dynamic IP-numbers)
- If this is the case, Sec. 11 et. seqq. TMG apply
- User profiles are allowed for certain purposes, but only with pseudonymised data
- User has to be informed and the right to object
- Data transfer is very restricted; consent recommended
- Statistical data of documents can be published, as long as they do not contain personal data

Thank you for your attention

hannes.obex@uni-muenster.de