

Hinweise und Empfehlungen zur Analyse von Internet-Angeboten mit „Piwik“

Metadaten:

Version:	1.2
Ausgabedatum:	15. März 2011
Status:	<input type="checkbox"/> in Bearbeitung <input type="checkbox"/> in Abstimmung <input checked="" type="checkbox"/> Freigegeben
Ansprechpartner juristisch:	Dr. Moritz Karg 0431/988-1651 ULD41@datenschutzzentrum.de
Ansprechpartner technisch:	Sven Thomsen 0431/988-1211 ULD3@datenschutzzentrum.de

Inhalt:

1.	Allgemeine Informationen zu Piwik.....	3
1.1	Was ist Piwik?	3
1.2	Wer steht hinter Piwik?	4
1.3	Wie funktioniert Web-Analyse mit Piwik?	5
1.4	Vorteile und Nachteile von Piwik.....	5
2.	Zur Installation von Piwik.....	6
2.1	Voraussetzungen.....	6
2.2	Vorbereitungen.....	6
2.3	Ablauf der Installation und Konfiguration.....	6
3.	Notwendige technische und organisatorische Maßnahmen.....	8
3.1	Plugin „AnonymizeIP“ einsetzen	8
3.2	Widerspruchsmöglichkeit bereitstellen	10
3.3	Referrer datensparsam verwenden.....	11
3.4	Verbot der Zusammenführung durchsetzen	11
3.5	Daten löschen.....	12
3.6	Lebensdauer von Piwik-Cookies setzen und verringern.....	12
4.	Rechtliche Hinweise zum Einsatz von Piwik	15
4.1	Cookies.....	15
4.2	Beschluss zur datenschutzkonformen Ausgestaltung von Analyseverfahren zur Reichweitenmessung bei Internet-Angeboten	16
5.	Fazit	17

1. Allgemeine Informationen zu Piwik

1.1 Was ist Piwik?

Piwik erstellt detaillierte Statistiken über die Besucher einer Webseite. Enthalten sind Informationen über die benutzten Suchmaschinen, die benutzten Suchbegriffe, die verwendeten Sprachen, die Herkunft der Besucher nach Ländern, die verwendeten Browser und ihre Plugins, die Referrer, die Aufenthaltsdauer, Einstiegsseiten, Ausstiegsseiten, Abbruchquoten, die IP-Adresse usw.

Piwik ist in der Programmiersprache PHP programmiert und wird quelloffen unter der GNU General Public License (GPL) bereitgestellt. Piwik benötigt zum Betrieb eine Datenbank; aktuell wird der Betrieb mit MySQL empfohlen.

Piwik kann unter <http://piwik.org/> (Aktualität des Links: 15. März 2011) kostenfrei bezogen werden.

Das Programm wird auf dem Webserver des Anwenders (im Folgenden: „Piwik-Verwender“) installiert und speichert Daten ebenfalls nur auf dem Server des Piwik-Verwenders. Es handelt sich also um eine sogenannte Inhouse-Lösung für Web-Analyse.

Piwik bietet eine Analyse der Nutzung einer oder mehrerer Webseiten in Echtzeit an. Dies bedeutet, dass aktuelle Besucher quasi sofort, nicht erst nach längerer Bearbeitungszeit, in den Statistiken erscheinen.

Eine Online-Demonstration (siehe auch Abb. 1) ist unter dem folgendem Link verfügbar:

<http://demo.piwik.org/> (Aktualität des Links: 15. März 2011).

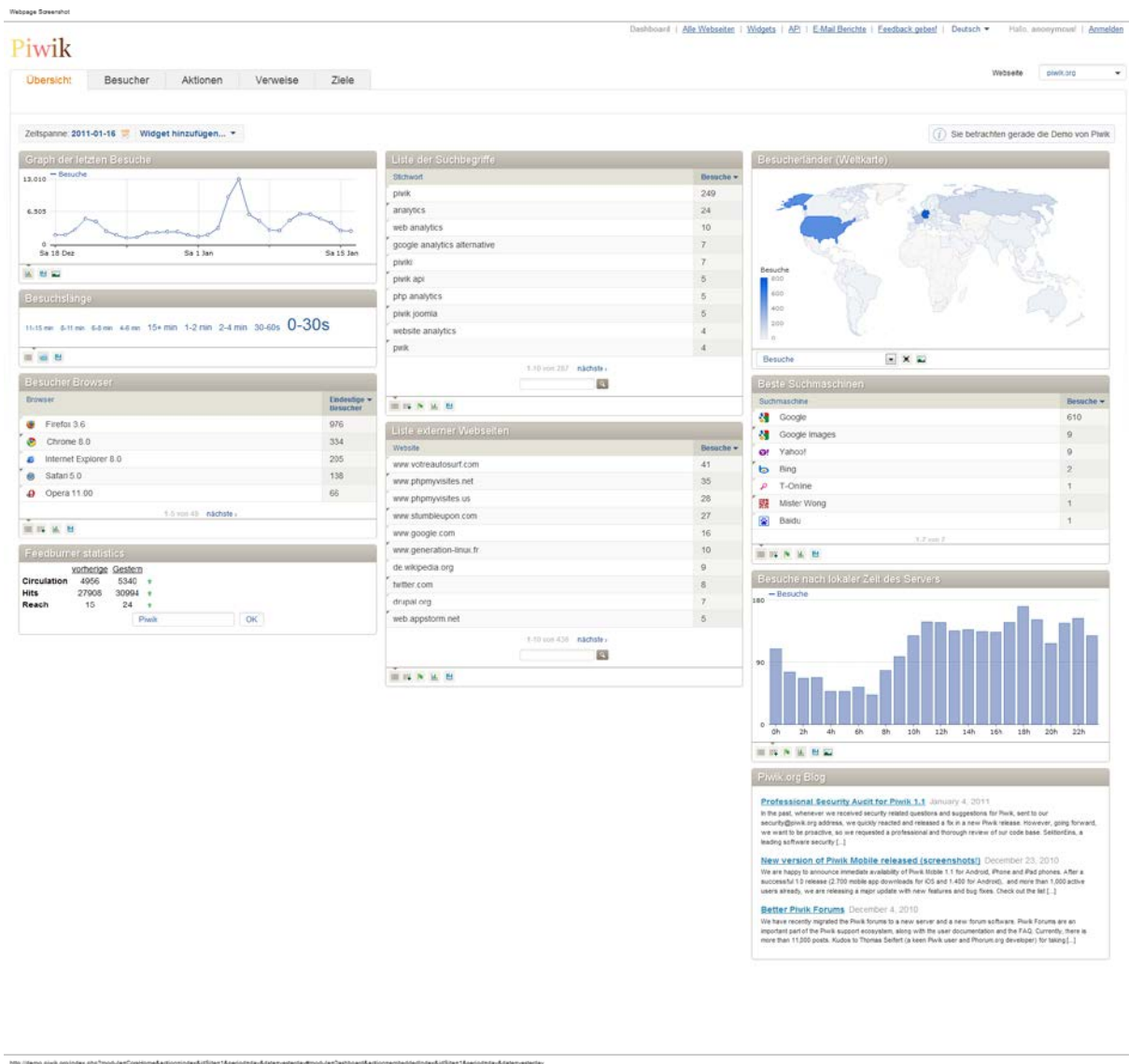


Abb. 1: Piwik-Screenshot

1.2 Wer steht hinter Piwik?

Hinter Piwik steht ein großes Entwicklerteam unter der Leitung von Matthieu Aubry. Eine Übersicht wichtiger Teammitglieder findet sich unter:

<http://piwik.org/the-piwik-team/> (Aktualität des Links: 15. März 2011).

Das Entwicklerteam scheint hinreichend stabil zu sein, so dass eine längerfristige Unterstützung und Verfügbarkeit von Piwik wahrscheinlich ist.

Die Software von Piwik wurde Ende 2010 einem Security Audit inklusive Code-Review unterzogen. Die Ergebnisse sind in die Version 1.1 eingeflossen.¹ Mittlerweile liegt Piwik in der Version 1.2 vor.

Piwik ist ein eingetragenes Warenzeichen.

¹ <http://piwik.org/blog/2011/01/professional-security-audit-in-piwik/> (Aktualität des Links: 15. März 2011).

1.3 Wie funktioniert Web-Analyse mit Piwik?

Der Piwik-Verwender installiert Piwik auf einem System, das die Skriptsprache PHP unterstützt und das Datenbankmanagementsystem MySQL bereitstellt. Er erhält von Piwik Programmanweisungen (JavaScript), die am Ende jeder zu analysierenden Webseite eingefügt werden müssen. Dies kann manuell erfolgen oder alternativ durch Plugins für eine Reihe von Content-Management-Systemen, Blog-Software etc. automatisiert durchgeführt werden.

1.4 Vorteile und Nachteile von Piwik

Der Hauptvorteil von Piwik besteht darin, dass eine Web-Analyse-Lösung unter der vollständigen Kontrolle und Steuerung der verantwortlichen Stelle umgesetzt werden kann und die Daten allein bei dem Piwik-Verwender verbleiben. Eine Weitergabe von Nutzerdaten an Dritte mit den entsprechenden datenschutzrechtlichen Folgeproblematiken (Übermittlung oder Auftragsdatenverarbeitung) erfolgt nicht.

Zudem ermöglicht die Verwendung von Piwik Statistiken in Echtzeit. Der Piwik-Verwender muss also nicht längere Zeit auf die Übernahme von Daten in die Statistik warten.

Zu den technischen Nachteilen von Piwik gehört es, dass die Besucherzählung zumindest in der Grundvariante des Programms davon abhängig ist, dass der Besucher JavaScript aktiviert hat. Piwik verwendet HTTP-Cookies. Genauer zur rechtlichen Situation rund um Cookies findet sich im Abschnitt 4.1 dieses Dokuments.

2. Zur Installation von Piwik

2.1 Voraussetzungen

Voraussetzungen für die Installation der Piwik-Version 1.2 sind²:

- PHP in der Version 5.1.3 oder darüber (5.3 wird empfohlen),
- MySQL in der Version 4.1 oder darüber,
- PHP Erweiterungen pdo und pdo_mysql.

Die PHP GD-Erweiterung wird benötigt, um die sogenannten Sparklines (kleine Grafiken) zu erstellen. Sie ist allerdings nicht nötig für die Grundfunktionalität, und die Installation kann auch ohne diese Erweiterung abgeschlossen werden.

2.2 Vorbereitungen

Die jeweils neueste Version von Piwik kann auf der Seite www.piwik.org heruntergeladen werden.

Man erhält einen ZIP-komprimierten Ordner. Wird dieser entpackt, erscheint ein Ordner namens „piwik“. Dieser wird auf den für die Durchführung der Analyse genutzten Webserver an die gewünschte Stelle gelegt. Dieses System muss nicht identisch mit dem Server sein, auf dem das zu analysierende Angebot liegt; im Folgenden wird dies jedoch aus Gründen der Einfachheit vorausgesetzt.

Anschließend muss zur weiteren Installation die entsprechende URL in einem Browser aufgerufen werden. Nun sollte die Willkommenseite der weiteren Piwik-Installation erscheinen. Es ist jedoch möglich, dass zunächst Fehlermeldung zu Ordnern ausgegeben werden, auf die Piwik nicht zugreifen kann, obwohl dies nötig wäre. Hier ist zunächst zu prüfen, ob die angegebenen Ordner auf dem Server vorhanden sind. Ist dies nicht der Fall, müssen sie angelegt werden. Anschließend muss überprüft werden, ob die entsprechenden Rechte korrekt gesetzt sind. Beispielsbefehle zum korrekten Setzen der Rechte werden von Piwik mit der Fehlermeldung mitgereicht.

2.3 Ablauf der Installation und Konfiguration

Der eigentliche Installationsprozess verläuft in 8 Schritten, die mit dem Anzeigen verschiedener Seiten einhergehen:

1. **Begrüßungsseite:** Zunächst erscheint eine Willkommenseite, und die Installation kann durch Klicken eines Links begonnen werden.
2. **System-Check:** Dann wird überprüft, ob der Server die nötigen Voraussetzungen erfüllt. Insbesondere wird angezeigt, wenn
 - der Integritätscheck hinsichtlich der hochgeladenen Dateien fehlschlägt. Dies geschieht zum Beispiel, wenn beim (allein aus Sicherheitsgründen schon nicht zu empfehlenden) Hochladen per FTP nicht der Binärmodus verwendet wurde.

² <http://piwik.org/docs/requirements/> (Aktualität des Links: 15. März 2011).

- nicht die richtigen Versionen von PHP, MySQL oder der pdo-Erweiterung vorliegen.

Auch die optionalen Voraussetzungen (siehe oben) werden überprüft und die Ergebnisse der Prüfung dargestellt. Zudem wird erneut die Zugriffsmöglichkeit auf bestimmte Ordner überprüft und im Fehlerfall eine Lösungsmöglichkeit angeboten.

Sind alle Mängel beseitigt worden, kann die Seite erneut geladen werden, die Fehleranzeigen verschwinden bei der Aktualisierung, und der Prozess kann fortgesetzt werden.

3. **Einrichten der MySQL Datenbank:** Als Nächstes erscheint eine Seite mit einem Formular zum Eingeben der für die Konfiguration der Datenbank benötigten Daten: MySQL-Server, Login, Passwort, Datenbankname und Tabellenpräfix.
4. **Datenbank und Tabellen erstellen:** Piwik erstellt die Tabellen und erforderlichenfalls die Datenbank automatisch. Ist dies geglückt, erscheint eine Seite mit einem entsprechenden Erfolgsbericht und der Möglichkeit, den nächsten Schritt durch Anklicken einzuleiten.
5. **Allgemeine Einstellungen:** Auf der nächsten Seite wird der sogenannte „Super User“ mit Login, Passwort und E-Mail-Adresse eingetragen. Der „Super User“ soll als Einziger die Rechte innehaben, um andere Benutzer anzulegen und zu verwalten, Webseiten zu integrieren und Plugins zu installieren und zu verwalten.
6. **Webseite einrichten:** Als Nächstes erscheint eine Seite, auf der die erste Webseite für die Analyse eingerichtet werden kann. Andere können nach der Installation in unbegrenzter Anzahl nachgetragen werden. Die Eintragung der Webseite erfolgt durch die Angabe eines Namens und die Angabe der URL.
7. **JavaScript-Code:** Im nächsten Schritt wird ein Stück JavaScript-Code angeboten, das kopiert und in den Quelltext aller Seiten eingefügt werden muss, die betrachtet werden sollen. Der Code soll am Ende der Seite eingefügt werden (z.B. direkt vor dem </body>-Tag).
8. **Abschlussseite:** Zum Abschluss erscheint eine Seite, die zu der erfolgreichen Installation gratuliert und einen Link zu den laufenden Statistiken bietet.

3. Notwendige technische und organisatorische Maßnahmen

3.1 Plugin „AnonymizeIP“ einsetzen

IP-Adressen dürfen als personenbezogene Daten nicht ungekürzt gespeichert werden³.

Piwik bietet in diesem Zusammenhang das Plugin AnonymizeIP an. Dieses wird bei der aktuellen Piwik-Version bereits mitgeliefert und muss lediglich vom „Super User“ aktiviert werden. Dies geschieht durch Aufruf des Menüpunkts „Einstellungen → Plugins“. Die Anzahl der zu maskierenden Oktette ist in `config/global.ini.php` als `ip_address_mask_length` mit dem Default-Wert 1 festgelegt.

Ändert man also nichts an der Einstellung, wird bei aktiviertem AnonymizeIP-Plugin das letzte Oktett der IP-Adresse auf 0 gesetzt. Eine gespeicherte IP-Adresse sollte dann beispielsweise so aussehen: 128.91.97.0.

Die Anzahl der zu maskierenden Oktette kann in `config.ini.php` konfiguriert werden. Zulässige Werte sind für IPv4-Adressen 1 bis 4. Die Einstellung erfolgt durch das Einfügen einer neuen Wertangabe unter [Tracker].

Um beispielsweise nicht nur ein Oktett, sondern zwei Oktette der IP-Adresse zu anonymisieren, muss die Datei `config/config.ini.php` wie folgt geändert werden:

```
[Tracker]
```

```
ip_address_mask_length = 2
```

Hat der Verwender die Rubrik [Tracker] schon wegen einer anderen Konfiguration eingefügt (beispielsweise beim Setzen der Lebensdauer für die Web-Analyse-Cookies, s.u.), kann die Zeile mit der Maskenlänge einfach nur mit in dieser Rubrik eingetragen werden.

Eine Anonymisierung der IP-Adressen hatte in den Tests des ULD keine Auswirkung auf eine hinreichend genaue Analyse der Herkunftsländer der Besucher.

Im internen Test des AnonymizeIP-Plugins ist folgendes Phänomen aufgetreten: War das Plugin aktiviert, erschien als IP-Adressangabe bei allen internen Besuchern (ULD-Mitarbeitern) die Angabe 255.255.255.255. Dasselbe galt zu Beginn für einen großen Teil der durch den Besuchergenerator (ein weiteres Plugin zur Generierung von Testdaten) erstellten Besucher. Ein kleinerer Teil der generierten Besuchsdaten erschien mit korrekt anonymisierter IP-Adresse, z.B. 41.39.0.0 bei einer Einstellung mit zwei zu anonymisierenden Oktetten. Im weiteren Verlauf erschienen die meisten der generierten IP-Adressen mit korrekt anonymisierten IP-Adressen. Für die internen Besucher wurden weiterhin die IP-Adressen mit 255.255.255.255 angegeben.

War das AnonymizeIP-Plugin deaktiviert, erschienen alle Besucher, auch die internen, mit einer vollständigen IP-Adresse. Die gleichzeitige Nutzung des Anonymisierungs- und des Testdaten-Plugins scheint unerwünschte Seiteneffekte zu produzieren. Das Testdaten-Plugin sollte nur für interne Tests auf einer eigenen Test-Instanz eingesetzt werden.

Im Rahmen der Besucherwiedererkennung benutzt Piwik eine Heuristik, die versucht, einen Besucher mit einem vorherigen Besuch zu identifizieren, indem bestimmte Daten berücksichtigt werden.⁴ Insbesondere sind dies die IP-Adresse, die Auflösung, der Browser,

³ Siehe dazu auch <https://www.datenschutzzentrum.de/ip-adressen/> (Aktualität des Links: 15. März 2011).

⁴ Siehe auch http://piwik.org/faq/general/#faq_43 (Aktualität des Links: 15. März 2011).

die verwendeten Plugins und das Betriebssystem. Die Daten werden kombiniert und zu einem Hashwert verrechnet. Diese Heuristik ist dazu bestimmt, Besucher auch dann wiederzuerkennen, wenn bei ihnen kein Cookie gespeichert ist, unabhängig davon, ob diese nicht angenommen werden oder gelöscht worden sind.

Auch bei aktiviertem AnonymizeIP-Plugin nutzt diese Heuristik intern die volle IP-Adresse⁵. Die IP-Adresse wird jedoch nicht isoliert verarbeitet und gespeichert.

Die IP-Adresse geht mit anderen Daten verknüpft in einen Hashwert ein. Ein Rückrechnen ist nur mit erheblichem Aufwand denkbar. Die Wiedererkennbarkeit von Internetnutzern hängt zudem nicht unbedingt an der IP-Adresse, sondern kann mit überraschend großer Zuverlässigkeit auch über andere Werte vorgenommen werden.⁶

Die Nutzung der IP-Adresse zur Erstellung der Pseudonyme kann als eine zu rechtfertigende Datenverarbeitung gesehen werden: § 15 Abs.3 TMG erlaubt die Erstellung von Nutzungsprofilen zum Zweck der Werbung, Marktforschung und bedarfsgerechten Gestaltung von Telemedien. Grundsätzlich ist nach der Auslegung der Aufsichtsbehörden die Verarbeitung personenbezogener Daten, zu der auch die IP-Adresse gehört, für die Analyse des Nutzungsverhaltens unzulässig. Der Wortlaut des zur Auslegung des § 15 Abs. 3 TMG heranzuziehenden Beschlusses lautet:

„Die Analyse des Nutzungsverhaltens unter Verwendung vollständiger IP-Adressen (einschließlich einer Geolokalisierung) ist aufgrund der Personenbeziehbarkeit dieser Daten daher nur mit bewusster, eindeutiger Einwilligung zulässig. Liegt eine solche Einwilligung nicht vor, ist die IP-Adresse vor jeglicher Auswertung so zu kürzen, dass eine Personenbeziehbarkeit ausgeschlossen ist.“⁷

Die IP-Adresse wird bei diesem Nutzungsvorgang **nicht** zur Analyse des Nutzungsverhaltens, sondern zur Erstellung des Pseudonyms verwendet. Eine Restrisikoanalyse des ULD ergibt, dass eine Rückrechnung aus dem gebildeten Hashwert bei Kenntnis der anderen zusätzlich genutzten Parameter und innerhalb der Lebenszeit des Cookies zwar theoretisch möglich, jedoch mit derartigem Aufwand verbunden wäre, dass das Restrisiko in diesem Fall hinnehmbar ist. Dieser Nutzungsvorgang dürfte noch von dem Wortlaut des § 15 Abs. 3 TMG gedeckt sein. Da letztlich ein Herausrechnen der IP-Adresse aus dem Hashwert nur mit erheblichem Aufwand möglich ist, ist das Gefährdungspotenzial der Verletzung der Persönlichkeitsrechte der Nutzerinnen und Nutzer als äußerst gering einzustufen.

Eine unzulässige Verarbeitung der IP-Adresse wird derzeit seitens des ULD nicht gesehen. Denn die Verarbeitung der IP-Adresse dient der Bildung eines Pseudonyms, das wiederum dem Ziel der Datenvermeidung und Datenminimierung dient, §§ 3 Abs. 6a, 3a S. 2 BDSG. Das ULD steht mit den Entwicklern diesbezüglich jedoch im Kontakt, um anzuregen, dass nur gekürzte IP-Adressen zur Hashwertbildung herangezogen werden.

Das ULD empfiehlt, das AnonymizeIP-Plugin mit der Einstellung `ip_address_mask_length=2` zu verwenden.

⁵ Siehe auch <http://forum.piwik.org/index.php?showtopic=16721> (Aktualität des Links: 15. März 2011).

⁶ Vgl. hierzu <https://panopticlclick.eff.org/> (Aktualität des Links: 15. März 2011).

⁷ Siehe „Datenschutzkonforme Ausgestaltung von Analyseverfahren zur Reichweitenmessung bei Internet-Angeboten“, Beschluss der obersten Aufsichtsbehörden für den Datenschutz im nicht-öffentlichen Bereich am 26./27. November 2009 in Stralsund, <http://www.lfd.m-v.de/dschutz/beschlue/Analyse.pdf> (Aktualität des Links: 15. März 2011).

3.2 Widerspruchsmöglichkeit bereitstellen

Bei der Web-Analyse mit Piwik werden Cookies auf dem Rechner des Nutzers gespeichert. Zur rechtlichen Situation beim Einsatz von Cookies sei auf die Ausführungen im Abschnitt 4.1 verwiesen.

Die aktuelle Piwik-Version setzt eine Widerspruchsmöglichkeit mit Hilfe eines IFrames um, der den Besuchern dann beim Erstaufwurf der Seite und beispielsweise generell in der Datenschutzerklärung auf der Webseite präsentiert wird. Der Einsatz von Reichweitenanalysediensten ohne Widerspruchsmöglichkeit stellt einen Verstoß gegen § 15 Abs. 3 TMG dar. Der Einsatz des Analysedienstes ohne die angebotene Widerspruchsmöglichkeit ist datenschutzrechtlich unzulässig.

Der notwendige Code muss entsprechend auf der Web-Seite eingebunden werden.

Beispiel:

```
<iframe      frameborder="no"      width="600px"      height="200px"
src="http://www.example.com/x_piwik/index.php?module=CoreAdmin
Home&action=optOut"></iframe>
```

Zu beachten ist hierbei, dass derzeit⁸ der Hinweistext auf Englisch erscheint. Eine Übersetzung muss manuell in der Datei /lang/en.php beziehungsweise /lang/de.php vorgenommen werden. Da § 15 Abs. 3 TMG in Verbindung mit § 13 Abs. 1 TMG eine Unterrichtung der Betroffenen verlangt, hat die Information über das Widerspruchsrecht in allgemein verständlicher Form zu erfolgen. Sollte sich das Angebot hauptsächlich an Deutsche richten und die Information über einen Hinweistext erfolgen, ist eine deutsche Sprachversion zu verwenden.

Der Widerspruch wird in Form eines Opt-Out-Cookies abgelegt, so dass beispielsweise nach einem Löschen aller Cookies das Opt-Out erneut erklärt werden muss.

Auf diesen Umstand sollten Nutzerinnen und Nutzer eindeutig und verständlich hingewiesen werden. Anderenfalls könnte der Eindruck entstehen, ein einmal eingelegter Widerspruch wäre in jedem Fall ausreichend.

Beim Einlegen eines Widerspruchs wird das bestehende Tracking-Cookie erst ab der Version 1.2 von Piwik gelöscht. Vorherige Versionen löschen das Tracking-Cookie nicht und sollten daher nicht eingesetzt werden.

Datenschutzrechtlich ist ein erklärter Widerspruch nur dann wirksam umgesetzt, wenn keine identifizierenden Informationen durch den Diensteanbieter gesetzt werden, d.h. bereits die Erstellung eines Pseudonyms nicht erfolgt. Denn die datenschutzrechtlichen Vorgaben zur Umsetzung des Widerspruchs zielen auf die Unterbindung der (Wieder-) Erkennung der Betroffenen. Diensteanbieter müssen daher bei dem Einsatz der Analyse- software sicherstellen, dass im Fall der Einlegung eines Widerspruchs identifizierende Merkmale nicht gesetzt bzw. bestehende Merkmale gelöscht werden.

Für die Handhabung und Umsetzung des Widerspruchs besteht momentan noch Raum für Verbesserungen. Trotzdem bietet Piwik mit den aktuell bestehenden Instrumenten weit mehr Kontroll- und Widerspruchsmöglichkeiten als andere Systeme.

⁸ Februar 2011.

Piwik bietet in der Version 1.2 eine zusätzliche Möglichkeit, um Widersprüche zu verarbeiten. Hierbei wird ein noch relativ neues Verfahren mittels „X-Do-Not-Track“⁹-Headern angewendet. Das ULD hat die Umsetzung dieser Funktionen in Piwik zum Zeitpunkt der Erstellung dieses Dokuments noch nicht abschließend geprüft.

Das ULD empfiehlt, die Widerspruchsmöglichkeit beim Erstbesuch der Seite prominent und ansonsten dauerhaft z.B. in der Datenschutzerklärung der Webseite anzubieten.

3.3 Referrer datensparsam verwenden

Die Referrer-Daten enthalten Informationen über den Anlass des Seitenaufrufs, z.B. zu den verwendeten Suchmaschinen und Suchbegriffen, den Herkunfts-Webseiten, Kampagnen und direkten Seitenaufrufen. Es besteht eine theoretische Möglichkeit, dass über die durch den Browser des Nutzers übertragenen Referrer-Informationen Rückschlüsse auf den Nutzer gezogen werden können und außerdem eine Zusammenführung der pseudonymen Nutzungsprofile mit identifizierenden Angaben erfolgen kann.

Das ULD weist darauf hin, dass eine derartige Profilbildung unzulässig sein kann, was auch einen Bußgeldtatbestand darstellt, § 16 Abs. 2 Nr. 5 TMG. Eine nach § 15 Abs. 3 TMG mögliche Nutzerprofilierung ist nach Ansicht des ULD nur zulässig, wenn die verantwortliche Stelle zweifelsfrei nachweist, dass die Erstellung der Profile für Werbung, Marktforschung oder die bedarfsgerechte Gestaltung des Dienstes tatsächlich erforderlich ist. So müssen z.B. nichtwirtschaftlich, gemeinnützig oder freiberuflich tätige Diensteanbieter nachweisen, dass die Nutzeranalyse tatsächlich den in § 15 Abs. 3 TMG genannten Zwecken dient und erforderlich ist.

Piwik bietet die Möglichkeit, die Verarbeitung der vom Browser der Nutzerinnen und Nutzer des Web-Angebots übermittelten Referrer-Daten nicht durchzuführen.

Hierzu ist das Referrer-Plugin zu deaktivieren (Einstellungen-> Plugins, Plugin „Referer“(sic)).

Das ULD empfiehlt, zu prüfen, ob für die geplante Web-Analyse die Nutzung des Referrers zwingend erforderlich ist, und ggfs. das Referrer-Plugin zu deaktivieren.

3.4 Verbot der Zusammenführung durchsetzen

Die bei der Analyse erhobenen Daten dürfen keinesfalls mit Daten über den Träger des Pseudonyms zusammengeführt werden, §§ 13 Abs. 4 Nr. 6, 15 Abs. 3 TMG. Dies kann beispielsweise durch andere Programme oder Dienste erfolgen, die eine Verknüpfung zu anderweitig erhobenen Daten, z.B. aus einem Bestellvorgang, ermöglichen. Durch geeignete technische und organisatorische Maßnahmen muss die Wahrung des Trennungsprinzips sichergestellt werden. Eine tatsächliche Zusammenführung muss ausgeschlossen sein.

Das ULD empfiehlt, soweit eine technische Trennung der Daten nicht zuverlässig sichergestellt werden kann, zumindest durch entsprechende Arbeitsanweisungen die Verwender sowie die Administratorinnen und Administratoren der Systeme auf dieses Verbot hinzuweisen.

⁹ <http://donottrack.us/> (Aktualität des Links: 15. März 2011).

3.5 Daten löschen

Die erhobenen Daten sind nach Erstellen der Statistik oder jederzeit auf Verlangen des betroffenen Nutzers zu löschen oder, soweit eine Löschung wegen der besonderen Art der Speicherung nicht oder nur mit unverhältnismäßigem Aufwand möglich ist, zu sperren.

Derzeit gibt es nach Angaben von Piwik auf den eigenen Webseiten kein Tool zum Löschen aller Analysedaten einer bestimmten Webseite. Auch ist bislang kein Tool zum Löschen einzelner Datensätze vorhanden. Die Daten müssen direkt aus der Piwik Datenbank gelöscht werden.

Sollen alle Analysedaten für eine Webseite gelöscht werden, sind nach Angaben des Piwik-Teams¹⁰ folgende Schritte auszuführen:

1. Löschen aller MySQL-Tabellen mit einem Namen, der mit `piwik_archive_*` beginnt; diese werden automatisch neu erstellt, wenn Piwik Berichte berechnet.
2. Ausführen der folgenden Anfrage: `DELETE FROM piwik_log_visit WHERE idsite = X`. Hierbei ist X durch die Website-ID der Seite zu ersetzen, für die der Reset ausgeführt werden soll.

Das ULD empfiehlt, durch unmittelbare Aufforderung durch Nutzerinnen und Nutzer die erforderliche Löschung und eine regelmäßige anlasslose Löschung direkt in der Datenbank durch auf den jeweiligen Anwendungsfall optimierte SQL-Anweisungen umzusetzen.

3.6 Lebensdauer von Piwik-Cookies setzen und verringern

Zur datenschutzkonformen Verwendung von Cookies gehört die angemessene Begrenzung ihrer Lebensdauer.

Für Verwender von Piwik ist es möglich, die Lebensdauer der von Piwik gesetzten Cookies zu beeinflussen. Ist die durch Piwik vorkonfigurierte Lebensdauer nicht erforderlich, muss diese durch den Verwender entsprechend angepasst werden.

Das Verfahren ist ähnlich der oben beschriebenen Einstellung der Zahl der zu anonymisierenden Oktette bei IP-Adressen. Die Standardlebensdauer des Authentifizierungs-Cookies (dieses wird gesetzt, wenn sich der Piwik-Verwender auf seinen Statistikseiten einwählt) beträgt 30 Tage. Die voreingestellte Lebensdauer des Tracking-Cookies für die Webseiten-Besucher (dieses wird bei dem Besucher der beobachteten Webseite gesetzt) beträgt zwei Jahre. Diese Default-Werte sind in der Datei `global.ini.php` im Ordner `config`, der im Haupt-Piwikordner liegt, festgelegt. Das sieht im Code, mit `piwik_auth` als Authentifizierungs-Cookie und `piwik_visitor` als Tracking-Cookie, folgendermaßen aus:

```
[General]
; login cookie name
login_cookie_name = piwik_auth
```

¹⁰ Quelle: http://de.piwik.org/piwik-haufig-gestellte-fragen/wie-mache-ich/#faq_73 (Aktualität des Links: 15. März 2011).

```
; login cookie expiration (30 days)
login_cookie_expire = 2592000
```

```
[Tracker]
```

```
; name of the cookie used to store the visitor information
cookie_name = piwik_visitor
```

```
; by default, the Piwik tracking cookie expires in 2 years
cookie_expire = 63072000
```

Will der Verwender diese Werte verändern, sollten diese nicht in der Datei global.ini.php editiert werden, sondern in der Datei config.ini.php. Dies geschieht durch das Einfügen entsprechender Zeilen in den Rubriken [General] und [Tracker]. Die Lebensdauer wird hierbei in Sekunden angegeben.

Um die Lebensdauer der zu setzenden Tracking-Cookies zum Beispiel von 2 Jahren auf 2 Stunden herabzusetzen, werden die folgenden Zeilen eingefügt:

```
[Tracker]
cookie_expire = 7200
```

Will man zum Beispiel die Lebensdauer der zu setzenden Authentifizierungs-Cookies von 30 Tagen auf zwei Wochen ändern, fügt man die folgenden Zeilen ein:

```
[General]
login_cookie_expire = 1209600
```

Hat der Verwender die Rubriken [Tracker] oder [General] schon bei einer früheren Konfiguration eingefügt (z.B. beim Setzen der gewünschten Oktettzahl für die Anonymisierung von IP-Adressen), müssen einfach nur die Zeilen mit den neuen Ablaufdaten in dieser Rubrik eingefügt werden. Die Rubrik [Tracker] könnte nun beispielsweise so aussehen:

```
[Tracker]
ip_address_mask_length = 2
cookie_expire = 7200
```

Es ist zu beachten, dass die in den Beispielen aufgeführten Zahlenwerte reine Beispielswerte sind und keinerlei Empfehlung beinhalten.

Die Lebensdauer des Tracking-Cookies wird bei jedem Besuch aktualisiert. Die Lebensdauer des Tracking-Cookies bezieht sich somit nicht direkt auf die maximale Dauer des Trackings nach einem Erstbesuch, sondern gibt lediglich an, für welchen Zeitraum Nutzer nach einem Besuch wiedererkannt werden.

Das ULD empfiehlt, die Lebensdauer der Tracking-Cookies so kurz wie möglich zu konfigurieren und insbesondere eine Lebensdauer von einer Woche nicht zu überschreiten.

4. Rechtliche Hinweise zum Einsatz von Piwik

4.1 Cookies

Der Einsatz von Cookies ist nur zulässig, wenn alle relevanten datenschutzrechtlichen Vorgaben eingehalten werden. Diesbezüglich besteht derzeit eine europarechtlich angestoßene Umbruchphase. Artikel 5 Absatz 3 der Datenschutzrichtlinie für elektronische Kommunikation (RL 2002/58/EG) wurde im Dezember 2009 durch das sogenannte Telekom-Reformpaket zur Reform der europäischen Telekommunikationsmärkte geändert. Nach der bislang geltenden Fassung war die Verwendung von Cookies rechtlich zulässig, wenn folgende Kriterien erfüllt waren:

- umfassende Information der Nutzer in der Datenschutzerklärung über den Dienst und ihr Recht, der Web-Analyse zu widersprechen,
- wirksame Umsetzung eingelegter Widersprüche (z.B. mittels eines Widerspruchs-Cookies; eine sogenannte Opt-Out-Lösung war zulässig),
- angemessene Begrenzung der Lebensdauer von Cookies.

Die neue Fassung geht über diese Ansprüche hinaus. Nach der neuen Version ist die Verwendung von Web-Analyse-Cookies nur zulässig, wenn der Nutzer vor dem erstmaligen Setzen eines Cookies alle relevanten Informationen erhält und auf dieser Grundlage seine Einwilligung erteilt (sog. Opt-In). Der späteste Zeitpunkt für eine Umsetzung in nationales Recht ist der 25. Mai 2011.

Für die Übergangszeit hat die Artikel 29-Datenschutzgruppe eine „Diskussionsperiode“ ausgerufen, da sich noch keine „Best Practices“ im Zusammenhang mit Opt-In-Lösungen herausgebildet hat. Ob und in welcher Form die Vorgaben der Richtlinie in deutsches Recht umgesetzt werden, ist derzeit noch offen (Stand: 15. März 2011).

Eine Möglichkeit, in der Zwischenzeit bereits so gut wie möglich den Anforderungen der neuen Fassung der Richtlinie gerecht zu werden, ist zunächst die Information der Nutzerinnen und Nutzer in prominenter Weise zu gewährleisten. Hierbei genügt die Bereitstellung in der Datenschutzerklärung gemäß der neuen Richtlinienfassung nicht mehr. Es kann beispielsweise ein Pop-Up-Fenster oder Ähnliches Verwendung finden. Es ist möglich, die Informationen im Sinne eines Stufenmodells bereitzustellen, den Nutzenden also zunächst knappe Informationen zu geben und die Möglichkeit zu bieten, erweiterte Informationen anzuwählen. Auch die Nutzung von (zu standardisierenden) Informationsgrafiken wäre denkbar.

Das ULD hält eine Beachtung der derzeit geltenden nationalen datenschutzrechtlichen Vorgaben bei dem Einsatz von Piwik für realisierbar. Die neuen europarechtlichen Vorgaben zum Opt-In sind mit der aktuellen Version von Piwik jedoch nicht umsetzbar. Die Ergebnisse der Diskussionsperiode zu „Opt-In“-Lösungen und deren Umsetzung durch den Gesetzgeber bleibt abzuwarten. Dies könnte eine Änderung des anzuwendenden § 15 Abs. 3 TMG erforderlich machen.

4.2 Beschluss zur datenschutzkonformen Ausgestaltung von Analyseverfahren zur Reichweitenmessung bei Internet-Angeboten

Werden Nutzerinnen und Nutzer beim Einsatz von Diensten zur Reichweitenanalyse individualisiert und Profile über ihr Nutzungsverhalten erstellt, liegt in der Regel eine Erhebung, Verarbeitung und Nutzung personenbezogener Daten vor. Eine Identifikation Einzelner, z.B. durch Klarnamen, ist nicht Voraussetzung für die Anwendung des Datenschutzrechts. Die Zulässigkeit der Erhebung, Verarbeitung und Nutzung personenbezogener Daten durch Telemediendiensteanbieter wird durch das Telemediengesetz geregelt.

Gemäß § 15 Abs. 3 TMG dürfen Diensteanbieter für Zwecke der Werbung, der Marktforschung oder zur bedarfsgerechten Gestaltung der Telemedien Nutzungsprofile bei Verwendung von Pseudonymen erstellen. Den Nutzern ist jedoch ein Widerspruchsrecht einzuräumen und der Diensteanbieter hat auf das Widerspruchsrecht hinzuweisen. Die erstellten Nutzungsprofile dürfen nicht mit Daten über den Träger des Pseudonyms zusammengeführt werden.

In ihrer Sitzung vom 26. und 27. November 2009 haben die obersten Aufsichtsbehörden für den Datenschutz im nicht-öffentlichen Bereich einen Beschluss¹¹ zu datenschutzkonformen Analyseverfahren zur Reichweitenmessung bei Internet-Angeboten gefasst, der als Interpretationshilfe für Diensteanbieter dienen soll.

Folgende Vorgaben werden diesbezüglich gemacht:

Den Betroffenen ist eine Möglichkeit zum Widerspruch gegen die Erstellung von Nutzungsprofilen einzuräumen. Derartige Widersprüche sind wirksam umzusetzen.

Diese Vorgabe ist durch Piwik umsetzbar, vgl. Abschnitt 3.2.

Die pseudonymisierten Nutzungsdaten dürfen nicht mit Daten über den Träger des Pseudonyms zusammengeführt werden.

Dies kann zurzeit nur in Verbindung mit organisatorischen Maßnahmen erfolgen.

Sie [die pseudonymisierten Nutzungsdaten] müssen gelöscht werden, wenn ihre Speicherung für die Erstellung der Nutzungsanalyse nicht mehr erforderlich ist oder der Nutzer dies verlangt.

Piwik selbst kann diese Löschung nicht umsetzen, hier muss die verantwortliche Stelle durch SQL-Skripte eine Löschung direkt in der Datenbank herbeiführen. Im Einzelfall kann auch eine Sperrung der Daten zulässig sein. Der Diensteanbieter muss dies jedoch plausibel darlegen. Wirksam anonymisierte Nutzungsstatistiken, die aus individuellen Nutzungsprofilen erstellt werden, unterfallen nicht mehr dem Datenschutzrecht.

Auf die Erstellung von pseudonymen Nutzungsprofilen und die Möglichkeit zum Widerspruch müssen die Anbieter in deutlicher Form im Rahmen der Datenschutzerklärung auf ihrer Internetseite hinweisen.

Dies muss auf der Webseite selbst gelöst werden. Das „Verstecken“ des Hinweises in unübersichtlichen Datenschutzhinweisen oder Nutzungsbedingungen erfüllt diese Anforderungen nicht.

Personenbezogene Daten eines Nutzers dürfen ohne Einwilligung nur erhoben und verwendet werden, soweit dies erforderlich ist, um die Inanspruchnahme von Telemedien zu ermöglichen

¹¹ <http://www.datenschutz-mv.de/dschutz/beschlue/Analyse.pdf> (Aktualität des Links: 15. März 2011).

und abzurechnen. Jede darüber hinausgehende Nutzung bedarf der Einwilligung der Betroffenen.

Die Analyse des Nutzungsverhaltens unter Verwendung vollständiger IP-Adressen (einschließlich einer Geolokalisierung) ist aufgrund der Personenbeziehbarkeit dieser Daten daher nur mit bewusster, eindeutiger Einwilligung zulässig. Liegt eine solche Einwilligung nicht vor, ist die IP-Adresse vor jeglicher Auswertung so zu kürzen, dass eine Personenbeziehbarkeit ausgeschlossen ist.

Piwik verarbeitet bei Nutzung des Anonymisierungs-Plugins keine IP-Adressen der Nutzerinnen und Nutzer eines Internet-Angebots.

5. Fazit

Die Software Piwik ist aus Sicht des ULD bei datenschutzkonformer Implementierung eine mögliche Alternative zur datenschutzrechtlich unzulässigen Anwendung anderer Analysedienste.

Beim Einsatz ist darauf zu achten, dass durch zusätzliche Plugins, eine datensparsame Grundkonfiguration und zusätzliche organisatorische Maßnahmen das Erheben und Verarbeiten von personenbezogenen Daten minimalisiert wird.

Die im Kapitel 3 genannten Maßnahmen sind Grundvoraussetzung für einen datenschutzkonformen Einsatz von Piwik.