

Bericht über Energiedaten im Internet

Wintersemester 2013/14

Andreas Krause

Januar 2014

1 Einleitung

Zur Verarbeitung von Energiedaten ist es erforderlich diese von einem Anbieter zu bekommen. Dank des Erneuerbare-Energien-Gesetz sind Netzbetreiber und Übertragungsnetzbetreiber verpflichtet, Daten von Anlagen zu veröffentlichen, die Strom aus so genannten erneuerbaren Energien erzeugen. Erneuerbare Energien haben den Ruf, besonders umweltfreundlich zu sein, da sie meist Strom aus “unerschöpflichen” Quellen gewinnen, wie zum Beispiel Sonnenergie oder Windkraft.

Die veröffentlichten Informationen dieser Anlagen bestehen aus Anlagestammdaten und Abrechnungs- bzw. Bewegungsdaten. In Kapitel 2 wird auf die Anlagestammdaten eingegangen, wobei die Gemeinsamkeiten und Unterschiede der Daten aus verschiedenen Quellen untersucht werden. Die Abrechnungsdaten werden anschließend in Kapitel 3 behandelt. Es folgt eine kurze Zusammenfassung in Kapitel 4.

2 Anlagestammdaten

2.1 Quellen der Anlagedaten

Nach dem Erneuerbare-Energien-Gesetz müssen Netzbetreiber Standort- und Leistungsinformationen der Anlagen veröffentlichen. Diese Daten werden auch Stammdaten genannt. Weiterhin bieten auch die Übertragungsnetzbetreiber diese Informationen an. Im Falle der Anlagen im Leipziger Stadtgebiet kommen also die hiesigen Netzbetreiber sowie der für diese Region zuständige Übertragungsnetzbetreiber zur Beschaffung der Stammdaten in Frage.

In Leipzig gibt es zwei Netzbetreiber, die Netz Leipzig GmbH [9] sowie die Mitteldeutsche Netzgesellschaft Strom mbH [6]. Beide Unternehmen bieten die Stammdaten als Download im PDF-Format an [11, 8]. Von der Homepage des Übertragungsnetzbetreibers 50Hertz Transmission GmbH kann man sich die Stammdaten der Anlagen im CSV-Format herunterladen [1], wobei es einige Filteroptionen gibt. Bei dem von uns verwendeten Datensatz wurden nur Anlagen aus Sachsen verwendet. Eine weitere Quelle der Stammdaten ist die Informationsplattform der deutschen Übertragungsnetzbetreiber, <http://www.eeg-kwk.net>. Hier findet man zahlreiche Informationen zum Erneuerbare-Energien-Gesetz und zur Kraft-Wärme-Kopplung. Es werden ebenfalls die Anlagestammdaten aller Übertragungsnetzbetreiber in Deutschland im CSV-Format zum Download angeboten [5].

Die in unserem Energieprojekt genutzten Daten der Anlagen stammen aus dem Jahr 2012 und wurden Mitte 2013 veröffentlicht. Lediglich die Daten, die von 50Hertz stammen, enthalten auch Anlagen aus dem Jahr 2013, da die Daten ständig aktualisiert werden. Die im Projekt verwendete Version des 50Hertz-Datensatzes stammt vom November 2013. In der nachfolgenden Tabelle ist eine Übersicht der Veröffentlichungsdaten und Aktualität der einzelnen Datensätze zu finden.

Quelle	Veröffentlichungsdatum	Aktualität der Daten
50Hertz	November 2013	November 2013
eeg-kwk.net	31.05.2013	2012
MitNetz Strom mbH	03.06.2013	2012
Netz Leipzig GmbH	08.10.2013	2012

Tabelle 1: Aktualität der Anlagestammdaten

2.2 Quantität der Anlagedaten

Die Menge der Anlagedaten der verschiedenen Quellen variiert stark. Das liegt daran, dass die Datensätze der Netzbetreiber nur Anlagen aus ihren Netzen enthalten und folglich zueinander disjunkt sind. Die Daten von 50Hertz hingegen enthalten alle Anlagen der Regelzone von 50Hertz und somit auch die Anlagen der Netzbetreiber. Ebenso verhält es sich mit den Anlagestammdaten von eeg-kwk.net. In Tabelle 2 ist eine Auflistung der Anzahl an Anlagen der einzelne Datensätze zu finden.

Quelle	Anlagen insgesamt	Anlagen in Leipzig
50Hertz (nur Anlagen aus Sachsen)	30217	1319 (davon 122 von 2013)
eeg-kwk.net	117598	1217
MitNetz Strom mbH	31190	579
Netz Leipzig GmbH	638	612

Tabelle 2: Anzahl der Anlagedaten verschiedener Quellen

Die in Tabelle 2 ersichtliche Differenz der Anzahl Leipziger Anlagen zwischen den Datensätzen von 50Hertz und eeg-kwk.net kommt zustande, da die Daten von 50Hertz über 100 Anlagen aus dem Jahr 2013 enthalten, die in den Daten von eeg-kwk.net nicht vorkommen, da diese aus dem Jahr 2012 stammen. Trotz der Berücksichtigung dieses Umstandes stimmen die Anzahlen aber trotzdem nicht überein. Auch die Summe der Leipziger Anlagen aus den Daten der Netzbetreiber ist geringer als die in den anderen Datensätzen vorhandene Anzahl Leipziger Anlagen, obwohl es nur diese zwei Netzbetreiber in Leipzig gibt. Insgesamt gibt es 1335 Leipziger Anlagen, die sich in einem der vier Datensätze befinden. Dadurch wird deutlich, dass es in den Datensätzen Anlagen gibt, die in anderen nicht vorhanden sind, obwohl sie es sein müssten. Ein Grund könnte sein, dass Anlagen aus den Datensätzen von 2012 mittlerweile außer Betrieb genommen wurden und somit in dem 50Hertz-Datensatz nicht mehr auftauchen. Eine weitere Möglichkeit besteht darin, dass die einzelnen Datensätze unvollständig sind.

2.3 Qualität der Anlagedaten

Die Stammdaten der verschiedenen Quellen unterscheiden sich nicht nur in der Menge der Anlagen, sondern auch darin, welche Informationen zu den Anlagen enthalten sind. Dabei gibt es Anlageninformationen, die in allen Quellen vorhanden sind und solche, die nur in einzelnen Datensätzen vorkommen. Tabelle 3 gibt einen Überblick über die Attribute der einzelnen Datensätze.

Datensatz	Attribute
Gemeinsame Attribute aller Datensätze	Anlagennummer, Ort, Strasse, Bundesland, Nennleistung, Einspeisungsebene, Energieträger, Datum der Inbetriebnahme
50Hertz.com	Netzbetreiber, Biomasseanlagen: KWK-Anteil, Technologie
eeg-kwk.net	Übertragungsnetzbetreiber, Netzbetreiber, Netzbetreiber Betriebsnummer, leistungsgemessene Anlage, Regelbarkeit, Zeitpunkt des Netzzuganges, Zeitpunkt des Netzabganges, Zeitpunkt der Außerbetriebnahme
MitNetz Strom mbH	Zeitreihentyp
Netz Leipzig GmbH	Zeitpunkt der Außerbetriebnahme

Tabelle 3: Attribute der Stammdaten

Es ist zu erkennen, dass in jedem Datensatz der Standort, die Leistung und die Art der Energieerzeugung der Anlagen enthalten ist. Weiterhin besitzt jeder Datensatz zusätzliche Attribute der Anlagen. So sind beispielsweise zusätzliche Informationen zu Biomasseanlagen beim 50Hertz-Datensatz zu finden, die bei den Datensätzen der Netzbetreiber fehlen. Einige zusätzliche Attribute haben nur sehr selten Werte bei den Anlagen. Ein Beispiel hierzu ist der Zeitpunkt der Außerbetriebnahme, da nahezu alle Anlagen in Betrieb sind.

Bei den Werten der Stammdaten müssen einige Besonderheiten beachtet werden. So fehlt beispielsweise die vorgestellte „0“ bei den Postleitzahlen im eeg-kwk.net Datensatz. Ebenfalls zu beachten ist, dass in den Datensätzen teilweise die Ortsteile beim Ortsnamen mit angegeben sind.

Da Flurstücke bei den Adressdaten angegeben sind, müssen diese eventuell gesondert behandelt werden. In den Anlagen im Stadtgebiet von Leipzig kommen solche Anlagen nicht vor. Wird allerdings das Gebiet auf das Leipziger Umland oder eine andere Region ausgeweitet, so gehören auch Anlagen auf Flurstücken zur relevanten Anlagenmenge.

2.3.1 Datenkonflikte

Es hat sich gezeigt, dass sich die Anlagedaten teilweise widersprechen. Eine automatisierte Analyse ergab 656 Konflikte zwischen den Stammdaten aller Datensätze. Eine Verteilung dieser Konflikte bezüglich der Attribute ist in Tabelle 4 zu sehen.

Den größten Teil der Konflikte machen unterschiedliche Leistungsangaben aus. Hierbei handelt es sich um Abweichungen, die auf unterschiedliche Genauigkeit der Leistungswerte zurückzuführen sind. Diese Konflikte können beispielsweise aufgelöst werden, indem bei der entsprechenden Anlage der Wert mit der höchsten Präzision genommen wird.

Attribut	Anzahl der Konflikte
Leistung	618
Energieträger	18
Zeitpunkt der Inbetriebnahme	10
PLZ	5
Strasse	3
Hausnummer	2

Tabelle 4: Verteilung der Datenkonflikte bezüglich Attribute

Die meisten anderen Konflikte beruhen auf abweichenden Schreibweisen für den gleichen Wert, bspw. „Wind“ und „Windenergie“ als Energieträger, oder Schreibfehler. Bei diesen kann man ebenfalls einen „richtigen Wert“ festlegen.

Lediglich 11 Konflikte können nicht ohne Weiteres aufgelöst werden. Hierbei handelt es sich zum einen um einen falschen Eintrag als Energieträger, bei dem einerseits „Solar“, andererseits „Bio“ als Wert in Frage kommt, sowie die 10 Konflikte bezüglich des Datums der Inbetriebnahme. Da bei diesen Konflikten nicht entscheidbar ist, welcher Wert der richtige ist, müssen diese Daten als fehlerhaft betrachtet werden. In der im Projekt erstellten Ontologie sind dabei beide abweichenden Werte enthalten mit einem Hinweis in Form eines `rdfs:comment`.

3 Abrechnungsdaten

Die Netzbetreiber sind nach dem EEG ebenfalls verpflichtet, Jahresberichte zu veröffentlichen, aus denen hervorgeht, wie viele Anlagen wie viel Strom produziert haben und wie hoch die Vergütung war. Die Netz Leipzig GmbH bietet diesen Bericht unter [10] zum Download als PDF an. Der Bericht der Mitteldeutschen Netzgesellschaft Strom mbH ist auf der Homepage veröffentlicht [7]. Bei diesen Berichten sind allerdings die Daten nicht anlagebezogen, sondern eine Zusammenfassung aller Anlagen der entsprechenden Netzbetreiber. Es lassen sich also keine Informationen zu den einzelnen Anlagen gewinnen.

Die Abrechnungsdaten der einzelnen Anlagen können aber von 50Hertz bezogen werden [2]. Hierbei können die schon von den Anlagestammdaten bekannten Filter eingesetzt werden. Das Resultat sind die Abrechnungsdaten jeder Anlage in einem bestimmten Jahr. In diesen Daten befinden sich die Anlagestammdaten der Anlagen sowie weitere Abrechnungsinformationen wie produzierter EEG-Strom, die Vergütungskategorie, die jeweilige Vergütung und vermiedene Netznutzungsentgelte. Es kann also genau nachvollzogen werden, welche Anlage wie viel Strom in einem Jahr produziert hat. Bei der Vergütung muss beachtet, dass die gelieferte Energiemenge je nach Vergütungskategorie unterschiedlich vergütet wird.

4 Zusammenfassung und Ausblick

Energiedaten von Erneuerbare-Energie-Anlagen können von den jeweiligen Netzbetreibern bzw. vom Übertragungsnetzbetreiber erhalten werden. Diese Daten werden entweder einmal im Jahr für das vorangegangene Jahr veröffentlicht oder ständig aktualisiert (50Hertz). Die gelieferten Informationen über die Anlagen bestehen zum einen aus Anlagestammdaten, die

hauptsächlich aus Standort-, Leistungs- und Energieerzeugungsinformationen bestehen, zum anderen werden auch Abrechnungsdaten veröffentlicht.

Die Anlagestammdaten sind zum Teil fehlerhaft, da sich Informationen aus verschiedenen Quellen zu einigen Anlagen teilweise widersprechen. Es ist auch nicht auszuschließen, dass die Menge der Anlagen bei bestimmten Quellen unvollständig ist. Es gibt also einen gewissen Qualitätsmangel bei den angebotenen Anlagestammdaten.

Die Abrechnungsdaten, die von den Netzbetreibern veröffentlicht werden, sind für alle Anlagen zusammengefasst und deswegen zu allgemein. Laut dem EEG müssen diese aber auch nicht mehr Daten veröffentlichen, weshalb sich das in Zukunft auch nicht ändern wird. Allerdings kann man zu einem gegebenen Geschäftsjahr anlagenbezogene Abrechnungsdaten von 50Hertz beziehen, wodurch man genaue Aussagen über die reale Leistungsfähigkeit einer Anlage bekommt – im Gegensatz zu der in den Anlagestammdaten angegebenen Nennleistung.

Da im Erneuerbare-Energien-Gesetz keine genauen Vorgaben existieren, welche Daten in welcher Form veröffentlicht werden sollen, kann man nicht ausschließen, dass sich Menge und Qualität der Anlagendaten verändert. So könnte es aus datenschutzrechtlichen Gründen beispielsweise dazu kommen, dass die Standortdaten eingeschränkt werden. Dies ist schon jetzt beim Übertragungsnetzbetreiber TransnetBw der Fall, bei dem sich die Standortdaten nur auf Ort und Postleitzahlen beschränken [12]. Änderungen des Erneuerbare-Energien-Gesetz könnten ebenfalls dazu führen, dass sich die Menge der angebotenen Energiedaten ändert. In der Politik ist das Erneuerbare-Energien-Gesetz ein aktuelles Streitthema, sodass dessen Änderung in absehbarer Zeit durchaus denkbar wäre. Es ist daher nicht absehbar, wie sich die Bereitstellung von Energiedaten in Zukunft entwickelt.

Literatur

- [1] Homepage des Übertragungsnetzbetreibers 50Hertz, letzter Zugriff am 30.01.2014. <http://www.50hertz.com>.
- [2] Abrechnungsdaten der Anlagen in der 50hertz Regelzone, letzter Zugriff am 31.01.2014. http://www.50hertz.com/cps/rde/xchg/trm_de/hs.xsl/166.htm.
- [3] CSV-Export der EEG-stammdaten von 50Hertz, letzter Zugriff am 30.01.2014. http://www.50hertz.com/cps/rde/xchg/trm_de/hs.xsl/165.htm.
- [4] Informationsportal der Übertragungsnetzbetreiber, letzter Zugriff am 30.01.2014. <http://www.eeg-kwk.net>.
- [5] Download der Anlagestammdaten vom Informationsportal der Übertragungsnetzbetreiber, letzter Zugriff am 30.01.2014. <http://www.eeg-kwk.net/de/Anlagenstammdaten.htm>.
- [6] Homepage des Netzbetreibers Mitteldeutsche Netzgesellschaft Strom mbH, letzter Zugriff am 30.01.2014. <https://www.mitnetz-strom.de>.
- [7] Jahresbericht des Jahres 2012 des Netzbetreibers Mitteldeutsche Netzgesellschaft Strom mbH, letzter Zugriff am 31.01.2014. <https://www.mitnetz-strom.de/Unternehmen/ZahlenFakten/ErneuerbareEnergien>.

- [8] Anlagestammdaten des Jahres 2012 des Netzbetreibers Mitteldeutsche Netzgesellschaft Strom mbH, letzter Zugriff am 30.01.2014. https://www.mitnetz-strom.de/irj/go/km/docs/z_ep_em_unt_documents/em/mitnetzstrom/Dokumente/Stammdaten_EEG-Jahresmeldung_2012_MITNETZ%20STROM.pdf.
- [9] Homepage des Netzbetreibers Netz Leipzig GmbH, letzter Zugriff am 30.01.2014. <http://www.netz-leipzig.de>.
- [10] Jahresbericht des Jahres 2012 des Netzbetreibers Netz Leipzig GmbH, letzter Zugriff am 31.01.2014. <http://www.netz-leipzig.de/netzanschluesse/erzeugungsanlagen/#c81>.
- [11] Anlagestammdaten des Jahres 2012 des Netzbetreibers Netz Leipzig GmbH, letzter Zugriff am 30.01.2014. http://www.netz-leipzig.de/netzanschluesse/erzeugungsanlagen/?eID=dam_frontend_push&docID=358.
- [12] Auswertung der EEG-Anlagenstammdaten für Photovoltaik-Anlagen: keine Verbesserung in Sicht, letzter Zugriff am 31.01.2014. <http://hsi-solar-invest.com/?p=5364>.