

Linked Open Data im Tourismus - Ordnung ins Buzzword-Chaos bringen

von Prof. Dr. Eric Horster – Fachhochschule Westküste in Heide

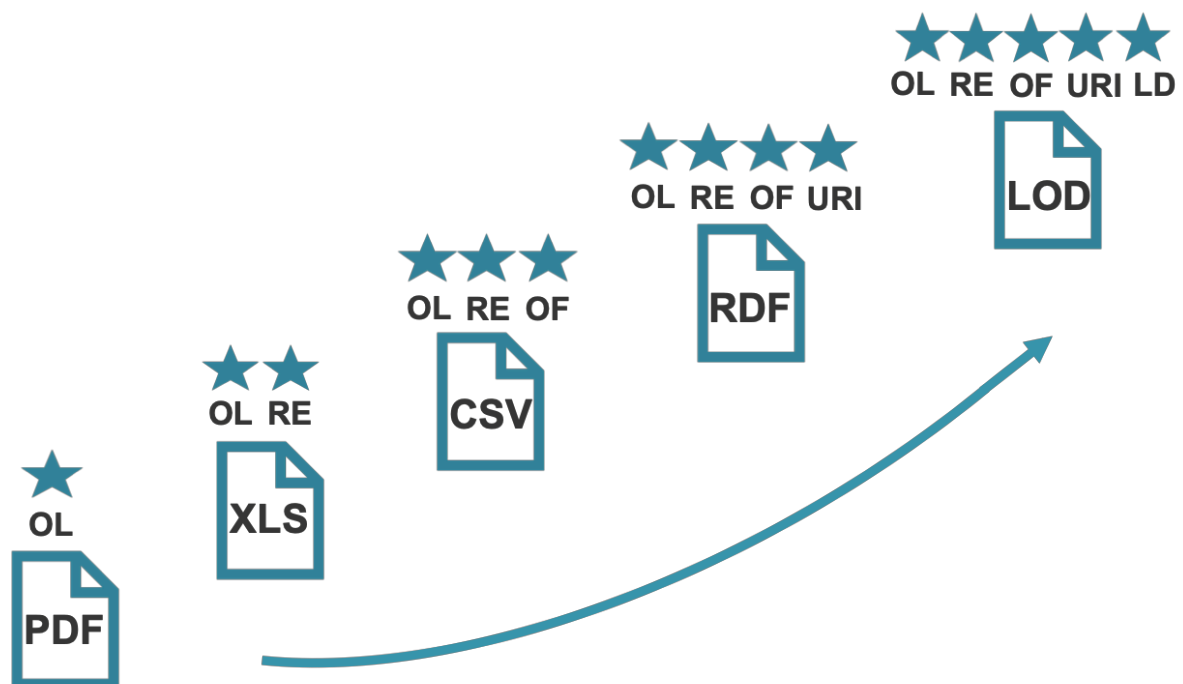
Das Thema Open Data wird im Tourismus aktuell breit diskutiert. Gleichwohl herrscht ein mehr oder weniger großes Durcheinander zu den Begrifflichkeiten, die oft als Schlagworte verwendet werden, ohne dass immer die Bedeutung hinter diesen von denjenigen, die sie verwenden, verstanden wird. In diesem Kurzbeitrag soll nun einmal der Überbegriff bei dieser Debatte geklärt werden: Linked Open Data.

Linked Data vs. Open Data

Zunächst ist es wichtig zu verstehen, dass mit Open Data und Linked Data zwei unterschiedliche Dinge gemeint sind, die jedoch in der Debatte oftmals synonym verwendet werden. Mit Open Data wird die Öffnung von Daten für jeden und für jegliche Zwecke bezeichnet. Um die unglaubliche Datenmenge der bereitgestellten Informationen für Mensch und Maschine durchsuchbar zu machen, sollte man diese Daten zudem Strukturieren, wie nachfolgend noch erläutert wird. Der Begriff Linked Data bezeichnet hingegen eine Art, die Daten im World Wide Web abzulegen, damit diese mittels Links in Beziehung zueinander gesetzt werden können. Dies ist ein wichtiger Unterschied, denn nur weil die Daten geöffnet sind, muss der Ort an dem Sie gespeichert sind sowie Ihre Beziehung zueinander nicht zwingend beschrieben worden sein. Andersherum müssen Daten, die miteinander verlinkt sind, nicht per se offen sein.

Das fünf Sterne-Modell

Tim Berners-Lee (der Erfinder des World Wide Web) hat für die Bereitstellung von Open Data ein fünf Stufen bzw. Sterne-Modell entwickelt, welches sehr anschaulich verdeutlicht, wie Linked Open Data erfolgreich sein kann.



Quelle: Eigene Darstellung nach <https://5stardata.info/de>

Diese Sterne sind wie folgt definiert¹:

★ Stellen Sie Daten im Web unter einer offenen Lizenz bereit, das Datenformat für die Bereitstellung ist Ihnen überlassen

★★ Stellen Sie Daten in einem strukturierten Format bereit

★★★ Verwenden Sie offene, nicht proprietäre Formate

★★★★ Verwenden Sie URIs, um Dinge zu bezeichnen, damit Daten verlinkt werden können

★★★★★ Verlinken Sie Ihre Daten mit anderen Daten, um Kontexte herzustellen

Es wird ersichtlich, dass sich die ersten drei Stufen primär auf die Öffnung der Daten beziehen und bei den höheren Stufen der Fokus auf der Verlinkung liegt. Es erfolgt hier somit eine Strukturierung der Veröffentlichung von Open Data mittels Linked Data, was vermutlich ein Grund für die Verwirrung hinsichtlich der Begrifflichkeiten ist. Wichtig ist es daher zu verdeutlichen, dass diese Begriffe nicht im Widerspruch zueinander stehen, sondern sich ergänzen.

Konkret wird in den ersten drei Stufen definiert, wie Daten aufbereitet werden sollten, damit diese zum einen weiterverwendet werden *dürfen* und zum anderen leicht weiterverwendet werden *können*. Auf der ersten Stufe wird eine offene Lizenzierung gefordert (bspw. mittels des Lizenzmodells der Creative Commons). Auf der zweiten wird die Struktur der Daten betont, die über ein bestimmtes Vokabular erfolgen kann, um die Daten zu beschreiben bzw. auszuzeichnen. Bekannt und im Tourismus mittlerweile verbreitet ist hierbei Schema.org. Das Ausgabeformat wird in der dritten Stufe adressiert. Im Kontext von Linked Open Data für maschinelle Verarbeitung kann hier JSON-LD genannt werden. Mit JSON-LD lassen sich Daten so darstellen, dass sie von Maschinen einfach verarbeitet werden können. Dies ermöglicht es Menschen, die Daten je nach Verwendungszweck dynamisch aufzubereiten. Über dieses Format kann auch die Bedingung der vierten Stufe realisiert werden, da Daten innerhalb von JSON-LD mit sogenannten URIs (ähnlich einer URL, also einer Webadresse) eindeutig bezeichnet werden. In der obersten Stufe (fünf Sterne) sollen Daten dann miteinander verlinkt werden, um Beziehungen und Kontexte herzustellen. Aber warum das ganze überhaupt?

Die Entwicklung des Web

Es stellt sich also die Frage, weshalb Daten miteinander in Beziehung gebracht werden sollten. Um dies nachzuvollziehen ist es zunächst wichtig zu verstehen, wie sich das Web voraussichtlich weiterentwickeln wird, bzw. sich aktuell schon entwickelt. Große Plattformen wie bspw. Google sowie insbesondere auch Entwicklungen im Bereich der sogenannten künstlichen Intelligenz nehmen Einfluss auf die Struktur des Web. Auch unabhängige Instanzen wie bspw. das W3C-Konsortium um Tim Berners-Lee sehen und stützen diese Entwicklung.

Aktuell überwiegt im Tourismus noch die Denkweise, dass die Website der zentrale Ausgabekanal von Inhalten ist und bleiben wird. Diese Inhalte können dann von Website zu Website verlinkt werden. Das Problem bei dieser Struktur ist jedoch, dass die Möglichkeiten sehr beschränkt sind, um diese Inhalte maschinell zu verarbeiten und somit in unterschiedlicher Form auszuspielen. Bei Suchmaschinen wie Google sieht man jedoch, dass

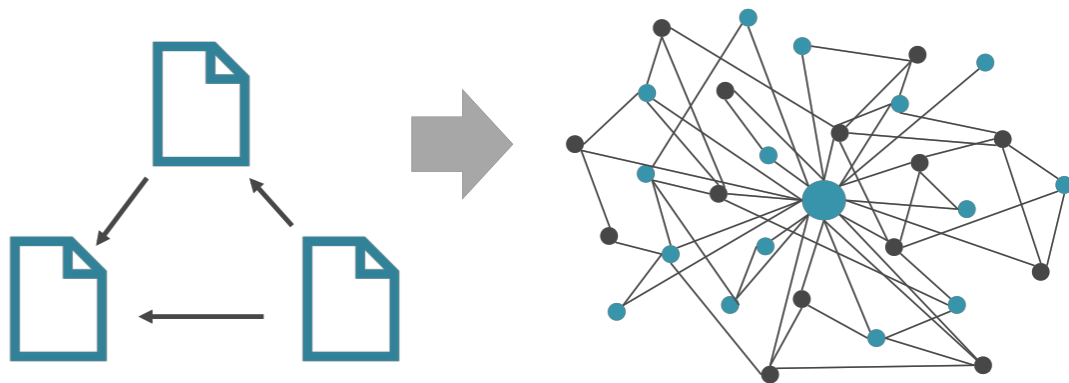
¹ Eine genaue Beschreibung der einzelnen Stufen/Sterne findet sich unter: <https://5stardata.info/de>

bei Suchanfragen versucht wird dem Nutzer die Antwort direkt zu geben – ohne den Umweg über eine Website. Wetter, Flugdaten oder Events werden oftmals unmittelbar oben auf der Suchergebnisseite angezeigt. Die Ergebnisse werden zunehmend auch von einem Sprachassistenten wie Alexa, Siri, Google, Cortana oder Bixby in auditive Rückmeldungen umgewandelt. Die Ausgabegeräte selbst werden ebenfalls immer mannigfaltiger und reichen von großen Screens in der Touristeninformation über klassische Monitore und Laptops bis hin zu mobilen Devices wie Tablets oder Smartphones. In der Konsequenz ist absehbar, dass künftig der Datenfluss wichtiger sein wird als der Datenkanal:

„Der Datenfluss ist wichtiger als der Datenkanal. Denn für die Nutzer ist es wichtig, dass sie die Informationen bekommen – und nicht von wem oder über welchen Kanal.“

Quelle: Sommer (2018): Herausforderungen und Chancen einer offenen, digitalen Dateninfrastruktur im Tourismus.

Aufgrund dieser Veränderungen müssen Daten künftig anders vorgehalten werden. Beispiele, wie diese Datenstruktur aussehen kann liefern Projekte wie die „Linked Open Data Cloud“ oder „Wikidata“. Hier sind nicht mehr Website-Dokumente über Hyperlinks miteinander vernetzt, sondern Inhalte werden aufgebrochen und stehen als einzelne Datenpunkte zur Verfügung und deren Beziehung zueinander wird beschrieben .



Quelle: Eigene Darstellung in Anlehnung an De Keyzer et al. (2013): Einführung in Linked Data.

Die Daten werden unabhängig vom Ausgabekanal gepflegt und sind damit interoperabel einsetzbar. Dies ermöglicht es, dass sie in verschiedenen Kontexten und auf unterschiedliche Art und Weise ausgegeben werden können.

Zu diesen grundsätzlichen Begriffen paaren sich weitere mehr, die in diesem kurzen Beitrag nicht angesprochen werden (konnten). Dies erschwert es, die Thematik für diejenigen, die es umsetzen sollen, greifbar zu machen. Gleichwohl sollte die Relevanz von Linked Open Data deutlich geworden sein. Es besteht sowohl auf Seiten der Betriebe, welche Ihre Angebote und Daten einem möglichst großen Personenkreis unterbreiten wollen, als auch für Destinationsmanagementorganisationen akuter Handlungsbedarf. Für Betriebe ergibt sich aktuell die Chance mit überschaubarem Aufwand und durch die Auszeichnung der eigenen Daten nach Schema.org mehr Sichtbarkeit bei Suchmaschinen zu erhalten. Für die Destinationsmanagementorganisationen kann die Auseinandersetzung mit dem Thema zu einer in die Zukunft gerichteten Aufgabe (das Datenmanagement) führen, welche ihre Daseinsberechtigung neu definiert.