

# GovTech Netzwerk: Eine netzwerkbasierte Untersuchung des GovTech-Ökosystems in Deutschland

Robin Dörnemann

## Zusammenfassung

Analyse und Kartierung des GovTech-Ökosystems in Deutschland: Eine netzwerkbasierte Untersuchung der Verbindungen Unternehmen und Gründern im Bereich der Verwaltungstechnologien.

## Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Definition von GovTech . . . . .	1
1.2	Relevanz des Themas . . . . .	2
1.3	Forschungsziel . . . . .	2
<b>2</b>	<b>Hintergrund</b>	<b>3</b>
2.1	GovTech-Landschaft in Deutschland . . . . .	3
2.2	Netzwerktheorie und ihre Anwendung auf Unternehmensökosysteme . . . . .	3
2.3	LinkedIn als Datenquelle . . . . .	4
<b>3</b>	<b>Methodik</b>	<b>4</b>
3.1	Theoretische Fundierung der Datenerhebung . . . . .	4
3.2	Datenerhebung mittels Web Scraping . . . . .	4
3.2.1	Techstack und Prozessablauf . . . . .	4
3.2.2	Scraper-Architektur und -Prozesse . . . . .	5
3.2.3	Datenbankstruktur . . . . .	5
3.3	Netzwerkanalyse . . . . .	6
3.4	Visualisierung der Daten . . . . .	7
<b>4</b>	<b>Ergebnisse und Diskussion</b>	<b>8</b>
4.1	Darstellung des GovTech-Netzwerks . . . . .	8
4.2	Identifikation von Schlüsselakteuren und Clustern . . . . .	8
4.3	Interpretation der Netzwerkstruktur . . . . .	8
<b>5</b>	<b>Schlussfolgerungen und Ausblick</b>	<b>8</b>

## 1 Einleitung

### 1.1 Definition von GovTech

Um eine fundierte Analyse des GovTech-Ökosystems durchführen zu können, ist es zunächst erforderlich, den Begriff "GovTech" präzise zu definieren und abzugrenzen. In der wissenschaftlichen Literatur und der Praxis existieren verschiedene Ansätze zur Begriffsbestimmung, die sich

in Nuancen unterscheiden. Für die vorliegende Arbeit wird die Definition von Niehaves et al. als Ausgangspunkt gewählt:

«**GovTech** als Begriff beschreibt die Zusammenarbeit von Technologie-orientierten Start-ups (<10 Jahre) und dem öffentlichen Sektor, vor allem Verwaltungen.»  
(Niehaves, Bauer und Klassen 2024)

Diese Definition bietet einen klaren Rahmen für die Untersuchung, bedarf jedoch einer näheren Betrachtung und Kontextualisierung. Zunächst ist hervorzuheben, dass GovTech explizit auf die Kooperation zwischen jungen, innovativen Unternehmen und öffentlichen Einrichtungen abzielt. Dies grenzt GovTech von traditionellen IT-Dienstleistern für den öffentlichen Sektor ab und unterstreicht den innovativen Charakter des Konzepts. Die Altersgrenze von zehn Jahren für Start-ups ist ein wichtiges Kriterium, das die Dynamik und Agilität des GovTech-Sektors betont. Es reflektiert die Tatsache, dass dieser Bereich noch relativ jung ist und sich in einer Phase rasanter Entwicklung befindet.

## 1.2 Relevanz des Themas

Die Digitalisierung der öffentlichen Verwaltung stellt eine zentrale Herausforderung für den modernen Staat dar. Um dieser zu begegnen, hat sich in Deutschland ein GovTech-Ökosystem entwickelt – ein Netzwerk von Akteuren, die innovative Lösungen für den öffentlichen Sektor entwickeln und implementieren.

Die Entwicklung von Lösungen für die digitale Transformation im öffentlichen Sektor weist im Vergleich zur Privatwirtschaft einige Besonderheiten auf. (Silve und Moszoro 2023) Während die öffentliche Verwaltung oft durch langsamere Prozesse aufgrund komplexer Zuständigkeiten und Fachverfahren gekennzeichnet ist (Bogumil, Grohs und Kuhlmann 2019), stehen GovTechs darüber hinaus vor zusätzlichen Herausforderungen. Mit einem maximalen Alter von etwa 10 Jahren steht die GovTech-Akteure im Gegenwind zur traditionellen Trägheit des öffentlichen Sektors. Gleichzeitig besteht ein massiver Bedarf an digitalen Lösungen auf verschiedenen Verwaltungsebenen, von kommunaler bis zur Bundesebene. Die oft nicht klar abgrenzbaren Domänen der Akteure erschwert zudem eine präzise Zuordnung derer GovTech-Lösungen. (Silve und Moszoro 2023) Diese Faktoren machen eine strukturelle Analyse der inhaltlichen Beziehungen zwischen den einzelnen Akteuren im GovTech-Ökosystem besonders herausfordernd. Dennoch bietet gerade diese Komplexität ein interessantes Forschungsfeld, um die Dynamik und Entwicklung dieses innovativen Sektors zu verstehen und dessen Potenzial für die Modernisierung der öffentlichen Verwaltung zu erfassen. (Nathalie Haug und Mergel 2024)

Die Identifikation zentraler Akteure ist entscheidend für das Erkennen von Innovationstreibern, Brückenbauern und struktureller Lücken in diesem Netzwerk. Die Analyse der Netzwerkstruktur kann zudem Aufschluss über die Stabilität und Resilienz des Ökosystems geben.

## 1.3 Forschungsziel

Die vorliegende Arbeit widmet sich eben jener Struktur des GovTech-Ökosystems in Deutschland, einem dynamischen und zunehmend wichtigen Bereich an der Schnittstelle von Technologie und öffentlicher Verwaltung. Im Zentrum der Analyse steht die Frage nach den Beziehungen und Funktionsweisen dieses komplexen Netzwerks. Durch die Anwendung von Methoden der sozialen Netzwerkanalyse auf Daten, die auf der beruflichen Plattform LinkedIn gesammelt wurden, zielt die Studie darauf ab, ein tiefgreifendes Verständnis der Beziehungen und Interaktionen zwischen den verschiedenen Akteuren im GovTech-Bereich zu erlangen. Daraus lässt sich folgende Forschungsfrage ableiten:

«Wie ist das GovTech-Netzwerk in Deutschland strukturiert, welche Akteure nehmen zentrale Positionen ein, welche Muster der Zusammenarbeit und des Wissensaustauschs lassen sich zwischen den verschiedenen Beteiligten erkennen, und welche strukturellen Lücken sind im Netzwerk

identifizierbar?»

## 2 Hintergrund

### 2.1 GovTech-Landschaft in Deutschland

Die GovTech-Landschaft in Deutschland hat in den letzten Jahren eine bemerkenswerte Entwicklung durchlaufen. GovTech, kurz für Government Technology, umfasst Technologien, digitale Produkte und Lösungen, die speziell für den öffentlichen Sektor entwickelt werden. Diese Branche gewinnt zunehmend an Bedeutung, da sie innovative Ansätze zur Modernisierung und Digitalisierung der öffentlichen Verwaltung bietet.

Ein zentrales Element der deutschen GovTech-Szene ist der jährlich stattfindende GovTech-Gipfel, der vom Handelsblatt und Possible (ehemals PUBLIC Deutschland) initiiert wurde. Diese Veranstaltung hat sich zum wichtigsten Treffpunkt für Entscheidungsträger aus Politik, Startups und Verwaltung entwickelt.

Trotz des wachsenden Interesses und der zunehmenden Bedeutung von GovTech-Lösungen gibt es noch Herausforderungen. Dazu gehören die komplexen Beschaffungsprozesse der öffentlichen Hand, die oft eine Hürde für junge Unternehmen darstellen, sowie der Mangel an Überblick über die vielfältige GovTech-Landschaft.

Die GovTech-Landschaft in Deutschland befindet sich in einer dynamischen Entwicklungsphase. Momentan lässt sich XXX GovTech Unternehmen zählen. Mit zunehmendem Fokus auf die Digitalisierung der öffentlichen Verwaltung und wachsendem Interesse an innovativen Lösungen ist zu erwarten, dass dieser Sektor in den kommenden Jahren weiter an Bedeutung gewinnen wird.

### 2.2 Netzwerktheorie und ihre Anwendung auf Unternehmensökosysteme

Die Netzwerktheorie bietet einen robusten Rahmen für die Analyse komplexer Beziehungen und Interaktionen in Unternehmensökosystemen. Grundlegend für diesen Ansatz ist die Konzeptualisierung von Ökosystemen als Netzwerke, in denen Akteure (Knoten) durch verschiedene Arten von Beziehungen (Kanten) verbunden sind (Battistella u. a. 2013). Diese Perspektive ermöglicht es, die Struktur und Dynamik von Ökosystemen quantitativ zu erfassen und zu analysieren.

Ein zentrales Konzept in der Anwendung der Netzwerktheorie auf Unternehmensökosysteme ist die Identifikation von Schlüsselakteuren und strukturellen Positionen. Hierbei spielen Zentralitätsmaße eine wichtige Rolle, die die relative Bedeutung von Akteuren im Netzwerk quantifizieren (Freeman 1978). Beispielsweise kann die Betweenness-Zentralität Akteure identifizieren, die als Brücken zwischen verschiedenen Teilnetzwerken fungieren und somit eine wichtige Rolle bei der Integration des Ökosystems spielen (Burt 2004).

Die Analyse der Netzwerkstruktur kann auch Aufschluss über die Robustheit und Anpassungsfähigkeit des Ökosystems geben. Dichte und gut vernetzte Ökosysteme können einerseits den Informations- und Ressourcenfluss erleichtern, andererseits aber auch zu einer gewissen Rigidität führen (Uzzi 1997). Im Gegensatz dazu können Ökosysteme mit einer ausgewogenen Mischung aus starken und schwachen Verbindungen sowohl Stabilität als auch Innovationspotenzial bieten (Granovetter 1973).

Die Netzwerktheorie ermöglicht es auch, die Grenzen von Ökosystemen zu untersuchen und die Verbindungen zwischen verschiedenen Ökosystemen zu analysieren. Dies ist besonders relevant in einer zunehmend vernetzten Wirtschaft, in der Unternehmen oft in mehreren, sich überschneidenden Ökosystemen agieren (Shipilov und Gawer 2015).

## 2.3 LinkedIn als Datenquelle

LinkedIn bietet sich als ideale Plattform für die strukturelle Analyse des GovTech-Ökosystems in Deutschland an. Als professionelles Netzwerk spiegelt es die Verbindungen und Interaktionen zwischen Akteuren im GovTech-Bereich wider und ermöglicht tiefe Einblicke in die Struktur und Dynamik dieses Sektors.

Die Architektur LinkedIns als socialmedia Plattform ermöglicht es die Nutzeraktivitäten, wie das Teilen von Beiträgen und Interaktionen, Rückschlüsse auf Wissensaustausch und Kommunikationsflüsse innerhalb des Netzwerks zu ziehen. Dies ist besonders relevant für das Verständnis der Dynamik und Entwicklung des jungen GovTech-Sektors. Durch das Scraping von LinkedIn-Daten lässt sich eine detaillierte Kartierung des GovTech-Ökosystems erstellen, die sowohl formale Strukturen als auch informelle Beziehungen abbildet. (Sachini u. a. 2022)

## 3 Methodik

### 3.1 Theoretische Fundierung der Datenerhebung

Die Wahl der Methodik zur Datenerhebung basiert auf fundierten theoretischen Überlegungen. Web Scraping hat sich in der Forschung als effektive Methode zur Datengewinnung im Kontext von Social-Media-Netzwerkanalysen etabliert. Insbesondere für die Analyse von Geschäftsnetzwerken bietet LinkedIn eine reichhaltige Datenquelle, da die Plattform primär für professionelle Vernetzung genutzt wird.

### 3.2 Datenerhebung mittels Web Scraping

Für die Datenerhebung und -verarbeitung der GovTech-Profile wurde ein robuster und skalierbarer Techstack entwickelt. Dieser basiert auf bewährten Technologien und Best Practices der Softwareentwicklung. Die Wahl der einzelnen Komponenten erfolgte unter Berücksichtigung ihrer Leistungsfähigkeit, Flexibilität und Kompatibilität.

#### 3.2.1 Techstack und Prozessablauf

Der gesamte Datenerhebungsprozess ist in Docker-Containern gekapselt, was eine konsistente Entwicklungs- und Produktionsumgebung gewährleistet. Diese Containerisierung ermöglicht nicht nur eine einfache Skalierung, sondern auch eine hohe Reproduzierbarkeit der Ergebnisse.

Kernkomponenten des Techstacks:

- **Containerisierung:** Docker
- **Datenbank:** PostgreSQL
- **Programmiersprache:** Python
- **Web Scraping:** Inoffizielle LinkedIn API ([tomquirk2023](#))
- **Orchestrierung:** Ansible für Deployment und Konfiguration

Der Datenerhebungsprozess gliedert sich in mehrere aufeinanderfolgende Schritte:

1. **Initialisierung:** Ausführung des Setup-Playbooks mittels Ansible zur Initialisierung der Entwicklungsumgebung.
2. **Datenbankaufbau:** Start von PostgreSQL und Erstellung des Datenbankschemas.
3. **Scraping-Prozess:** Extraktion von Profildaten der GovTech-Unternehmen mittels Python-basiertem Scraper.

4. **Datenverarbeitung:** Identifikation von Erwähnungen und Verlinkungen zwischen Unternehmen.
5. **Datenspeicherung:** Speicherung der verarbeiteten Daten in der PostgreSQL-Datenbank.
6. **Aktivitätsbündelung:** Zusammenführung von Post-Aktivitäten der Unternehmen und ihrer Gründer.

Diese strukturierte Vorgehensweise ermöglicht eine effiziente und zuverlässige Datenerhebung, die die Grundlage für die anschließende Netzwerkanalyse bildet.

### 3.2.2 Scraper-Architektur und -Prozesse

Die Architektur des LinkedIn-Scrapers wurde modular konzipiert, um eine hohe Flexibilität und Wartbarkeit zu gewährleisten. Sie besteht aus mehreren Hauptkomponenten, die in ihrer Gesamtheit einen effizienten und skalierbaren Scraping-Prozess ermöglichen.

Hauptkomponenten des Scrapers:

- **CookieController und CookieManager:** Verwaltung der Authentifizierung
- **InputController und InputLoader:** Verarbeitung der Eingabedaten
- **ScrapingController:** Orchestrierung des Scraping-Prozesses
- **ScrapingLogic:** Implementierung der Scraping-Logik
- **ScrapingDataManager:** Verwaltung der gescrapten Daten
- **Logger:** Konsistentes Logging

Der Scraping-Prozess selbst läuft in mehreren Phasen ab. Nach der Initialisierung und Konfiguration des Systems erfolgt die Cookie-Verwaltung für die API-Authentifizierung. Anschließend werden die Eingabedaten verarbeitet, um die zu scrapenden LinkedIn-Profile zu identifizieren. Für jedes Profil wird dann der eigentliche Scraping-Vorgang durchgeführt, wobei kontinuierlich Daten gespeichert und Fehler behandelt werden.

### 3.2.3 Datenbankstruktur

Die Datenbankstruktur für das GovTech-Netzwerk-Projekt besteht aus sechs Haupttabellen:

#### 3.2.3.1 Persons

Diese Tabelle enthält Informationen über individuelle Personen im Netzwerk.

Spaltenname	Datentyp	Beschreibung
id	VARCHAR(255)	Primary Key
publicId	VARCHAR(255)	Öffentliche ID
trackingId	VARCHAR(255)	Tracking-ID
profileId	VARCHAR(255)	Profil-ID
occupation	VARCHAR(255)	Beruf/Position
firstName	VARCHAR(100)	Vorname
lastName	VARCHAR(100)	Nachname
picture	VARCHAR(512)	URL des Profilbildes

Spaltenname	Datentyp	Beschreibung
id	VARCHAR(255)	Primary Key
url	VARCHAR(512)	URL des Firmenprofils
fullName	VARCHAR(255)	Vollständiger Firmenname
image	VARCHAR(512)	URL des Firmenlogos
followers	INT	Anzahl der Follower

### 3.2.3.2 Companies

Diese Tabelle speichert Informationen über Unternehmen im Netzwerk.

### 3.2.3.3 Person2Company

Diese Tabelle stellt die Beziehung zwischen Personen und Unternehmen dar.

Spaltenname	Datentyp	Beschreibung
personId	VARCHAR(255)	Fremdschlüssel zu Persons
companyId	VARCHAR(255)	Fremdschlüssel zu Companies

### 3.2.3.4 Posts

Diese Tabelle enthält Informationen über einzelne Posts im Netzwerk.

Spaltenname	Datentyp	Beschreibung
id	VARCHAR(255)	Primary Key (aus URN)
text	TEXT	Inhalt des Posts
url	VARCHAR(512)	URL des Posts
postedAtTimestamp	BIGINT	Zeitstempel des Postings
postedAtISO	TIMESTAMP	ISO-formatierter Zeitpunkt

### 3.2.3.5 Author2Post

Diese Tabelle verknüpft Autoren (Personen oder Unternehmen) mit ihren Posts.

### 3.2.3.6 Attributes (Mentions)

Diese Tabelle speichert Informationen über Erwähnungen in Posts.

## 3.3 Netzwerkanalyse

Die Netzwerkanalyse des GovTech-Ökosystems basiert auf einem komplexen Verständnis von Akteuren und deren Interaktionen. Ein Akteur in diesem Kontext kann entweder ein Unternehmen selbst oder ein Angestellter sein. Dabei unterscheiden wir zwischen internen Akteuren des GovTech-Netzwerks (GovTech-Unternehmen oder deren Mitarbeiter) und externen Akteuren (z.B. Behörden, Politiker oder andere Unternehmen außerhalb des GovTech-Bereichs).

Die Netzwerkstruktur wird durch Verlinkungen zwischen den Akteuren definiert. Jede Verlinkung, sei es eine Erwähnung in einem Post, ein Kommentar oder eine direkte Verknüpfung, erzeugt eine Kante im Netzwerk. Diese Kanten repräsentieren die Beziehungen und Interaktionen zwischen den Akteuren.

Für die aktuelle Analyse betrachten wir diese Kanten als ungewichtet und ungerichtet. In zukünftigen Analysen könnte jedoch eine Gewichtung der Kanten implementiert werden, um

Spaltenname	Datentyp	Beschreibung
authorId	VARCHAR(255)	Fremdschlüssel zu Persons oder Companies
postId	VARCHAR(255)	Fremdschlüssel zu Posts

Spaltenname	Datentyp	Beschreibung
id	VARCHAR(255)	Primary Key
postId	VARCHAR(255)	Fremdschlüssel zu Posts
start	INT	Startposition der Erwähnung
length	INT	Länge der Erwähnung
reference	VARCHAR(255)	ID der erwähnten Person oder Firma

die Intensität und Qualität der Beziehungen genauer abzubilden. Mögliche Gewichtungsfaktoren könnten sein:

- Art des Akteurs (Unternehmen vs. Angestellter)
- Typ der Interaktion (Verlinkung, Like, Kommentar)
- Häufigkeit der Interaktionen
- Reichweite und Einfluss des verlinkenden Akteurs

Die Netzwerkanalyse umfasst verschiedene Metriken zur Charakterisierung der Netzwerkstruktur:

- Zentralitätsmaße (Degree-, Betweenness-, Eigenvector-Zentralität)
- Clustering-Koeffizient
- Pfadlängen und Durchmesser des Netzwerks
- Identifikation von Schlüsselakteuren und Brückenakteuren

### 3.4 Visualisierung der Daten

Die Visualisierung der Netzwerkdaten spielt eine entscheidende Rolle für das Verständnis und die Interpretation der Ergebnisse. Wir nutzen verschiedene gängige Methoden zur Darstellung des GovTech-Netzwerks:

1. **Knotenpunkt-Kanten-Diagramme:** Diese klassische Netzwerkvisualisierung stellt Akteure als Knoten und Beziehungen als Kanten dar. Wir verwenden Algorithmen wie Fruchterman-Reingold oder ForceAtlas2 zur optimalen Anordnung der Knoten.
2. **Heatmaps:** Zur Darstellung der Interaktionsintensität zwischen verschiedenen Akteuren oder Gruppen von Akteuren.
3. **Circos-Plots:** Diese kreisförmigen Diagramme eignen sich besonders gut, um Beziehungen zwischen verschiedenen Kategorien von Akteuren zu visualisieren.
4. **Ego-Netzwerke:** Für die Darstellung der direkten Verbindungen einzelner wichtiger Akteure.
5. **Dynamische Netzwerkvisualisierungen:** Zur Veranschaulichung der zeitlichen Entwicklung des Netzwerks.

Bei der Visualisierung nutzen wir verschiedene visuelle Attribute zur Kodierung zusätzlicher Informationen:

- **Knotengröße:** Repräsentiert die Zentralität oder Wichtigkeit eines Akteurs.
- **Knotenfarbe:** Unterscheidet zwischen internen und externen Akteuren oder verschiedenen Typen von Akteuren.
- **Kantendicke:** Stellt die Stärke oder Häufigkeit der Interaktionen dar.
- **Kantenfarbe:** Kann verschiedene Arten von Beziehungen kodieren.

Interaktive Visualisierungen ermöglichen es den Nutzern, das Netzwerk zu erkunden, in bestimmte Bereiche zu zoomen und zusätzliche Informationen zu einzelnen Akteuren oder Beziehungen abzurufen. Tools wie Gephi, NetworkX in Kombination mit Matplotlib, oder D3.js für webbasierte Visualisierungen kommen zum Einsatz.

## 4 Ergebnisse und Diskussion

### 4.1 Darstellung des GovTech-Netzwerks

### 4.2 Identifikation von Schlüsselakteuren und Clustern

### 4.3 Interpretation der Netzwerkstruktur

## 5 Schlussfolgerungen und Ausblick

## Literatur

- Battistella, Cinzia u. a. (2013). “Methodology of business ecosystems network analysis: A case study in Telecom Italia Future Centre”. In: *Technological Forecasting and Social Change* 80.6, S. 1194–1210.
- Bogumil, Jörg, Stephan Grohs und Sabine Kuhlmann (2019). “Ergebnisse und Wirkungen kommunaler Verwaltungsmodernisierung in Deutschland – Eine Evaluation nach zehn Jahren Praxiserfahrung”. In: *Reflex - Ergebnisbericht*. URL: [https://www.tu-chemnitz.de/wirtschaft/bwl9/forschung/fprojekte/reflex/instReflex/ergebnisse/download/PVSSH37\\_07Bogumil.pdf](https://www.tu-chemnitz.de/wirtschaft/bwl9/forschung/fprojekte/reflex/instReflex/ergebnisse/download/PVSSH37_07Bogumil.pdf).
- Burt, Ronald S (2004). “Structural holes and good ideas”. In: *American journal of sociology* 110.2, S. 349–399.
- Freeman, Linton C (1978). “Centrality in social networks conceptual clarification”. In: *Social networks* 1.3, S. 215–239.
- Granovetter, Mark S (1973). “The strength of weak ties”. In: *American journal of sociology* 78.6, S. 1360–1380.
- Nathalie Haug, Sorin Dan und Ines Mergel (2024). “Digitally-induced change in the public sector: a systematic review and research agenda”. In: *Public Management Review* 26.7, S. 1963–1987. DOI: 10.1080/14719037.2023.2234917. eprint: <https://doi.org/10.1080/14719037.2023.2234917>. URL: <https://doi.org/10.1080/14719037.2023.2234917>.
- Niehaves, Björn, Luca T. Bauer und Gerhard Klassen (Apr. 2024). *Government Technology*. Vorlesung. Sommersemester 2024.
- Sachini, Evi u. a. (2022). “Lessons for science and technology policy? Probing the LinkedIn network of an RDI organisation”. In: *Journal of Innovation and Entrepreneurship* 11.1, S. 53. DOI: 10.1186/s13731-022-00246-z.
- Shipilov, Andrew und Annabelle Gawer (2015). “Network dynamics and ecosystem performance”. In: *Strategic Management Journal* 36.4, S. 559–583.



- Silve, Arthur und Mariano Moszoro (2023). “The Political Economy of GovTech”. In: *IMF Notes* 2023.003, A001. DOI: 10.5089/9798400246500.068.A001. URL: <https://www.elibrary.imf.org/view/journals/068/2023/003/article-A001-en.xml>.
- Uzzi, Brian (1997). “Social structure and competition in interfirm networks: The paradox of embeddedness”. In: *Administrative science quarterly*, S. 35–67.