

Netzwerkanalyse

AUTUMN SCHOOL „DIGITALE EDITION – VERTIEFUNG UND NUTZUNG“

Bernhard Geiger
Wien, 3. Oktober 2018

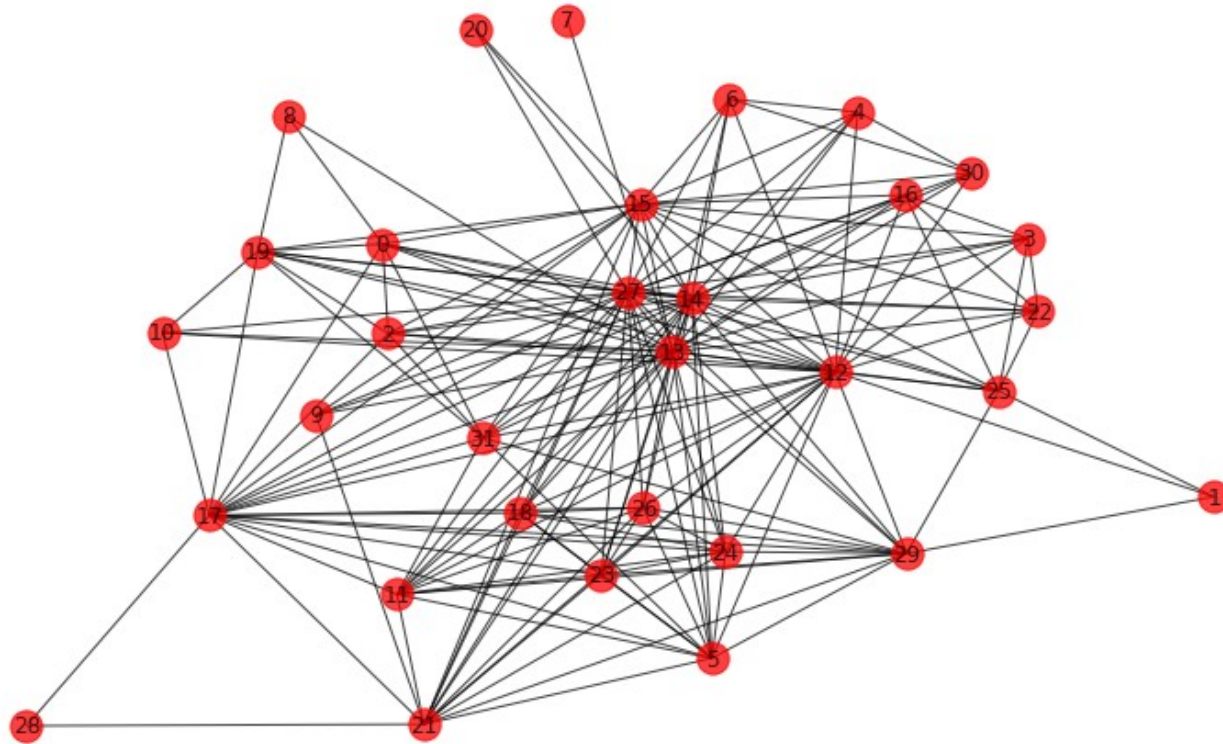




- Vorbereitende Übung: count-basierte Zentralität
- Netzwerke – Graphen
 - Gerichtet, gewichtet oder bipartit?
- Netzwerkbasierte Zentralität
 - Betweenness, closeness, eigenvector?
- Community Detection
 - Zielfunktion und Algorithmus?



NETZWERK – GRAPH

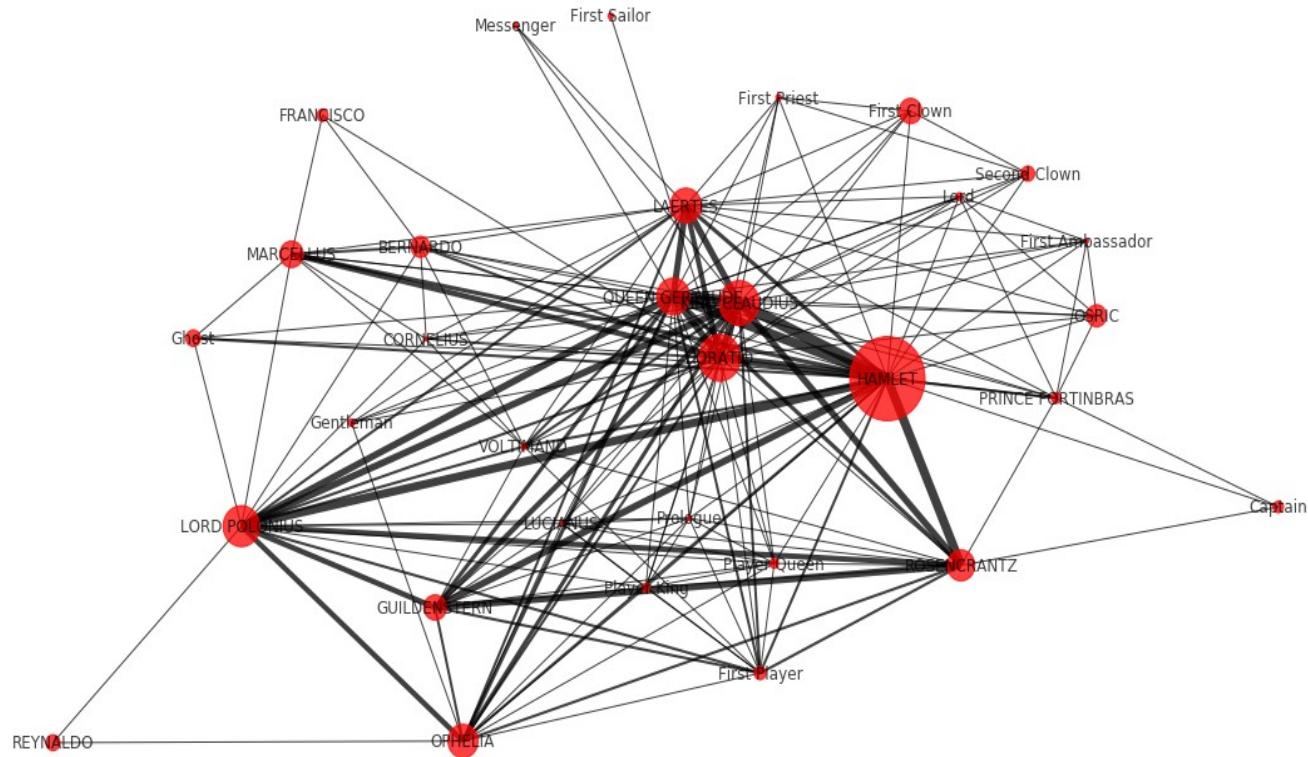




- **Gewichteter Graph:** Jede Kante hat ein Gewicht, das die Stärke der Verbindung zwischen den Knoten widerspiegelt.
- **Gerichteter Graph:** Jede Kante hat eine Richtung, wird also durch einen Pfeil dargestellt.
- **Bipartiter Graph:** Die Knotenmenge besteht aus zwei Teilen (= bipartit), und Kanten existieren nur zwischen Knoten unterschiedlicher Teile.

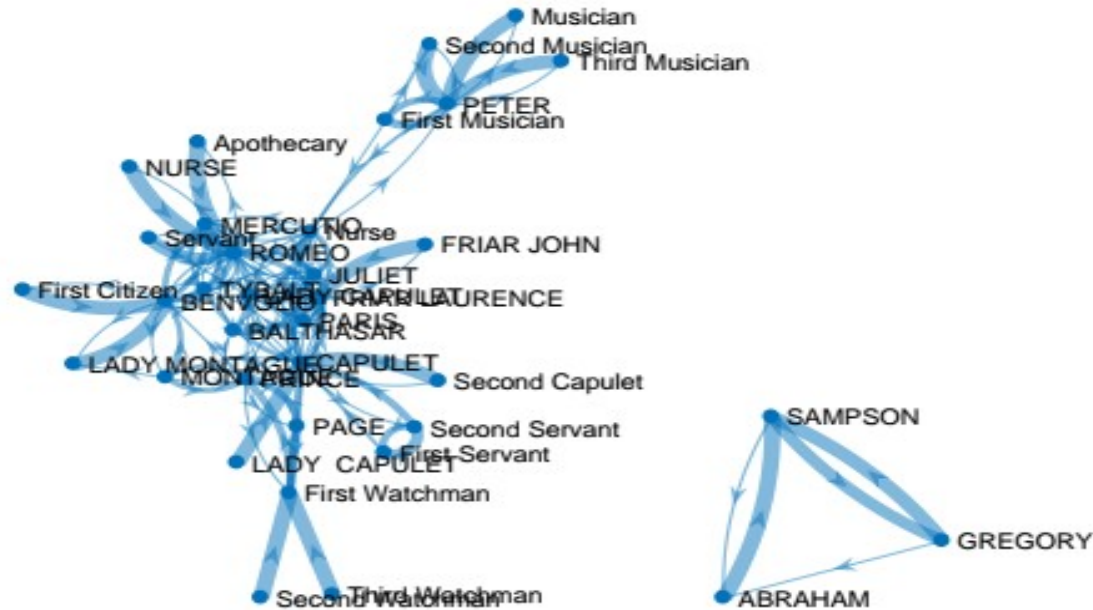


NETZWERK – GRAPH



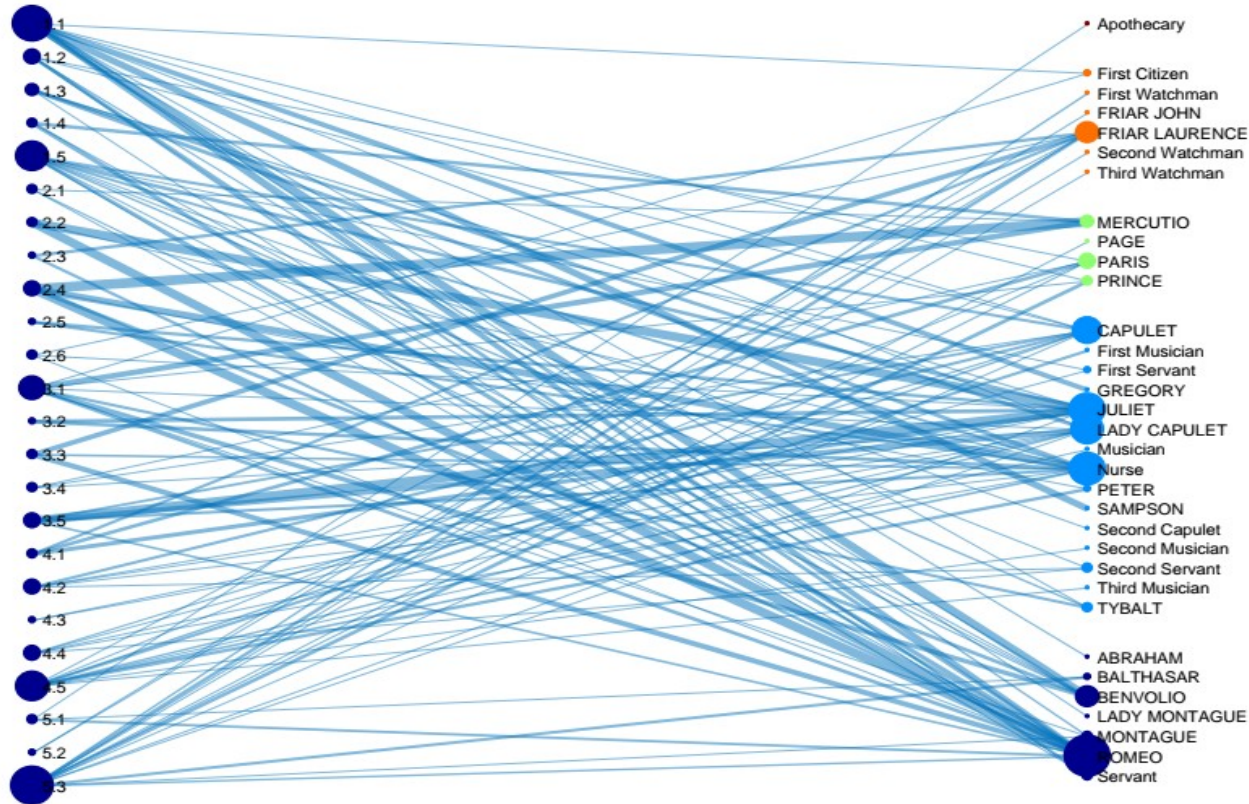


NETZWERK – GRAPH





NETZWERK – GRAPH





ÜBUNG: HAMLET ALS GEWICHTETER GRAPH



NETZWERKBASIERTE ZENTRALITÄT



WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction

Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools

What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item
Cite this page

Print/export

Create a book
Download as PDF
Printable version

Languages

Čeština
Español
فارسی

Article Talk

Centrality

From Wikipedia, the free encyclopedia

For the statistical concept, see [Central tendency](#).

In [graph theory](#) and [network analysis](#), indicators of **centrality** identify the most important [vertices](#) within a graph. Applications include identifying the most influential people in the Internet or urban networks, and [super-spreaders](#) of disease. Centrality concepts were first developed in [social network analysis](#), and many of the terms used to They should not be confused with [node influence metrics](#), which seek to quantify the influence of every node in the network.

Contents [hide]

- 1 Definition and characterization of centrality indices
 - 1.1 Characterization by network flows
 - 1.2 Characterization by walk structure
 - 1.3 Radial-volume centralities exist on a spectrum
 - 1.4 Game-Theoretic Centrality
- 2 Important limitations
- 3 Degree centrality
- 4 Closeness centrality
 - 4.1 Harmonic centrality
- 5 Betweenness centrality
- 6 Eigenvector centrality
 - 6.1 Using the adjacency matrix to find eigenvector centrality
- 7 Katz centrality
- 8 PageRank centrality
- 9 Percolation centrality
- 10 Cross-clique centrality
- 11 Freeman centralization
- 12 Dissimilarity based centrality measures
- 13 Extensions
- 14 See also
- 15 Notes and references
- 16 Further reading

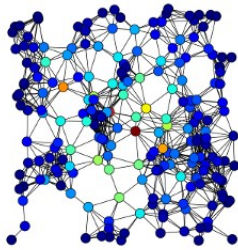


(Fokus auf ungewichtete, ungerichtete Graphen)

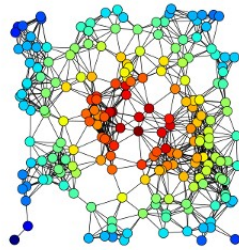
- **Degree centrality:** Wieviele Nachbarn hat ein Knoten?
- **Closeness centrality:** Wie nahe sind (im Durchschnitt) alle anderen Knoten?
- **Betweenness centrality:** Wie schwer kommt man an diesem Knoten vorbei?
- **Eigenvector centrality:** Frobenius-Perron-Vektor der Adjazenzmatrix des Graphen, hängt irgendwie mit unendlich langen Pfaden im Netzwerk zusammen...



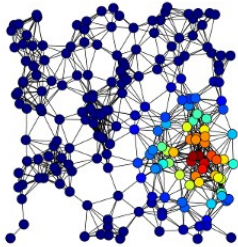
NETZWERKBASIERTE ZENTRALITÄT



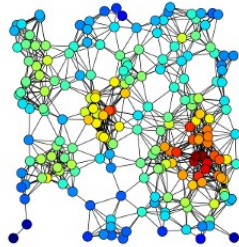
A



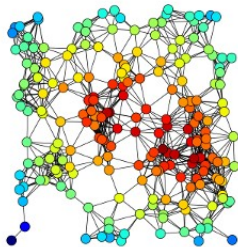
B



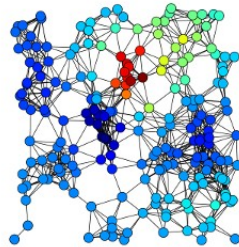
C



D



E



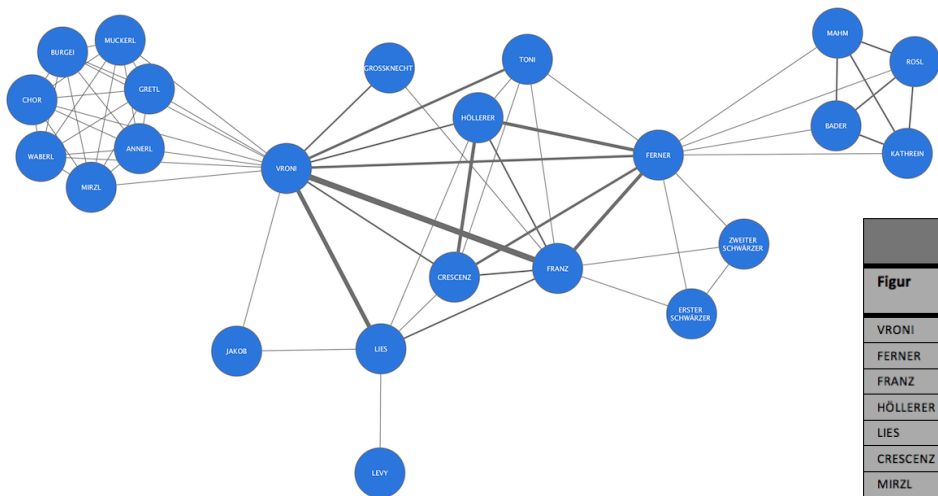
F

- A: Betweenness
- B: Closeness
- C: Eigenvector
- D: Degree
- E: Harmonic
- F: Katz

Taken from Wikipedia, created by user Tapiocozzo under CC BY-SA 4.0



NETZWERKBASIERTE ZENTRALITÄT



Anzengrubers „Der Meineidbauer“

Taken and adapted from Frank Fischer, Peer Trilcke, Christopher Kittel, Carsten Milling, Daniil Skorinkin, “To Catch a Protagonist: Quantitative Dominance Relations in German Language Drama (1730–1930)”, In Digital Humanities 2018 (Mexico City, 16--29 June 2018).

[Link](#)

	Countbasiert			Netzwerkbasiert				
Figur	Häufig- keit	Sprech- akte	Wörter	Degree	Weighted Degree	Betweenness	Closeness	Eigenvector
VRONI	1	1	1	1	1	1	1	1
FERNER	3	2	2	2	2	2	2	3
FRANZ	2	2	3	3	2	3	3	2
HÖLLERER	5	5	8	11	4	5	4	4
LIES	4	4	6	11	6	4	7	6
CRESCENZ	6	9	12	11	5	5	4	5
MIRZL	13	14	13	4	7	7	8	11
TONI	7	7	10	14	7	7	6	7
BURGEI	13	12	16	4	7	7	8	11
MAHM	9	6	5	15	7	7	19	19
GROSSKNECHT	7	8	4	21	19	7	15	8
MUCKERL	13	14	15	4	7	7	8	14
BADER	9	11	11	15	7	7	19	19
ROSL	9	12	14	15	7	7	19	19
WABERL	13	14	17	4	7	7	8	14
ANNERL	13	14	18	4	7	7	8	14
CHOR	13	21	23	4	7	7	8	11
GRETJ	13	20	20	4	7	7	8	14
KATHREIN	9	14	19	15	7	7	19	19
JAKOB	13	9	7	21	22	7	16	18
LEVY	13	14	9	23	23	7	23	23
1. SCHWÄRZER	13	21	21	19	19	7	17	9
2. SCHWÄRZER	13	21	21	19	19	7	17	9



ÜBUNG: NETZWERKBASIERTE ZENTRALITÄT



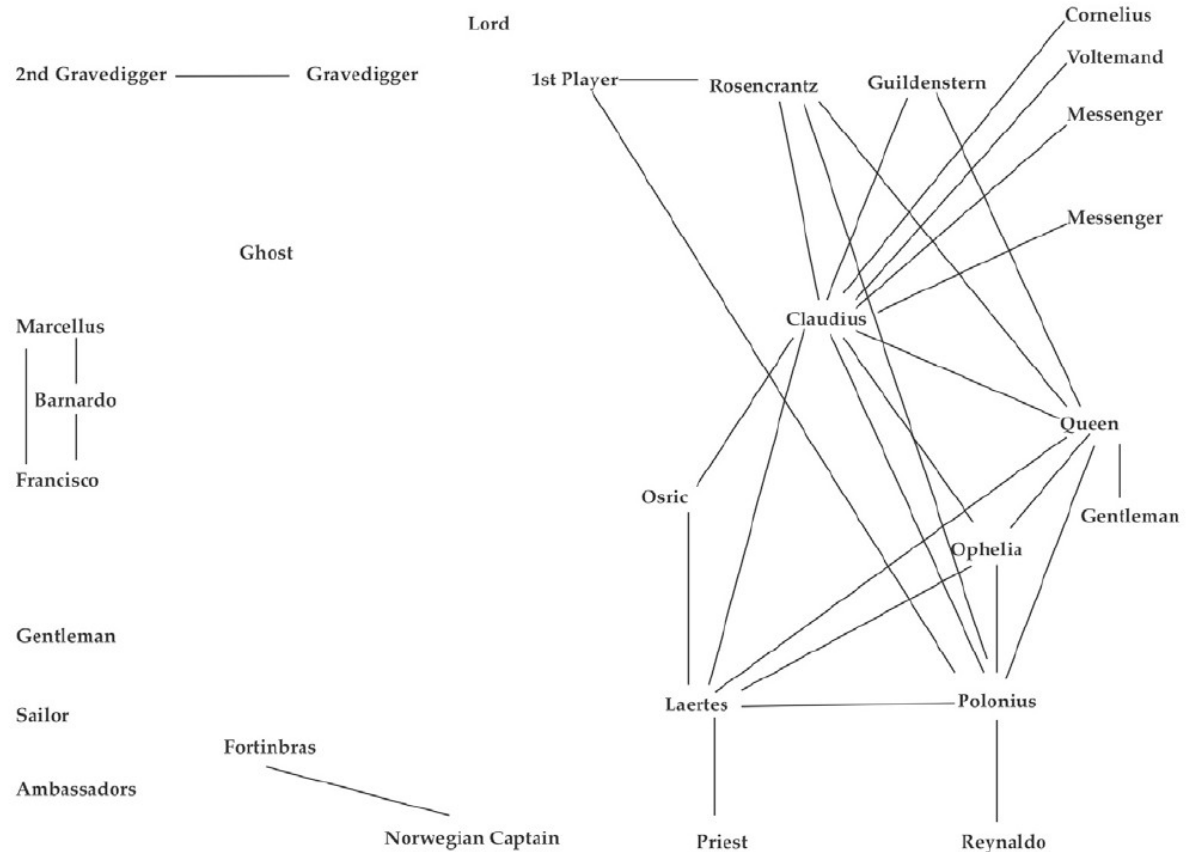
WAS STIMMT MIT HORATIO NICHT?

„Hamlet“ ohne Hamlet und Horatio

Taken from Franco Moretti,
“Network Theory, Plot
Analysis.” New Left Review,
no. 68(March):80. 2011

[Link](#)

*„I may be making too
much of this; or, Horatio
may really be a fantastic
half-intuition on
Shakespeare’s part“
(ibid., p. 7)*





COMMUNITY DETECTION

„In the study of complex networks, a network is said to have community structure if the nodes of the network can be easily grouped into (potentially overlapping) sets of nodes such that each set of nodes is densely connected internally.“

Taken from Wikipedia, „Community Structure“

[Link](#)



Es gibt eine Vielzahl von Kenngrößen dafür, wie gut eine Community-Struktur zu einem gegebenen Graphen passt:

- **Modularity:** Kanten innerhalb der Community vs. Kanten zwischen Communities.
- **Minimum Cut:** Anzahl bzw. Gewicht der Kanten zwischen Communities
- **Stochastic Block Modeling:** basiert auf der Theorie zufälliger Graphen...
- Etc.



Methode = Zielfunktion + Algorithmus

z.B. Modularität

- Clauset-Newman-Moore
- Louvain
- Label propagation
- Relaxierung auf Eigenwertproblem
- Etc.



ÜBUNG: COMMUNITY DETECTION



Netzwerke zur Visualisierung

Eine neue Sicht auf die Daten - aber:

- Welche Darstellung wählen wir?

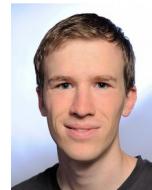
Netzwerkanalyse auf großen Korpusen

Stellen und Beantworten neuer Forschungsfragen - aber:

- Welche Darstellung wählen wir?
- Welche Zielfunktion wählen wir?
- (Falls nötig) Welchen Algorithmus wählen wir?



AUSTRIA'S LEADING RESEARCH CENTER
FOR DATA-DRIVEN BUSINESS AND BIG DATA ANALYTICS



Bernhard Geiger
Knowledge Discovery
geiger@ieee.org

Know-Center GmbH

Research Center for Data-Driven
Business and Big Data Analytics
Inffeldgasse 13/6
8010 Graz, Austria

Firmenbuchgericht Graz
FN 199 685 f
UID: ATU 50367703

gefördert durch das Programm COMET (Competence Centers for Excellent Technologies), wir danken unseren Fördergebern:

