

Netzwerkanalyse I

Kai Arzheimer | Vorlesung Forschungsmethoden

Einführung

Was sind Netzwerke?

Netzwerkdaten

Netzwerkanalyse (SNA)

Grundzüge der Netzwerkanalyse

Design und Datenerhebung

Konzepte

Perspektiven

Zentrale Konzepte

Linien, Richtung, Dichte

Zentralität und Zentralisierung

Komponenten, Kerne und Cliques

Zusammenfassung

Lernziele

1. Netzwerkanalyse als Ansatz
2. Basale Konzepte

Einführung

- Politikwissenschaft klassische Sozialwissenschaft („zoon politikon“)
- Klassisches „Badewannenmodell“ (Coleman etc.)
 - Makro-Ebene
 - Meso-Ebene
 - Mikro-Ebene
- Gruppen bzw. deren Strukturen und Interaktionen als Analysegegenstand
- Methodologischer Individualismus

- Politikwissenschaft klassische Sozialwissenschaft („zoon politikon“)
- Klassisches „Badewannenmodell“ (Coleman etc.)
 - Makro-Ebene
 - Meso-Ebene
 - Mikro-Ebene
- Gruppen bzw. deren Strukturen und Interaktionen als Analysegegenstand
- Methodologischer Individualismus
- „Sozial“ – *Beziehungen zwischen Menschen*

Wo ist das Problem?

- Konzentration auf Individuum
- Verstärkt durch moderne Surveyforschung: ein- oder mehrstufige Zufallsauswahl von Individuen
- „Bundesweit repräsentativ ausgewählte Befragte“ = 2 000 atomisierte Individuen
 - Informationen über politischen, räumlichen etc. Kontext: Mehr-Ebenen-Analyse
 - Informationen über Interaktionspartner: Netzwerkanalyse („Klumpen“)
- Standardprozedur: (lineare) Analyse von $y = \beta_0 + \beta_1 x_1 + \dots$
- Viele interessante Fragen so nicht zu untersuchen

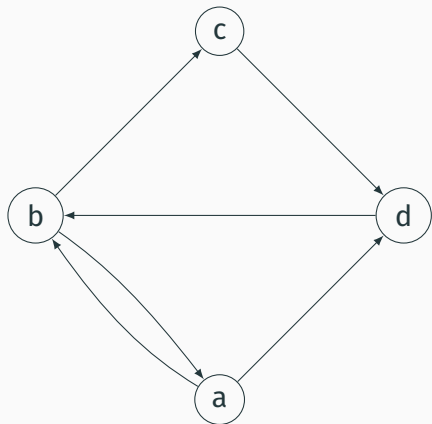
Was ist ein Netzwerk?

- Set von Akteuren
- Plus deren *Beziehungen* (Dominanz, Kommunikation, Kooperation, Verwandtschaft ...)
- Netzwerkperspektive:
 - Rückkehr zu den Ursprüngen der Sozialforschung
 - Kein neues/konkurrierendes, sondern komplementäres Paradigma
- Besondere Daten und Analyseverfahren
- Besondere Theorien?
- Ebook von Hanneman/Riddle:
<http://www.faculty.ucr.edu/~hanneman/nettext/>

Wie sehen „normale“ Daten aus?

- Daten-Matrix
 - „Fälle“ in den Zeilen
 - „Variablen“ (Information über *einzelne* Fälle) in den Spalten
 - Ergänzung um Kontextdaten möglich
 - Ergänzung um Informationen über ego-zentrierte Netzwerke möglich (ALLBUS)
 - Analyse mit „normalen“ Statistikprogrammen (Stata etc.)
- Keine Informationen über Beziehungen zwischen den Fällen in der Matrix

Wie sieht ein Netzwerk aus?



- Komplexere Datenstruktur; besondere Software

- Knoten, Kanten, Netz
- Knoten
 - Globale Eigenschaften
 - Relationale Eigenschaften?
- Kanten
 - (Bedeutung?)
 - Gerichtet?
 - Gewichtet bzw. mehrere?
- Netz: Struktur

Netzwerkanalyse (SNA)

Worum geht es?

- *Strukturelle* Beziehungen zwischen
- Akteuren (individuelle oder kollektive)
 - Natürliche Personen
 - Staaten
 - Gruppen
 - Organisationen
 - ...
- Ziel:
 - Messen und Repräsentieren
 - Erklärungen für Auftreten und Konsequenzen der Beziehungen

1. Strukturelle Beziehungen oft wichtiger als individuelle globale Variablen
 - Globale Variablen weitgehend konstant
 - Handeln kontextabhängig
 2. Soziale Netzwerke beeinflussen Wahrnehmungen, Überzeugungen, Handlungen
 - Kommunikationsflüsse (strong vs. weak ties)
 - Bridging and Bonding social capital
 3. Strukturelle Beziehungen sind dynamisch (Zeitperspektive)
 - Rückkopplung
 - Makro – Mikro – Makro
-
- Kein Gegensatz, sondern Verbindung von Akteurs- und Strukturperspektive

- Beziehung: *gemeinsame* Eigenschaft der Dyade
- Gerichtet/ungerichtete Beziehungen
- Enorme Vielfalt von Beziehungen
 - Charakter, Wahrnehmung, Stabilität, Intensität, (Richtung), ...
 - Problematik von Auswahl und Messung
- Interpretation von Realität durch Forscher (kein spezifisches SNA-Problem)

- Ursprünge; Arbeiten von Moreno (1934) über Beziehungen in Spielgruppen und Klassen
 - Mathematische und/oder graphische Repräsentation von
 - Akteuren und Beziehungen bzw. Knoten und Kanten
- Abstrakter Apparat mit inhaltlicher Interpretation (ähnlich wie Regressionsanalyse)
- Gleiche Akteure: Verschiedene Beziehungen – unterschiedliche Netzwerke (Freundschaft vs. Hierarchie)
- Komplette vs. partielle Netzwerke – Design und Sampling wichtig

1. Sampling units: Individuen, Gruppen, Organisationen, Schichten, Staaten
 - Einstufige/zweistufige Auswahlen
 - Kleine/große Sets von Akteuren
2. Inhalt und Form der Beziehungen:
 - *Richtung* und *Stärke/Häufigkeit/Intensität* als wichtige Formen
 - *Ökonomische* oder *Interessen*, *Kommunikation*, etc. etc. als Inhalt

- Transaktion
- Kommunikation
- Mitgliedschaften
- Instrumentelle Beziehungen
- Gefühle
- Machtbeziehungen
- Verwandtschaftsbeziehungen

1. Sampling units: Individuen, Gruppen, Organisationen, Schichten, Staaten
 - Einstufige/zweistufige Auswahlen
 - Kleine/große Sets von Akteuren
2. Inhalt und Form der Beziehungen:
 - *Richtung* und *Stärke/Häufigkeit/Intensität* als wichtige Formen
 - Ökonomische oder Interessen, Kommunikation, etc. etc. als Inhalt
3. Analyseebene
 - Egozentrische Netzwerke
 - Dyaden $N \times (N - 1)$ bzw. $N \times (N - 1)/2$
 - Triaden $\binom{N}{3} = \frac{N \times (N-1) \times (N-2)}{3 \times 2 \times 1}$
 - Komplettes Netzwerk

Problem

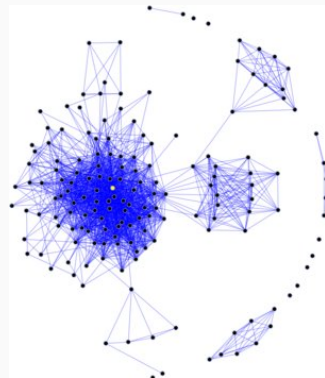
Wo liegt die Grenze eines Netzwerkes?

- Realistische (Wahrnehmung der Akteure) vs. nominalistische (Festlegung durch Forscher) Strategie
 1. Positionale Strategie (nominalistisch)
 2. Ereignisbasiert (nominalistisch)
 3. Relationale (realistische) Strategien (Reputation, snowball)
- „Generatoren“ und Archive
- Weitere Probleme
 - Cognitive maps und bias
 - Reliabilität
 - Missing Data

- Community Power Studies
- Eliten
- Organisationen
- Soziale Bewegungen
- Wahlen und politische Kommunikation
- Internationale Beziehungen
- Politische Ökonomie

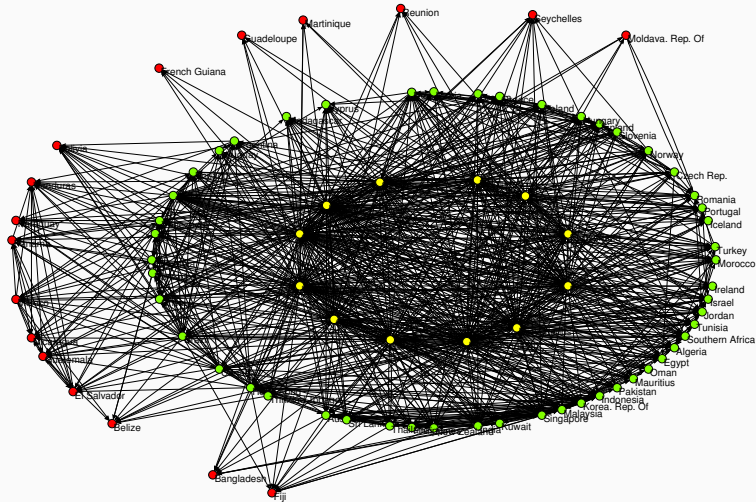
Konzepte

- Wurzeln der Terminologie
 - Soziologie
 - Mathematik (Graphentheorie)
- Alle relevanten Texte in Englisch, inzwischen aber deutschsprachige (Lehrbuch-)Literatur mit Übersetzungen
- Gelegentliche Konfusion

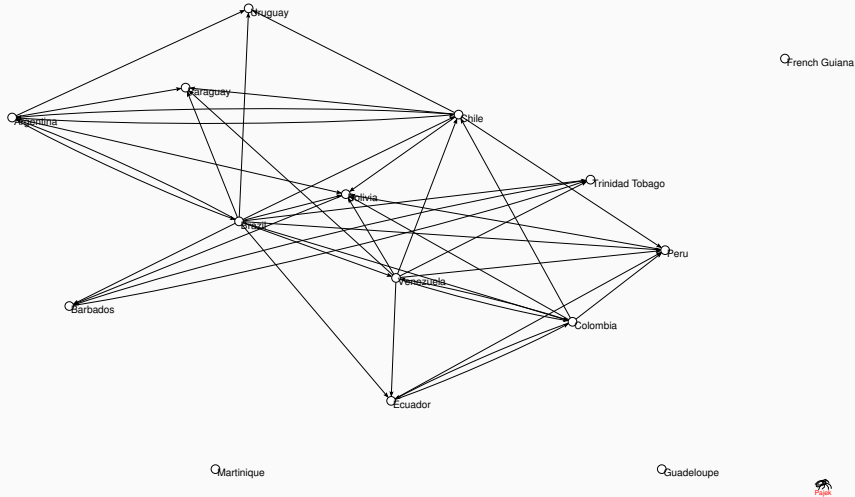


Perspektive auf das Netzwerk?

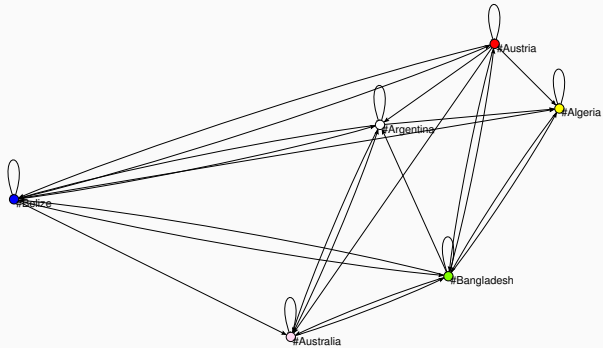
Gesamtnetzwerk



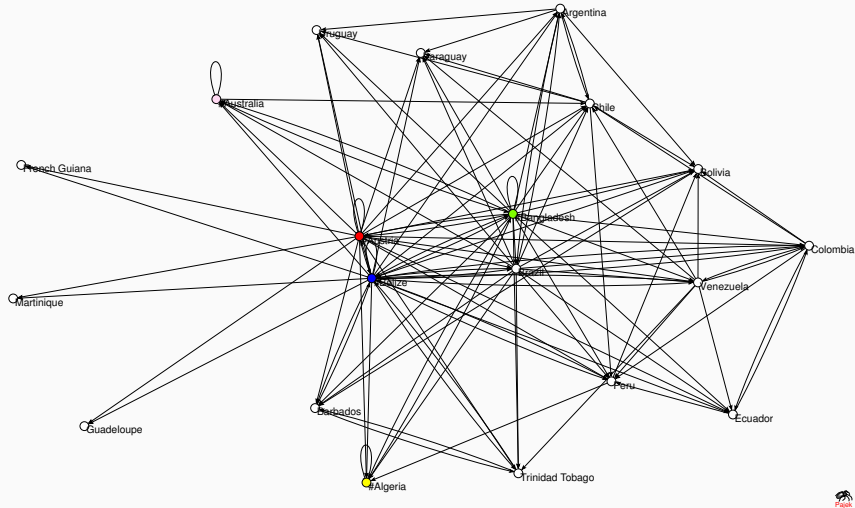
Lokale Perspektive: Südamerika



Globale Perspektive: Kontinente



Kontextuelle Perspektive: Südamerika ↔ Kontinente



Far Right Online Network in Germany (June 2015)

1. Use snowball/keyword sampling to map/crawl far-right network with HYPHE during June/July 2015
 - Start with six parties, various blogs, magazines etc.
 - Identify #2249 entities (sites, blogs, social media profiles ... nodes)
 - Connected by #8437 hyperlinks (edges)
 - What a (diligent) human would see
2. Most entities affiliated

Far Right Online Network in Germany (June 2015)

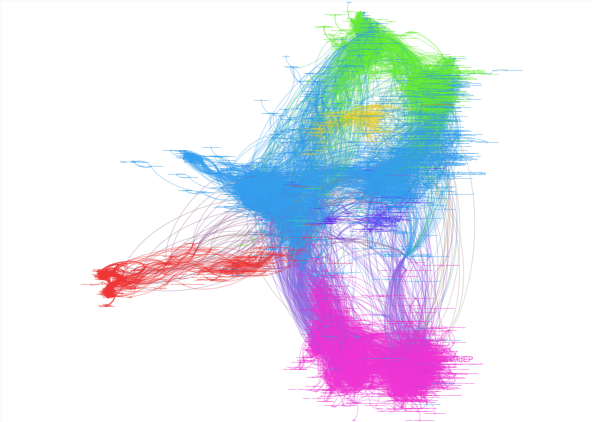
1. Use snowball/keyword sampling to map/crawl far-right network with HYPHE during June/July 2015

- Start with six parties, various blogs, magazines etc.
- Identify #2249 entities (sites, blogs, social media profiles ... nodes)
- Connected by #8437 hyperlinks (edges)
- What a (diligent) human would see

2. Most entities affiliated

Party	Count
FREIHEIT	6
PRO	58
RECHTE	72
REP	198
NPD	371
AfD	831
not affiliated	1013

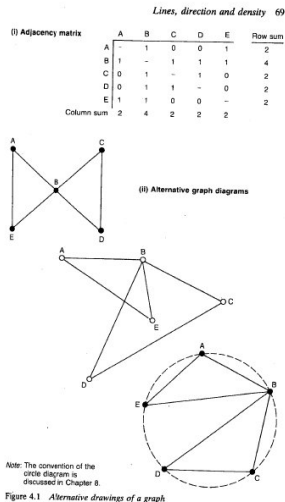
Far Right Online Network in Germany (June 2015)



AfD ein zentraler Akteur

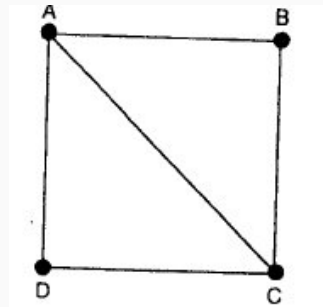
Graph und Matrix

- Netzwerk: Knoten entsprechen Akteuren, Linien Beziehungen zwischen den Punkten
 - Visualisierung nicht immer realisierbar
 - Exaktes Aussehen des Graphs irrelevant
 - *Wichtig sind Beziehungen, nicht Lage der Punkte*
- Matrixform
 - Gerichtete vs. ungerichtete Beziehungen
 - Gewichtete vs. einfache Beziehungen



Degree und Path

- Benachbarte/verbundene Punkte: adjacent
- Zahl der Nachbarn: degree (in vs. out)
- Walk (jede Verbindung) vs. *path* (jeder Knoten/jede Beziehung maximal einmal genutzt)
- A-D (1), A-C-D (2), A-B-C-D (3)
- Kürzester möglicher Pfad: geodesic
- Gerichtete Beziehungen und Pfade
- Richtung ignorieren: semi-paths



Density und Inclusiveness

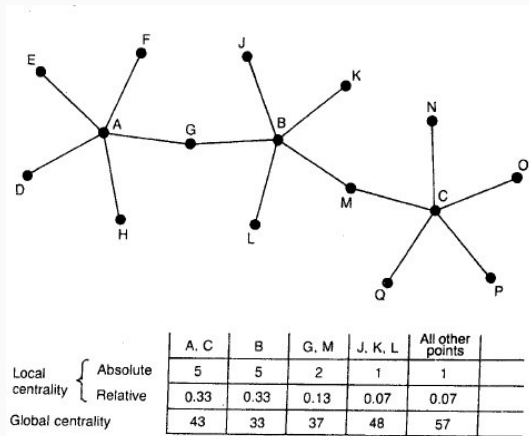
- Inclusiveness: wieviel Prozent der Punkte eingebunden?
- Density: Wieviel Prozent der theoretisch möglichen Verbindungen sind realisiert?
 - Besonderheiten: ego-zentrierte Netzwerke
 - In großen Netzwerken *ceteris paribus* niedriger – warum?
- Density als abhängige Variable

						
No. of connected points	4	4	4	3	2	0
Inclusiveness	1.0	1.0	1.0	0.7	0.5	0
Sum of degrees	12	8	6	4	2	0
No. of lines	6	4	3	2	1	0
Density	1.0	0.7	0.5	0.3	0.1	0

- Viele (eingehende) Verbindungen – hohes Prestige
- Lokale Zentralität vs. globale Zentralität
- Eigenschaft eines Knotens vs. Eigenschaft des Netzwerks

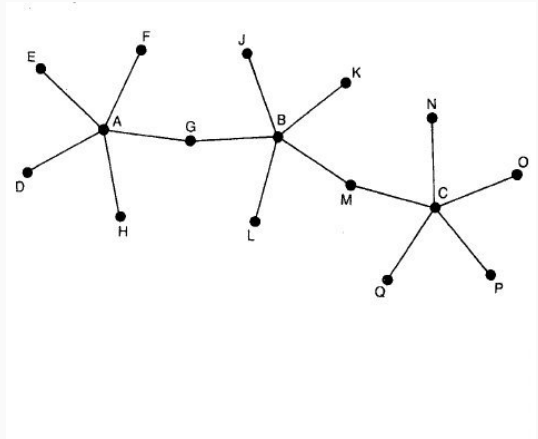
Degree/Closeness centrality (global & local)

- Degree: lokale Zentralität (degree centrality)
 - Absolut: Zahl der Nachbarn
 - Relativ: Zahl der Nachbarn / Zahl der Kanten
- Summe der kürzesten Entfernungen (geodesic) zu allen anderen Punkten: globale Zentralität (closeness centrality)
- Abhängigkeit von Netzwerkgröße



Betweenness centrality

- Knoten, der zwei andere Knoten verbindet (via geodesic), ist ein „Broker“
 - Verbindung kann unterbrochen werden
 - Filterfunktion
- Wieviele der kürzesten Verbindungen zwischen allen Paaren (X,Y) von Knoten laufen über Knoten Z?



Centralisation & Density

- Density: Kohäsion des Graphs
- Zentralisierung: Wie stark basiert Kohäsion auf „fokalen Punkten“
- Allgemeine Operationalisierung von Zentralisierung: Wie ungleich ist Zentralität verteilt?
- Verhältnis tatsächliche Ungleichheit/maximale Ungleichheit
 - degree centrality
 - closeness centrality

- betweenness centrality

- Maximale Ungleichheit: „Stern“, maximale Gleichheit: „kompletter“ Graph

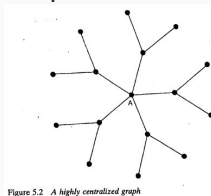
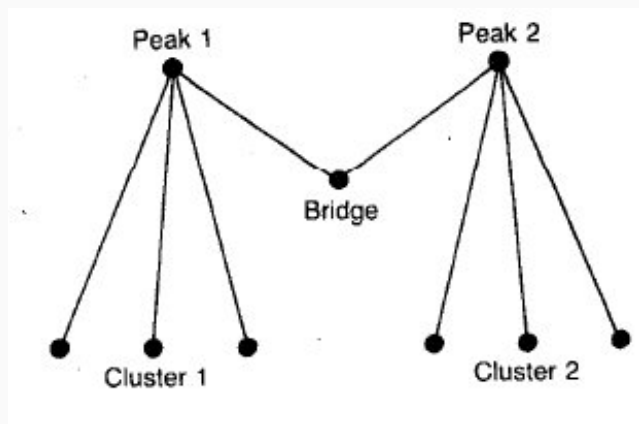


Figure 5.2 A highly centralized graph

- Ein oder mehrere zentrale Punkte (absolutes Zentrum)?

Peaks & Bridges

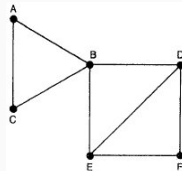
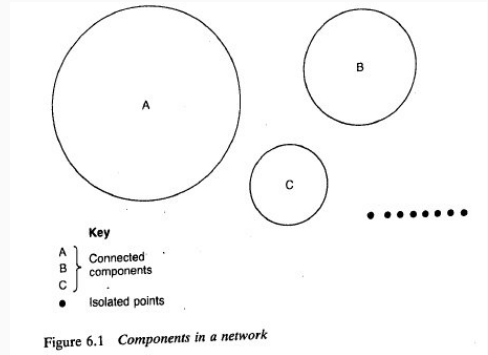


Was ist eine Clique?

- Ein „inner circle“
- „pocket of high density“
- Eine kohesive Subgruppe
- Im engsten Sinne: eine „kompletter“ und „maximaler“ Sub-Graph

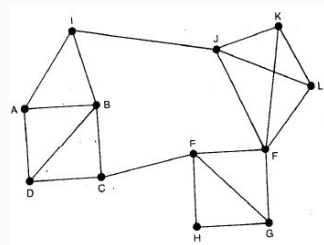
Component

- Maximal
- Connected
- Sub-Graph
- Starke vs. schwache Komponenten
- Innere Struktur
 - Zyklische Komponenten (3- oder 4-Ringe)
 - Knots and cut-points



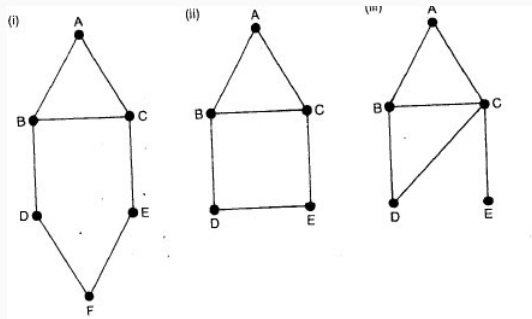
Nesting & k-cores

- Nesting: Zerlegen von Komponenten durch zusehends restriktivere Kriterien
- k -core: alle Mitglieder haben degree von mindestens k
- Schrittweises Erhöhen von $k \rightarrow$ core schrumpft/zerfällt $\rightarrow$ Rückschluß auf Netzwerkstruktur
- (m -cores: basieren auf Wert/Gewicht der Beziehung)



Cliquen

- Maximal und komplett (vs. components), starke und schwache Cliquen
- Beide extrem selten
- n -cliques: Mitglieder maximal n Beziehungen voneinander entfernt (für $n > 2$ nicht sehr informativ)



- Viele schwächere Definitionen, z.B. *k*-Plex: Knoten ist mit $n - k$ Mitgliedern der Gruppe direkt verbunden
- Soziale Kreise: überlappende Cliques (Anteil gemeinsamer Mitglieder)

Zusammenfassung

- **SNA: Perspektive**
 - Beziehungen zwischen Akteuren manchmal wichtiger als deren Eigenschaften
 - Beziehungsnetzwerk = strukturelle Eigenschaft des sozialen Systems
 - Klassische Sozialwissenschaftliche Betrachtungsweise
- **SNA: Verfahren**
 - Besondere Daten → besondere Verfahren
 - Ähnlich wie Regression (weitgehend) agnostisch gegenüber theoretischen Annahmen
 - Grundlagen seit acht Jahrzehnten bekannt
 - In jüngerer Zeit (ca. 25 Jahre) extrem dynamische Entwicklung