

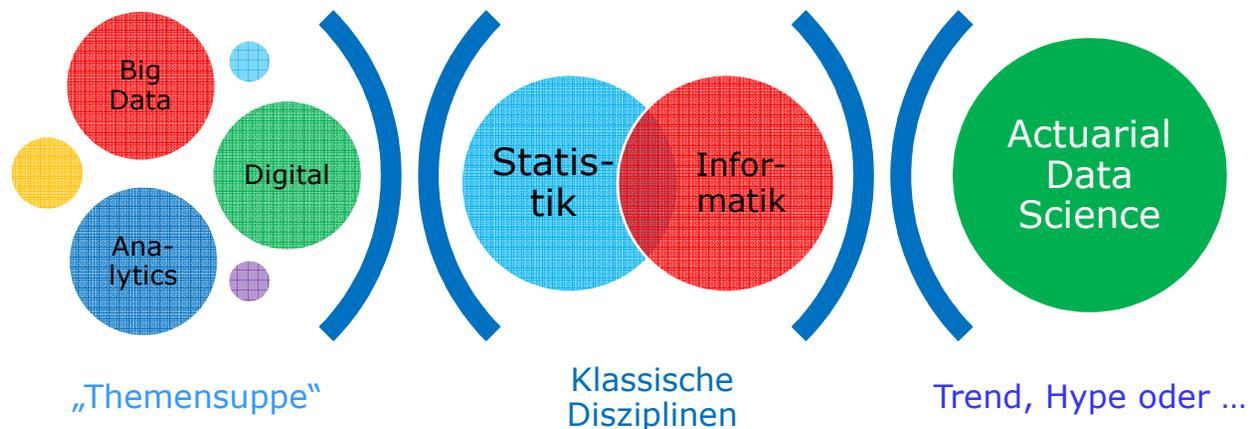
# Actuarial Data Science in Versicherungen – mehr als nur Statistik und Informatik

Treffpunkt AMMO, FH Bielefeld, 07. Dezember 2017

Dr. Stefan Nörtemann, msg life central europe gmbh  
Leiter der DAV-Fachgruppe „Actuarial Data Science“

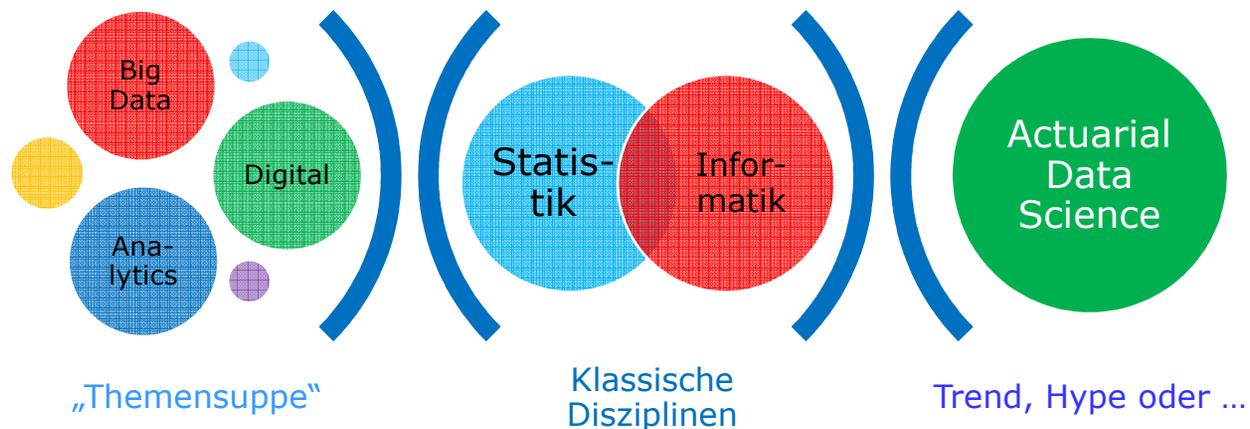
# Agenda

1. Moderne Zeiten – Big Data, Digitalisierung, Analytics
2. Actuarial Data Science
3. Business Cases für Versicherungen – Überblick
4. Ethik, Datenschutz, Wertschöpfung
5. Ausgewählte Business Cases für Versicherungen – Einblick
6. Hype um Big Data & Co. in der Versicherungswirtschaft – Ausblick



# Agenda

1. **Moderne Zeiten – Big Data, Digitalisierung, Analytics**
2. Actuarial Data Science
3. Business Cases für Versicherungen – Überblick
4. Ethik, Datenschutz, Wertschöpfung
5. Ausgewählte Business Cases für Versicherungen – Einblick
6. Hype um Big Data & Co. in der Versicherungswirtschaft – Ausblick



# Big Data, Analytics & Co. – ausgewählte Zitate

Die Themen Big Data, Insurance Analytics, etc. werden (nicht nur) unter Aktuaren durchaus kontrovers diskutiert!

## Die Euphorie ist eher ungleich verteilt!

„Hoffentlich bin ich schon in Rente, wenn es richtig losgeht.“

„Das sind doch alles Gedankenspiele – Geld verdienen kann man damit sicher nicht!“

„Schade, dass ich nicht (noch mal) 30 bin – ich wäre sicher der neue Zuckerberg der Versicherungsbranche.“



## Jeder versteht etwas anderes darunter!

„Big Data ist doch alles über Tera-Byte – oder?“

„Heutzutage ist doch alles irgendwie Big Data!“



## Das Heilsversprechen Big Data ist durchaus umstritten!

„Big Data ist die Rettung der Versicherungsbranche!“

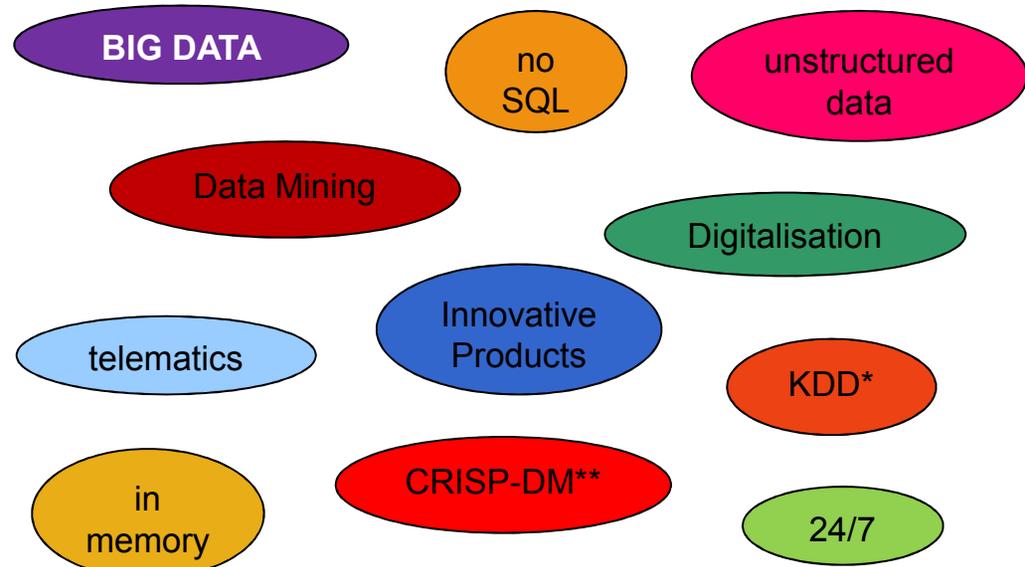
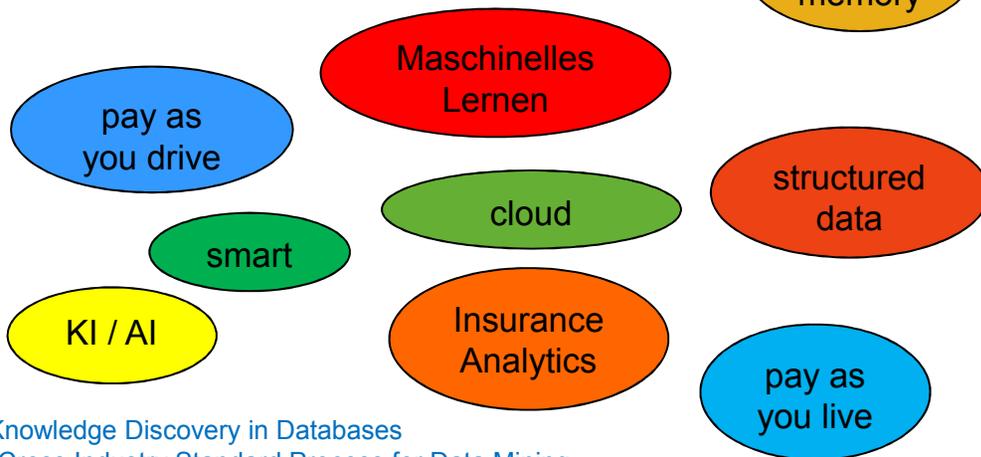
„Big Data ist doch nur eine Luftnummer!“

„Big Data bedeutet das Ende der Versicherung, wie wir sie heute kennen!“

# Vielfalt der Begriffe



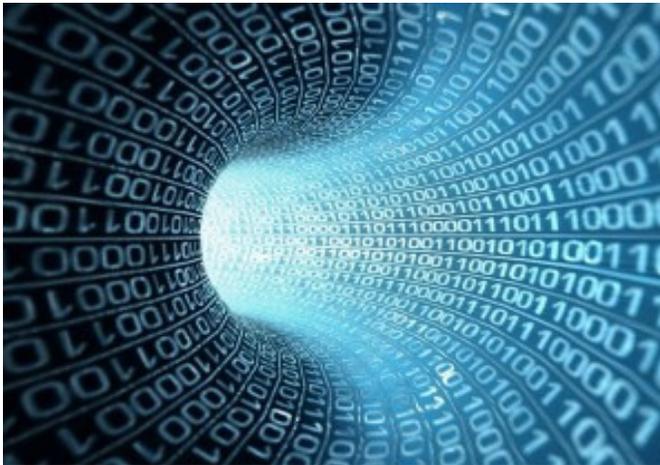
EL 70% DE LAS EMPRESAS INVERTIRÁ EN BIG DATA DURANTE LOS PRÓXIMOS DOS AÑOS von Marcos Gasparutti unter CC BY-SA 2.0



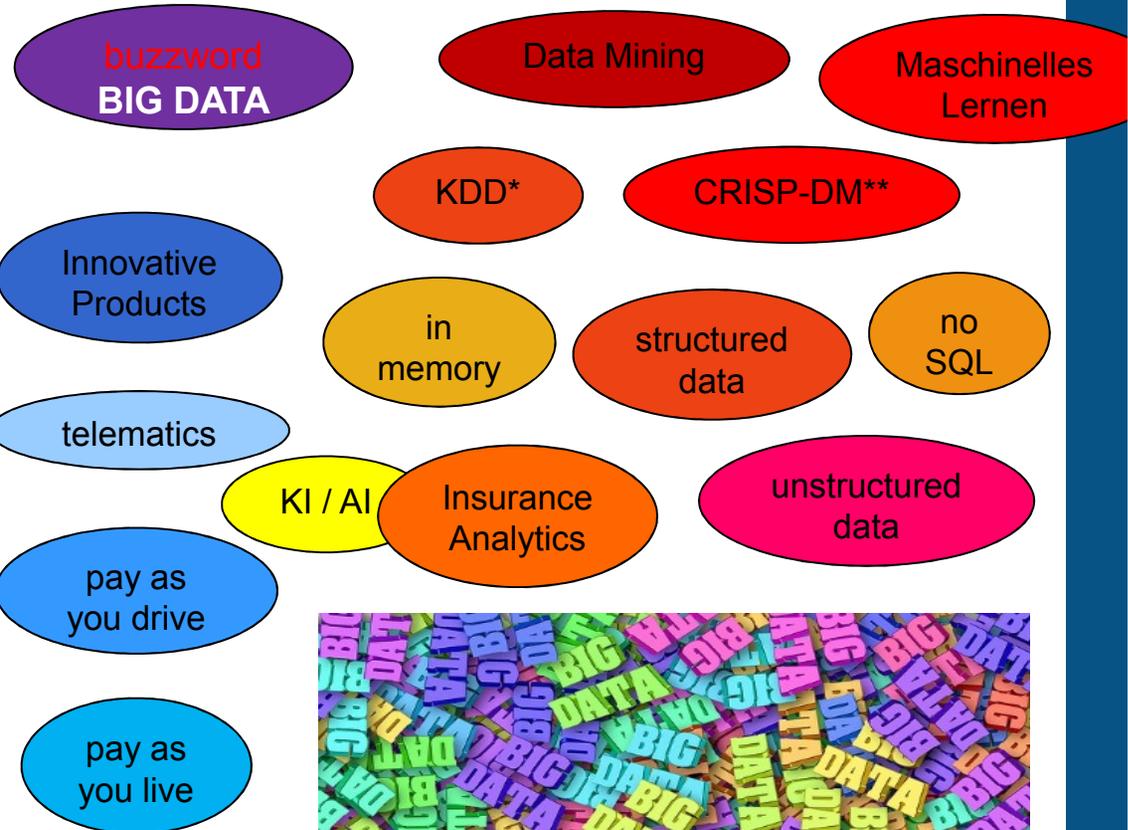
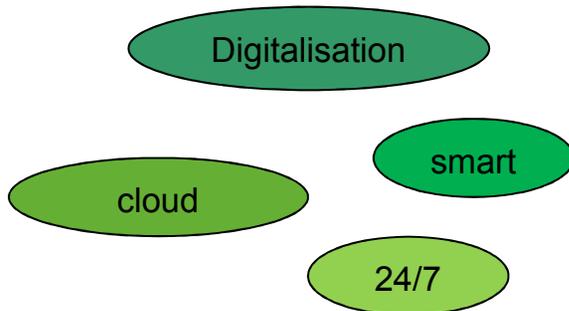
\*) Knowledge Discovery in Databases

\*\*\*) Cross Industry Standard Process for Data Mining

# Vielfalt der Begriffe – Versuch einer Ordnung



EL 75% DE LAS EMPRESAS INVERTIRÁ EN BIG DATA DURANTE LOS PRÓXIMOS DOS AÑOS von Marcos Gasparutti unter CC BY-SA 2.0

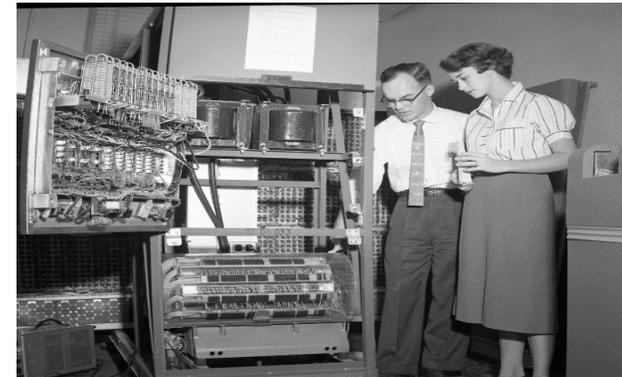


\*) Knowledge Discovery in Databases  
\*\*) Cross Industry Standard Process for Data Mining

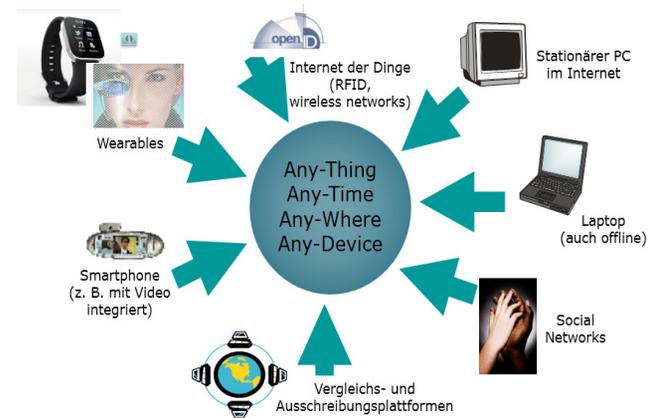
# Digitalisierung

## Digitalisierung =

- *(eigentlich, im engeren Sinne)*: Überführung analoger Größen in diskrete Werte, zu dem Zweck, sie elektronisch zu speichern oder zu verarbeiten
- Dies geschieht seit mehr als 60 Jahren in der Versicherungswirtschaft!
- *(im weiteren Sinne)*: Die durch das Internet geschaffenen Möglichkeiten, Daten zu jeder Zeit an jedem Ort abzurufen, weiterzuverarbeiten und zu speichern.
- *(für die Versicherungsbranche)*: Möglichkeiten & Anforderungen
  - ❖ an die Vertriebswege (Internet, Smartphones)
  - ❖ weiterer Services (z.B. Leistungsabrechnung)
  - ❖ Kommunikation mit dem Kunden (Stichwort 24/7)
  - ❖ ...



## Systems of Engagement (SoE)



# Data Mining & Insurance Analytics

## Data Mining =

- *(allgemein)*: Gewinnung von Wissen aus Daten
- *(konkret)*: Analyse großer Datenbestände (Big Data!)
  - ❖ (häufig) auf Basis statistischer Methoden, maschinellem Lernen, etc.
  - ❖ mit dem Ziel aus den Daten neue Erkenntnisse (Korrelationen, Trends, Muster) zu gewinnen



## Insurance Analytics =

- Data Mining mit versicherungsspezifischer Fragestellung
- Möglichkeit & Anforderung an die Auswertung großer Datenbestände für verschiedene Zwecke
  - ❖ Kundenverhalten
  - ❖ Marketing
  - ❖ Preisgestaltung
  - ❖ ...

Insurance  
Analytics



# Innovative Produkte

## Big Data ermöglicht neue Produkte:

- Individualisierung bzgl. des Zeitraums und des Beginns einer Versicherung
  - ❖ „spontane“ Policierung via Smartphone-App (*pay per use*)
  - ❖ Versicherung für einen kurzen Zeitraum
    - z.B. Unfall-Versicherung für Karneval
- Individualisierung bzgl. des Beitrags
  - ❖ *pay as you drive / Telematik*
  - ❖ *pay as you live*

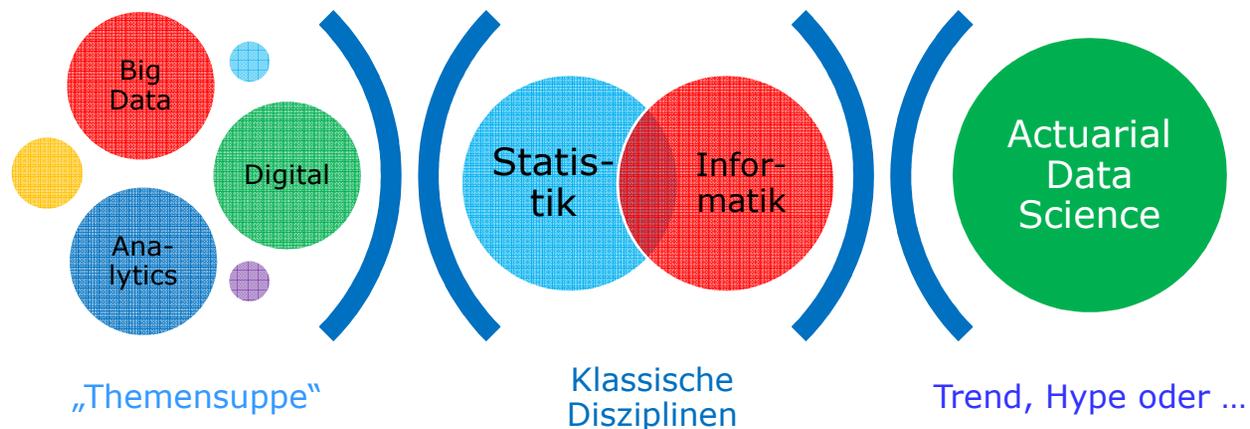
## Big Data erfordert neue Produkte:

- Versicherung für selbstfahrende Autos
- Schutz gegen neue Risiken: Cyber Risks, ...
- ...



# Agenda

1. Moderne Zeiten – Big Data, Digitalisierung, Analytics
- 2. Actuarial Data Science**
3. Business Cases für Versicherungen – Überblick
4. Ethik, Datenschutz, Wertschöpfung
5. Ausgewählte Business Cases für Versicherungen – Einblick
6. Hype um Big Data & Co. in der Versicherungswirtschaft – Ausblick



# Data Scientist



EL 75% DE LAS EMPRESAS INVERTIRÁ EN BIG DATA DURANTE LOS PRÓXIMOS DOS AÑOS von Marcos Gasparutti unter CC BY-SA 2.0

Ein Data Scientist ist jemand, der mehr über Statistik weiß als ein Informatiker und mehr über Informatik als ein Statistiker. Joel Grus

DATA

## Data Scientist: The Sexiest Job of the 21st Century

by Thomas H. Davenport and D.J. Patil

FROM THE OCTOBER 2012 ISSUE

aus: Harvard Business Review, Oktober 2012



Ken Duken als **Data Scientist** David Bogmann im Tatort „HAL“, ARD, 28.08.2016

# Data Science



EL 75% DE LAS EMPRESAS INVERTIRÁ EN BIG DATA DURANTE LOS PRÓXIMOS DOS AÑOS von Marcos Gasparutti unter CC BY-SA 2.0

Informatik /  
Programmierung

Mathematik &  
Statistik

Experten- / Fachwissen

Visualisierung &  
Kommunikation

Gesellschaftliches  
& rechtliches  
Umfeld

## Data Science =

- Informatik, Programmierung, Datenmanagement
- & Mathematik, Statistik, Stochastik
- & Expertenwissen in einem spezifischen Fachgebiet
- & gesellschaftlicher, ethischer und rechtlicher Rahmen im Fachgebiet
- & Kommunikation, Visualisierung, adressatenbezogene Ergebnispräsentation

# Actuarial Data Science



Informatik /  
Programmierung

Mathematik &  
Statistik

Aktuarielles Fachwissen

Visualisierung &  
Kommunikation

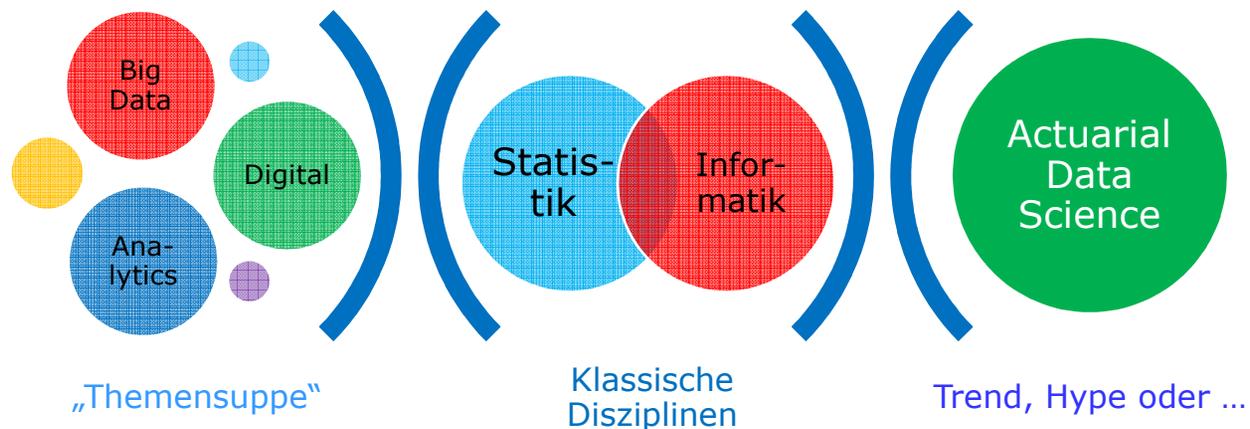
Gesellschaftliches  
& rechtliches  
Umfeld

## Actuarial Data Science =

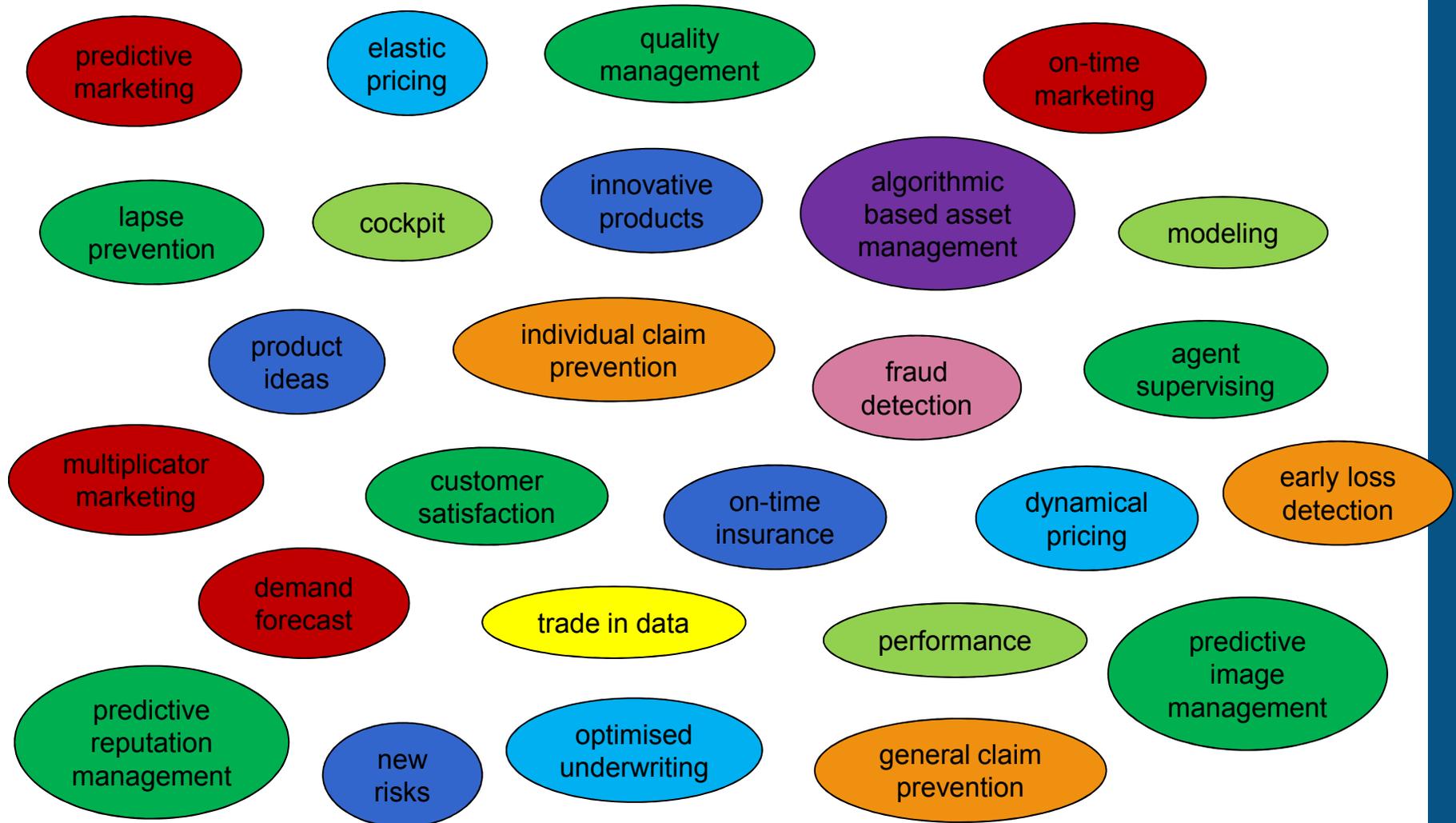
- Informatik, Programmierung, Datenmanagement
- & Mathematik, Statistik, Stochastik
- & **Aktuarielles / versicherungsspezifisches Fachwissen**
- & gesellschaftlicher, ethischer und rechtlicher Rahmen im **Versicherungsumfeld**
- & Kommunikation, Visualisierung, adressatenbezogene Ergebnispräsentation

# Agenda

1. Moderne Zeiten – Big Data, Digitalisierung, Analytics
2. Actuarial Data Science
- 3. Business Cases für Versicherungen – Überblick**
4. Ethik, Datenschutz, Wertschöpfung
5. Ausgewählte Business Cases für Versicherungen – Einblick
6. Hype um Big Data & Co. in der Versicherungswirtschaft – Ausblick

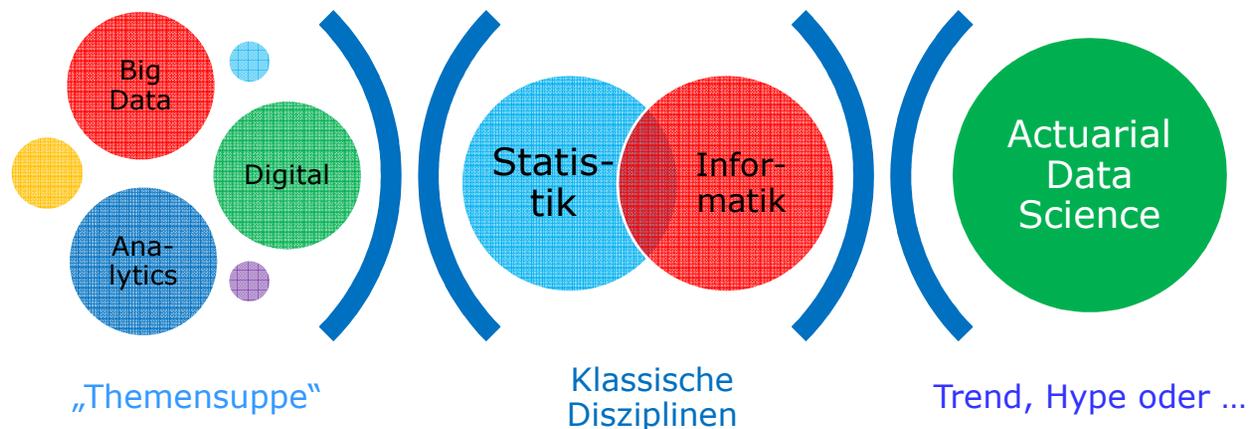


## 26 Business cases (Auswahl)



# Agenda

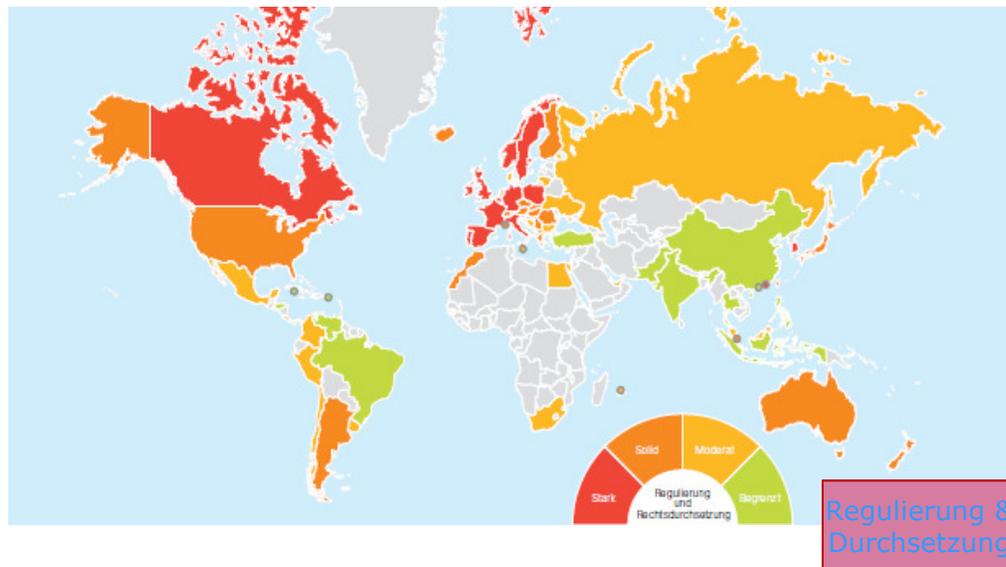
1. Moderne Zeiten – Big Data, Digitalisierung, Analytics
2. Actuarial Data Science
3. Business Cases für Versicherungen – Überblick
- 4. Ethik, Datenschutz, Wertschöpfung**
5. Ausgewählte Business Cases für Versicherungen – Einblick
6. Hype um Big Data & Co. in der Versicherungswirtschaft – Ausblick



# Datenschutz

## Datenschutz =

- Schutz vor missbräuchlicher Datenverarbeitung
- Schutz des Rechts auf informationelle Selbstbestimmung
- Aufgrund der technologischen Entwicklung im Zusammenhang mit der Digitalisierung gewinnt der Begriff der Datenschutz zunehmend an Bedeutung ...
- ... jedoch nicht überall auf der Welt



# Datenschutz in der Europäischen Union – heute

*DIRECTIVE 95/46/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL  
of 24 October 1995  
on the protection of individuals with regard to the processing of personal data and on the free  
movement of such data*

- Die Richtlinie 95/46/EG definiert einen Mindeststandard für den Datenschutz
- Die Verarbeitung personenbezogener Daten wird dort **grundsätzlich (!) verboten**.
- Es gibt lediglich **zwei Ausnahmen**:
  - ❖ ausdrückliche Zustimmung der betroffenen Person
  - ❖ falls wichtige Gründe vorliegen
- Es bestand die Pflicht zur Umsetzung der Richtlinie in nationales Recht.

*DIRECTIVE 2002/58/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL  
of 12 July 2002  
concerning the processing of personal data and the protection of privacy in the electronic commu-  
nications sector (Directive on privacy and electronic communications)*

- Ergänzung für elektronische Kommunikation in Directive 2002/58/EC

# Datenschutz-Grundverordnung

Die neue **Datenschutz-Grundverordnung** der EU

- soll die Richtlinien 95/46/EC und 2002/58/EC (Datenschutzrichtlinien) ersetzen
- wird ohne Umsetzungsakt (!) in gleicher Weise unmittelbar in allen EU-Mitgliedsstaaten gelten
- wurde im April 2016 vom Europäischen Parlament und dem EU-Ministerrat beschlossen
- wird nach 24 Monaten in Kraft treten, also Mitte des Jahres 2018

REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

of 27 April 2016

on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

- wird zu EU-weit einheitlichen Standards im Datenschutz führen (keine Datenschutzinseln mehr)
- wird insgesamt zu strengeren Regeln führen; in einzelnen Staaten (auch in Deutschland) wird eine Abschwächung des Datenschutzes befürchtet

# Datenschutz-Grundverordnung – Grundsätze

Grundsätze werden in Artikel 5 definiert:



Erweiterung der **Ausnahmen** vom Verbot der Verarbeitung personenbezogener Daten:

- Zustimmung der betroffenen Person
- Verarbeitung ist für die Erfüllung eines Vertrags, dessen Vertragspartei die betroffene Person
- zur Erfüllung einer rechtlichen Verpflichtung
- um lebenswichtige Interessen der betroffenen Person zu schützen
- zur Erfüllung einer Aufgabe, die im öffentlichen Interesse liegt
- zur Wahrung der berechtigten Interessen des Verantwortlichen oder eines Dritten

## Code of Conduct Datenschutz (CoC)

- Selbstverpflichtung der deutschen Versicherungswirtschaft:  
**„Verhaltensregeln für den Umgang mit personenbezogenen Daten durch die deutsche Versicherungswirtschaft“**
- initiiert vom  
***Gesamtverband der deutschen Versicherungswirtschaft e.V.***
- veröffentlicht im März 2013
- fast alle deutschen Versicherungsunternehmen sind dem CoC beigetreten
- Zweck ist die Konkretisierung der Datenschutzbestimmungen in Bezug auf das Versicherungswesen.
- Die Unternehmen gewährleisten als verantwortliche Stellen, dass die Anforderungen des Datenschutzes und der Datensicherheit beachtet werden (Selbstverantwortung).



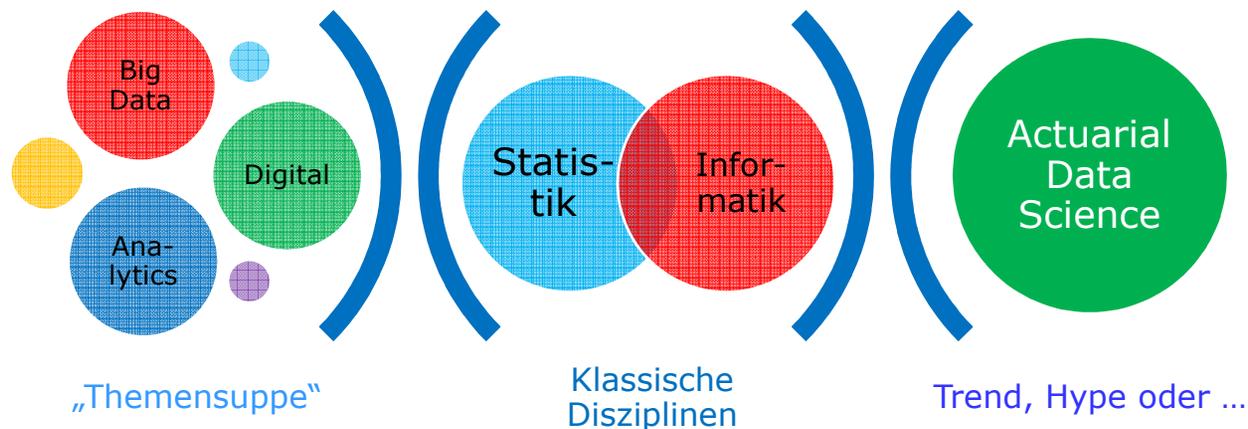
# Code of Conduct Datenschutz (CoC) – Prinzipien

## Die wichtigsten Prinzipien (Auswahl):

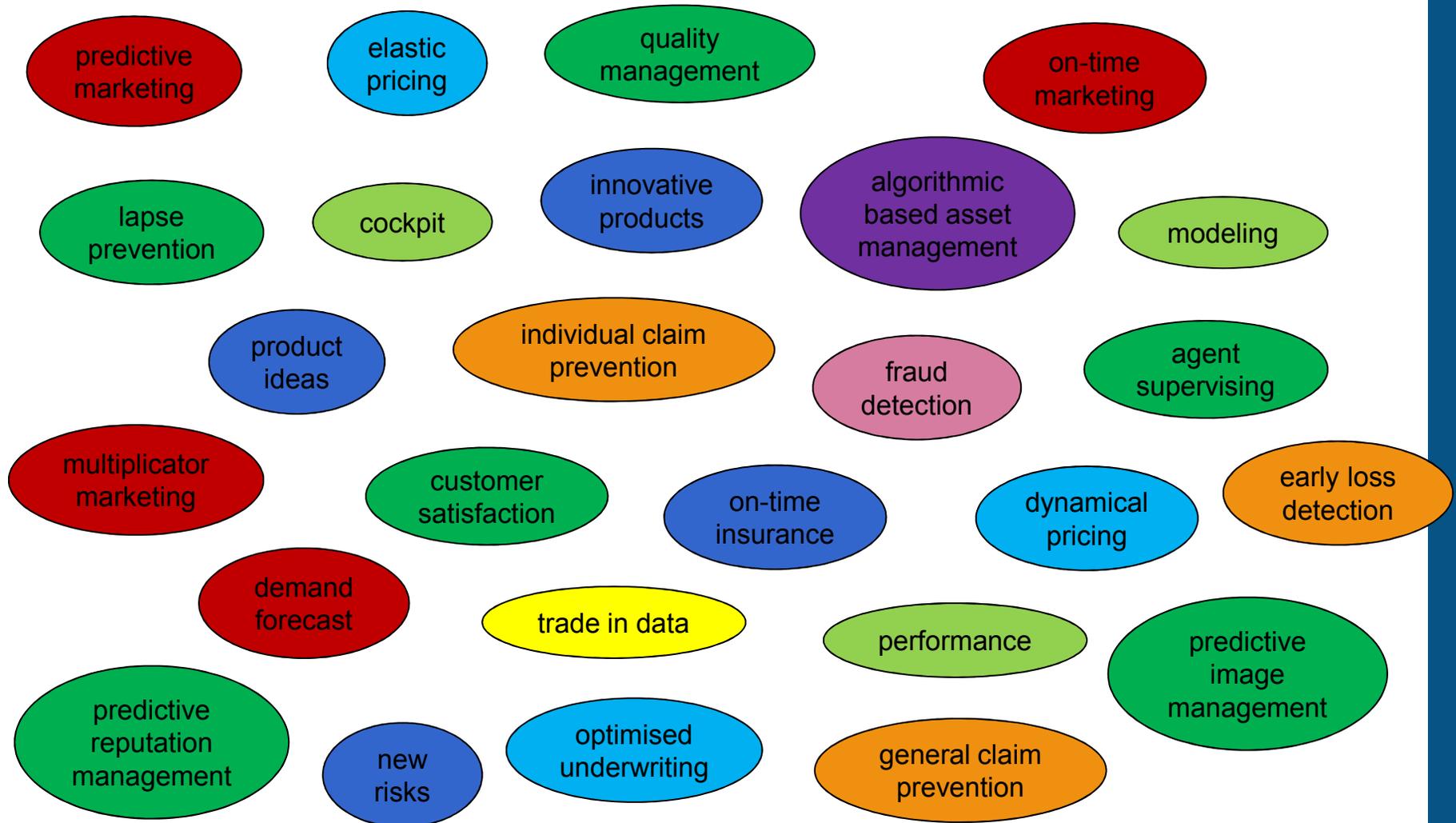
- **aus Artikel 2 (1):** Die Erhebung, Verarbeitung oder Nutzung personenbezogener Daten erfolgt grundsätzlich **nur**, soweit dies zur Begründung, Durchführung oder Beendigung eines Versicherungsverhältnisses erforderlich ist.
- **aus Artikel 2 (2):** Die personenbezogenen Daten werden grundsätzlich im Rahmen der den Betroffenen bekannten Zweckbestimmung verarbeitet oder genutzt.
- **aus Artikel 6:** Die Verarbeitung besonderer Arten personenbezogener Daten (zum Beispiel Gesundheitsdaten) bedürfen der gesonderten expliziten Zustimmung des Betroffenen.
- **aus Artikel 7:** Personenbezogene Daten werden grundsätzlich bei den Betroffenen selbst erhoben. (Ausnahmen in engen Grenzen: **Artikel 8 & 14 (für HIS)**)
- **aus Artikel 18:** Für Zwecke der Werbung dürfen personenbezogene Daten nur mit expliziter Zustimmung der Betroffenen verwendet werden. (vgl. § 28 BDSG).
- **aus Artikel 23:** Auskunftsanspruch (über die gespeicherten Daten)
- **aus Artikel 24:** Anspruch auf Berichtigung (unrichtiger oder unvollständiger Daten) sowie Löschung und Sperrung (unzulässig erhobener Daten)

# Agenda

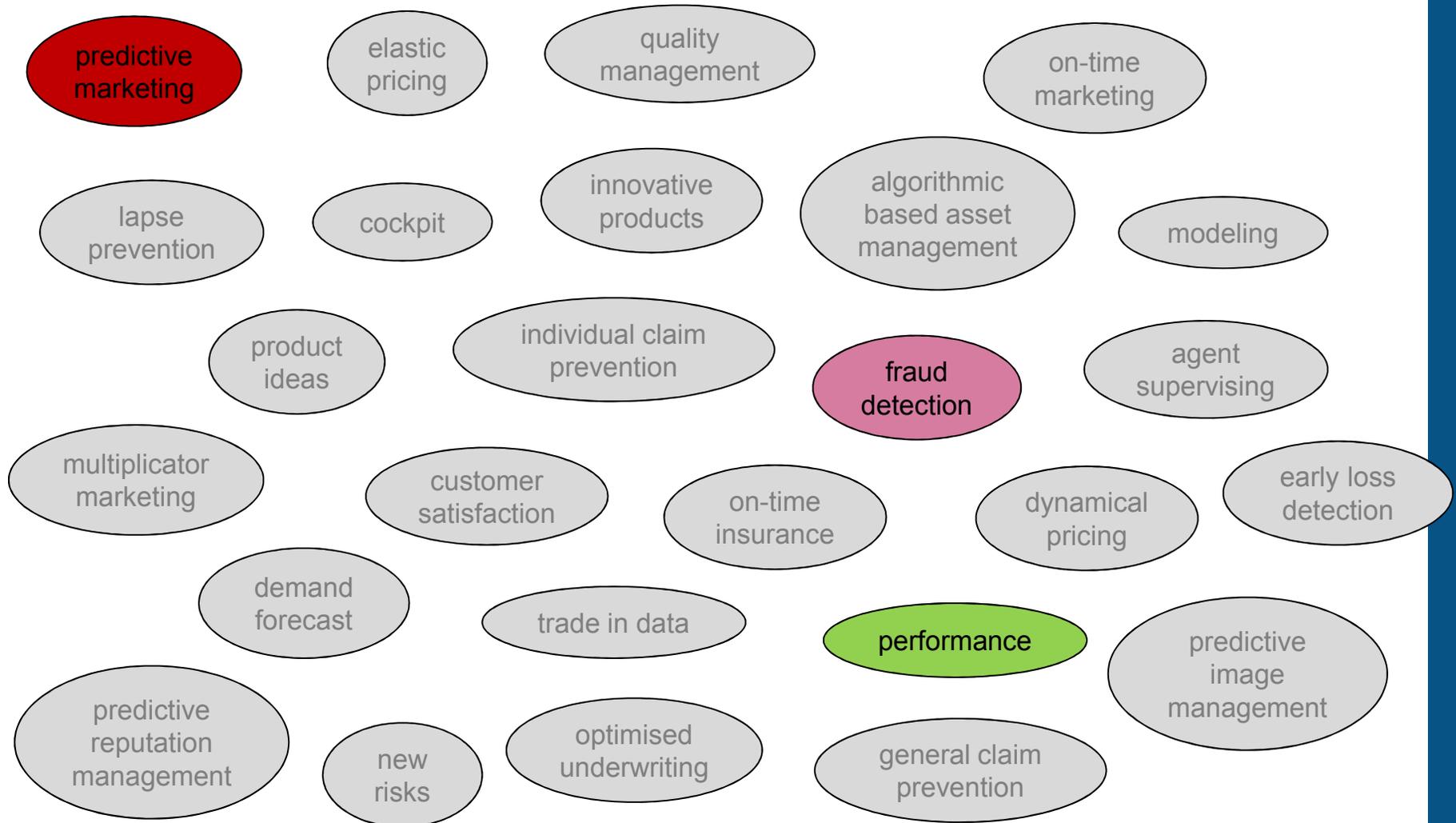
1. Moderne Zeiten – Big Data, Digitalisierung, Analytics
2. Actuarial Data Science
3. Business Cases für Versicherungen – Überblick
4. Ethik, Datenschutz, Wertschöpfung
- 5. Ausgewählte Business Cases für Versicherungen – Einblick**
6. Hype um Big Data & Co. in der Versicherungswirtschaft – Ausblick



## 26 Business cases (Auswahl)



## 26 Business cases (Auswahl)



# Predictive Marketing

predictive  
marketing

## **Predictive marketing =**

- Erkennung von Kaufmustern, gezielte individualisierte Werbung, gezielte individuelle Rabatte, individuelle Kaufanreize
- Bekannt durch 
- **Methode:** Insurance Analytics basierend auf unterschiedlichen Datenquellen
- **Datenbasis** verschiedene Ebenen ([Beispiel](#)):
  1. Kundeninformationen (aus bestehenden Verträgen / Kundenbeziehungen)
  2. Externe öffentliche Informationen
  3. Daten aus sozialen Medien
  4. Verhalten im Internet
- **Aufgabe:** Korrelationen zum Kaufverhalten anderer Kunden finden
- **Ziel:** Individualisierte Werbung und Angebote

# Fiktives Fallbeispiel

predictive  
marketing

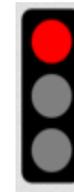
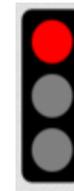
- **level 1 (nur Kundeninformationen):** Herr X., 50 Jahre alt, Kunde seit 10 Jahren, Risikolebensversicherung, VS = 50.000 €
  - ⌘ 67% der 50 Jahre alten Kunden mit einer Risikolebensversicherung haben zusätzlich eine Berufsunfähigkeitsversicherung -> spezielles Angebot an Herrn X.?
- **level 2 (zusätzl. öffentliche Informationen):** Aktuar seit 15 Jahren und verdient (wahrscheinlich) mehr als 100.000 € im Jahr\*
  - ⌘ 88% der vergleichbaren Kunden haben mindestens eine Versicherungssumme in Höhe von 250.000 € -> Angebot?
- **level 3 (zusätzl. Daten sozialer Medien):** Vater dreier Kinder und liebt Fallschirmspringen\*\*
  - ⌘ 68% der vergleichbaren Kunden haben eine Unfallversicherung -> spezielles Angebot?
- **level 4 (zusätzl. Surfverhalten):** Herr X. meistbesuchte Internetseite des letzten Monats ist **Porsche**\*\*\*
  - ⌘ 42% der vergleichbaren Kunden haben eine dread disease Versicherung -> spezielles Angebot

\*) Quelle: DAV Gehaltsstudie 2014

\*\*\*) Quelle: Facebook



\*\*\*) Quelle: Google Analytics



## Datenschutz:

- ◆ Nur mit expliziter Zustimmung!
  - Artikel 18 CoC, § 28 BDSG
- ◆ **Konflikt mit:**
  - Artikel 2 CoC
  - ❖ nur zur Begründung, Durchführung oder Beendigung eines Versicherungsverhältnisses
  - ❖ im Rahmen der den Betroffenen bekannten Zweckbestimmung
- Artikel 7 CoC
  - ❖ grundsätzlich vom Betroffenen
- § 3a BDSG
  - ❖ Datenvermeidung & Datensparsamkeit

# Fiktives Fallbeispiel

predictive  
marketing

- **level 1 (nur Kundeninformationen):** Herr X., 50 Jahre alt, Kunde seit 10 Jahren, Risikolebensversicherung, VS = 50.000 €
  - ⌘ 67% der 50 Jahre alten Kunden mit einer Risikolebensversicherung haben zusätzlich eine Berufsunfähigkeitsversicherung -> spezielles Angebot an Herrn X.?
- **level 2 (zusätzl. öffentliche Informationen):** Aktuar seit 15 Jahren und verdient (wahrscheinlich) mehr als 100.000 € im Jahr\*  
\*) Quelle: DAV Gehaltsstudie 2014
  - ⌘ 88% der vergleichbaren Kunden haben mindestens eine Versicherungssumme in Höhe von 250.000 € -> Angebot?
- **level 3 (zusätzl. Daten sozialer Medien):** Vater dreier Kinder und liebt Fallschirmspringen\*\*  
\*\*\*) Quelle: Facebook
  - ⌘ 68% der vergleichbaren Kunden haben eine Unfallversicherung -> spezielles Angebot?
- **level 4 (zusätzl. Surfverhalten):** Herr X. meistbesuchte Internetseite des letzten Monats ist **Porsche**\*\*\*  
\*\*\*) Quelle: Google Analytics
  - ⌘ 42% der vergleichbaren Kunden haben eine dread disease Versicherung -> spezielles Angebot



# Zum Beispiel: Betrugserkennung

fraud  
detection

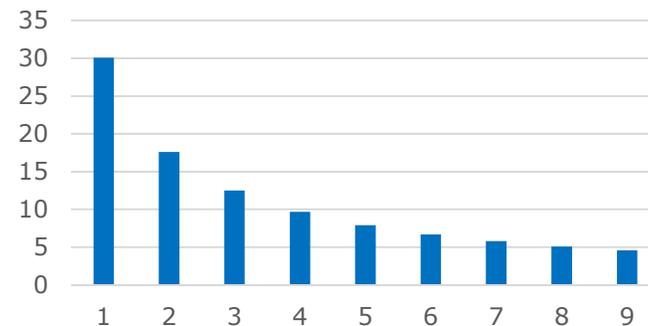
Statistische Betrugsdetektion ist ein „weites Feld“, in dem ausgefeilte statistische Methoden\* zur Anwendung kommen

- Ein einfaches univariates Konzept ist die Ziffernanalyse zur Auffälligkeitsdetektion
- Für „Basic“ geeignet (**Beispiel**): Analyse führender Ziffern gegen das **Newcomb-Benford's Law (NBL)**

Genügt eine Menge von Dezimalzahlen empirischer Daten der Benford-Eigenschaft („Benford-Variable“), so ist die Wahrscheinlichkeit  $P(d)$  für das Auftreten der Ziffer  $d$  an erster Stellen gleich:

$$P(d) = \log_{10}\left(1 + \frac{1}{d}\right)$$

Benford-Verteilung



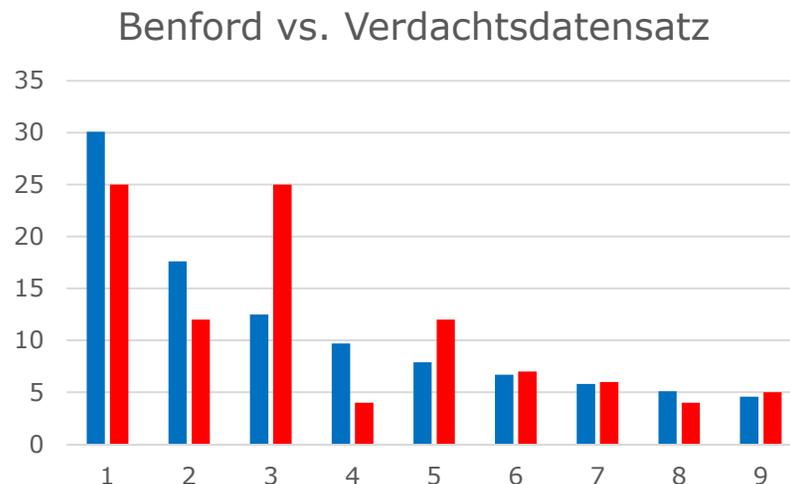
\*) siehe zum Beispiel: **Prof. Dr. Peter Ruckdeschel**: „Besser / Anders als Sherlock Holmes – was die moderne Statistik für die Betrugserkennung in der Versicherung leisten kann. . . (und was nicht)“; Tagung der Fachgruppe Versicherungsmathematik im deutschen Verein für Versicherungswissenschaft, Köln, 16. November 2016

# Beispiel aus der Praxis

fraud  
detection

## Fallbeispiel: Kasko-Versicherung (Kfz)\*

- Überprüfung von Schadenzahlungen für definierte Zeitintervalle auf Ebene der Sachbearbeiter / Gruppe
- Vergleich der führenden Ziffer gegen das **Newcomb-Benford's Law**



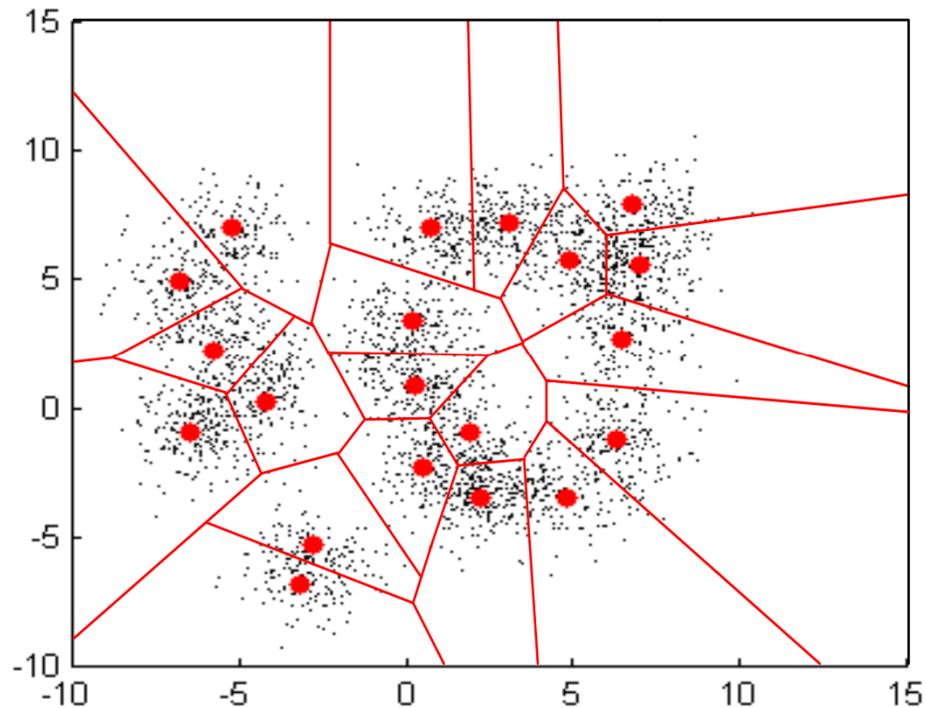
Der Vergleich legt nahe, dass die Schadenhöhen **schlecht** erfunden sein könnten.

- ➔ Suche nach Gründen für eine Verletzung der Benford-Eigenschaft
- ➔ weitere Untersuchungen

\*) realer Fall aus der Praxis eines großen deutschen Versicherungsunternehmens

# Publikumsfrage

⌘ Was ist das / Was könnten sich dahinter verbergen?



# Beispiel: Clusterbasierte Bestandsverdichtung (1)

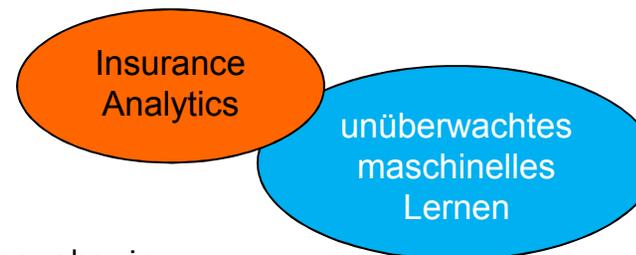
performance

**Herausforderung:** Laufzeiten bei Projektionsrechnungen

- Zum Beispiel: mittelgroßer Lebensversicherungsbestand (1-2 Mio. Verträge), stochastische Projektion (10.000 Pfade), Projektion über 60 Jahre, monatliche Projektionsschritte, dynamische Managementregeln
- ➔ Laufzeit bis zu mehreren Tagen!

**Lösungsansatz:** Bestandsverdichtung

- **„Trick“** der Verdichtung ist die Erzeugung eines Teilbestandes (des Versichertenbestandes), der dieselben (oder ähnliche) Eigenschaften aufweist, wie der Ausgangsbestand.\*
- Verfahren zur Bestandsverdichtung:
  - ⌘ Aktuarielle (meist „händische“) Verfahren
  - ⌘ Stichprobenverfahren
  - ⌘ Analytische Verfahren
  - ⌘ **Clusterbasierte Verfahren**



\*) Das geht naturgemäß mit einem Informationsverlust einher aber in gewissem Rahmen ist dies (vielleicht) tolerierbar.

## Beispiel: Clusterbasierte Bestandsverdichtung (2)

performance

### Vorgehen (vereinfacht):

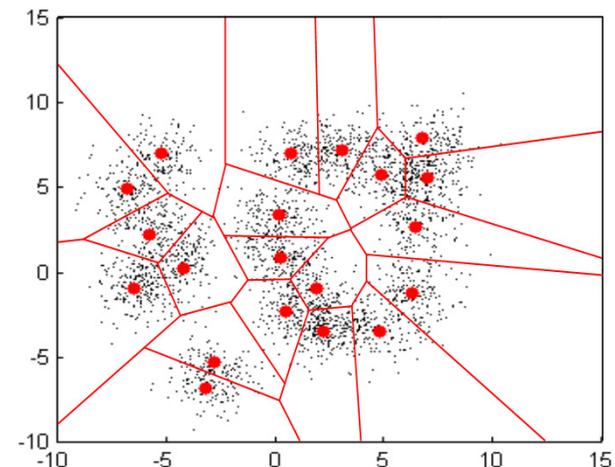
- Bestandsaufbereitung: Prüfung der Datenqualität, Bestandsbereinigung, ...
- Digitalisierung: Kodierung der Versicherungsverträge durch sog. Merkmalsvektoren
- Standardisierung  $(x - \mu_x) / \sigma_x$  und Gewichtung der Merkmale
- Wahl einer geeigneten Metrik
- ➔ Euklidische Vektorraum (Dimension = Anzahl der Merkmale)

### Aufgabe (vereinfacht):

- Partitionierung des Bestands in  $k$  Cluster, so dass die Summe der quadratischen Abstände jedes Vektors zum Clustermittelpunkt („Centroid“) minimal wird
- Existenz einer Lösung ist leicht beweisbar, aber das Problem ist **NP-schwer!** (NP-hard)

### Lösung (vereinfacht):

- Näherungsverfahren, z.B. k-means, o.a.



# K-means zur Clusterverdichtung

- Für Clusterberechnungen mit dem **K-Means-Algorithmus\***

SSE

- Seien:

- K = Anzahl Cluster
- N = Anzahl Merkmalsvektoren
- m = „Membership“ Zuordnung
- $c_i$  = Centroid des i-ten Clusters

➔ Minimierung der Abstände innerhalb der Cluster

$$\arg \min_{m(i,j)} \sum_{i=1}^k \sum_{j=1}^N m(i,j) \|x_j - c_i\|^2$$

$$m(i,j) : \{1, \dots, k\} \times \{1, \dots, N\} \mapsto \{0, 1\}$$

$$m(i,j) := \begin{cases} 1 & : x_j \in c_i \\ 0 & : \text{Sonst} \end{cases}$$

$$\sum_{j=1}^N m(i,j) = 1 \quad \text{für } i = 1, \dots, k$$

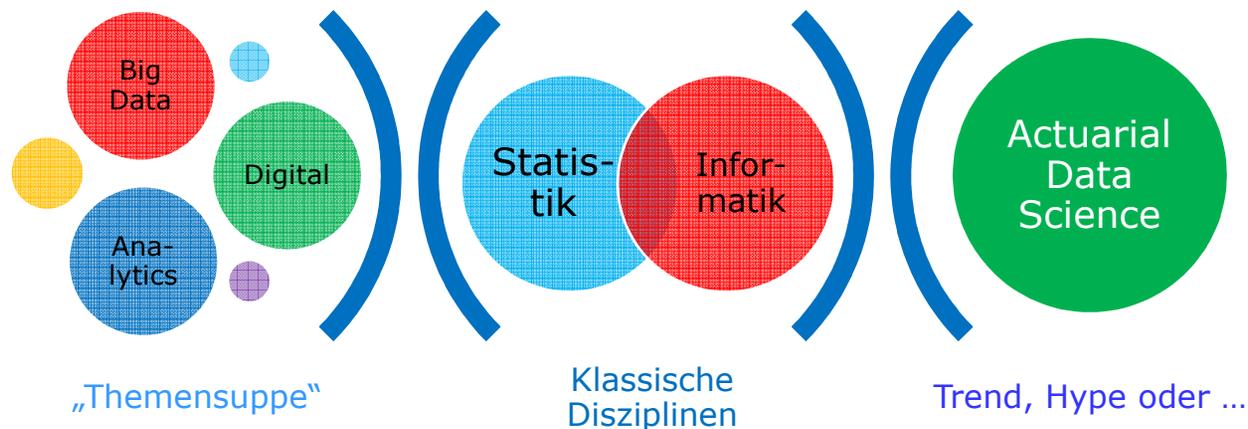
- Algorithmische Formulierung (vereinfacht)

- 1) Wähle K anfängliche Cluster-Centroide,  $c_1, c_2, c_3, \dots, c_K$
- 2) Für jeden Datenpunkt  $x$ , finde den nächsten (nearest neighbour) Cluster-Centroiden
- 3) Für jedes Cluster (gegeben durch vorherigen Schritt), aktualisiere den Cluster-Centroiden (bilde den Mittelwert über alle Cluster-Elemente, d.h. dem Cluster zugeordnete Datenpunkte)
- 4) Gehe zu (2) bis sich die Cluster nicht mehr ändern (z.B. SSE konstant) oder die Änderung nur noch minimal ist (vorgegebenes Epsilon)

\*) MacQueen, 1967

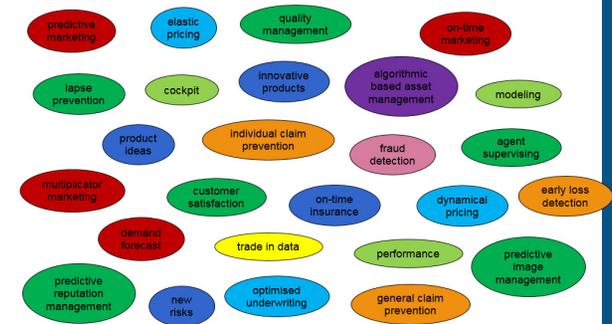
# Agenda

1. Moderne Zeiten – Big Data, Digitalisierung, Analytics
2. Actuarial Data Science
3. Business Cases für Versicherungen – Überblick
4. Ethik, Datenschutz, Wertschöpfung
5. Ausgewählte Business Cases für Versicherungen – Einblick
- 6. Hype um Big Data & Co. in der Versicherungswirtschaft –  
Ausblick**



# ADS – Hype oder Zukunft?

- Warum ist Actuarial Data Science ist aktuell populär?
  - ❖ technische Machbarkeit / neue technische Möglichkeiten
  - ❖ breiter (niedrigschwelliger) Zugang durch open-source Software
  - ❖ Analysen einfach(er) möglich (ohne tieferes Fachverständnis)
  - ❖ Vorbild: andere Branchen
  - ❖ ...
- Status im actuariellen Umfeld:
  - ❖ Einiges machen wir schon (seit langem), ohne es ADS zu nennen
  - ❖ Von unserem Selbstverständnis „fühlen“ wir uns zuständig für Mathe-Themen
- Künftige Entwicklungen:
  - ❖ Ideen für ADS-Anwendungen gibt es viele ->
  - ❖ neue Produkte werden sich verbreiten (PAYD, PAYL, ...)
  - ❖ Vertrieb, Marketing, Kommunikation mit dem Kunden sind im Wandel



**Aber:** Die große Disruption hat sich noch nicht gezeigt!

## Epilog: Innovative Products – Pay as you live

### Pay as you live - Wie sieht das konkret aus?

& Smart-Watch, wearable, apps, ...

- ❖ **fitness wearables** (z.B. Fitbit, Nike+, Runtastic, ...): sammeln & übertragen Fitness-Daten (Gegangene Schritte, Laufradius, Treppenstufen, ...)
- ❖ **medical wearables** (z.B. AliveCor, iBG Star ...): sammeln & übertragen medizinische Daten (Puls, Blutdruck, zugeführte/verbrannte Kalorien,...) in die cloud
- ❖ Alternativ oder ergänzend könnten zudem soziale Daten (Facebook-Einträge, Kaufverhalten, Umgang mit Dritten, etc.) über die versicherte Person gesendet werden.
- ❖ Aufgrund dieser Daten werden Risiko Zu-/Abschläge auf den Lebensversicherungstarif ermittelt, in beliebiger Taktung.



Aktivitätsapp  
Spüre das Erfolgserlebnis, wenn du die Ringe „Bewegen“, „Trainieren“ und „Stehen“ schließt.

## (Nicht nur) ein Witz

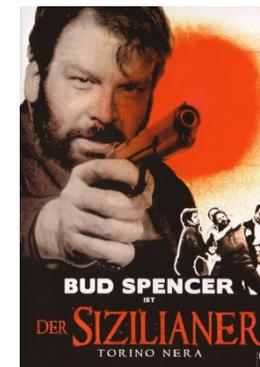
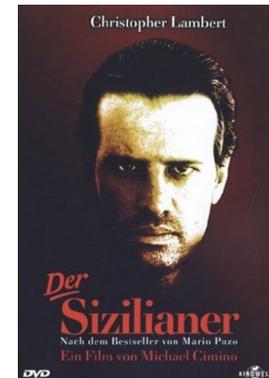
„Was unterscheidet den deutschen vom sizilianischen Aktuar?

Der deutsche Aktuar weiß ungefähr, wie viele Personen in seinem Bestand im kommenden Jahr versterben werden –

der sizilianische Aktuar weiß sogar wer!“

These: In Zukunft werden wir alle sizilianische Aktuare sein und wissen, wer aus unserem Bestand ausscheiden wird!

→ Wirklich: Rechnen mit der Zukunft!



## Literatur (1)

- ⌘ Alpaydin, Ethem; *Maschinelles Lernen*; Oldenbourg Wissenschaftsverlag, 2008
- ⌘ Banthorpe, Peter; *Modelling mortality using external or non-biometrical data*; DAV-Jahrestagung 2016 in Bremen, 57. Tagung der Leben-Gruppe, 28. April 2016
- ⌘ Bättig, Daniel; *Angewandte Datenanalyse*; Springer Spektrum, 2015
- ⌘ Ernst & Young, Australia; *Introducing ‚Pay As You Live‘ (PAYL) Insurance*; 08/2015 (EY-introducing-pay-as-you-live-payl-insurance.pdf)
- ⌘ Fondermann, Bernd; Spichale, Kai; George, Lars; *Big Data – Apache Hadoop*, entwickler.press, 2012
- ⌘ Gesamtverband der Deutschen Versicherungswirtschaft e.V. (GDV); *Verhaltensregeln für den Umgang mit personenbezogenen Daten durch die deutsche Versicherungswirtschaft*; 09/2012
- ⌘ Gesamtverband der Deutschen Versicherungswirtschaft e.V. (GDV); *Verzeichnis der Versicherungsunternehmen, die die Beitrittserklärung zu den Verhaltensregeln für den Umgang mit personenbezogenen Daten durch die deutsche Versicherungswirtschaft unterzeichnet haben*; 05/2015
- ⌘ Grus, Joel; *Einführung in Data Science: Grundprinzipien der Datenanalyse mit Python*; O'Reilly, 2016
- ⌘ ...

## Literatur (2)

- ⌘ ...
- ⌘ Hiendlmeier, Stefan & Hertting, Mark; *Auswirkungen der Digitalisierung auf die Steuerung von Versicherungsunternehmen*; White Paper, Horváth & Partners, 03/2015
- ⌘ Hoge, Wilfried & Stark Jochen; *Big Data Analytics für Aktuare*, Webinar 16-13 der Deutschen Aktuar Akademie, 28.06.2016
- ⌘ Kiermaier, Axel; *Kausalität digitalisiert (?)*; Deutsche Aktuar Akademie, max.99, Köln, 19. September 2017
- ⌘ Kohl, Tobias & Schlender, Mark; *Pflichten und Chancen des „Code of Conduct Datenschutz“*; Zeitschrift für Versicherungswesen 22 | 2013
- ⌘ Kurowski, Oliver; *NoSQL Einführung – CouchDB, MongoDB und Redis*, entwickler.press, 2012
- ⌘ Lobo, Sascha; *Die Mensch-Maschine: Bevormundet durch die Zahnbürste*; Spiegel Online, 12/2014
- ⌘ Makowski, Alexander; *Bedeutung und Nutzenpotenziale von Big Data für Versicherungsunternehmen*; Leipziger Masterarbeiten, Verlag Versicherungswirtschaft, Karlsruhe, 09/2016
- ⌘ Meck, Georg & Weiguny Bettina; *Disroption, Baby, Disruption*; Frankfurter Allgemeine Sonntagszeitung, 27. Dezember 2015
- ⌘ ...

## Literatur (3)

- ⌘ ...
- ⌘ Meckel, Miriam; *NEXT: Erinnerungen an eine Zukunft ohne uns*; Rowohlt, 2013
- ⌘ Nassehi, Armin (HG); *Kursbuch 177, Privat 2.0*; Murmann-Verlag, 03/2014
- ⌘ Neumann, Alf; *Digitalisierung und künstliche Intelligenz als Turbo für die Lebensversicherung?*; Versicherungsmathematisches Kolloquium der Ludwig-Maximilian Universität München; München, 29. Mai 2017
- ⌘ Nörtemann, Stefan; *Big Data & Sizilianische Aktuare*, WiMa-Kongress, Universität Ulm, 07. November 2015
- ⌘ Nörtemann, Stefan; *Big Data & Insurance Analytics versus Data Protection*, European Congress of Actuaries (eca 2016), Brüssel, 21. April 2016
- ⌘ Nörtemann, Stefan; *Structured and unstructured data – Insurance analytics, products and risk management of the future*; 3. Weiterbildungstag der DGVMF, Hannover, 16. Juni 2016
- ⌘ Nörtemann, Stefan; *Was ist dran am Hype um Big Data & Co. in der Versicherungsbranche?*, WiMa-Kongress, Universität Ulm, 12. November 2016
- ⌘ Ruckdeschel, Peter; *Besser / Anders als Sherlock Holmes – was die moderne Statistik für die Betrugserkennung in der Versicherung leisten kann. . . (und was nicht)*; Tagung der Fachgruppe Versicherungsmathematik im deutschen Verein für Versicherungswissenschaft, Köln, 16. November 2016
- ⌘ ...

## Literatur (4)

- ⌘ ...
- ⌘ Runkler, Thomas, A.; *Data Mining: Modelle und Algorithmen intelligenter Datenanalyse*, Springer Vieweg; Auflage 2, 2015
- ⌘ Schrempp, Mirko; *Big Data – Technologiegrundlagen*, entwickler.press, 2012
- ⌘ Swiss Re; *Lebensversicherung im digitalen Zeitalter: Ein grundlegender Wandel steht bevor*; sigma Nr. 6/2015, 11/2015
- ⌘ Tschirk, Wolfgang; *Statistik: Klassisch oder Bayes*; Springer Spektrum, 2014
- ⌘ Versicherungsforen Leipzig, adesso AG, Dortmund; *Studie: Geschäftsmodelle 4.0 – Was die Assekuranz von anderen Branchen lernen kann*; 11/2015
- ⌘ Weidner, Wiltrud; *Über Telematik zu Big Data - Überrollt uns die Datenflut oder lernen wir rechtzeitig schwimmen?*; Deutsche Aktuar Akademie, max.99, Köln, 19. September 2017
- ⌘ Winter, Fabian; *Einführung Big Data und statistische Modelle in der Versicherungswirtschaft*; Webinar 16-08 der Deutschen Aktuar Akademie, 11.02.2016
- ⌘ Zeh, Juli; *Der Selber-Schuld-Gedanke macht uns alle unfrei*; Süddeutsche Zeitung, 11/2014

Actuarial Data Science –  
mehr als nur Statistik und Informatik