

What Analytics can do in the Health Domain DeepHealth 2016

Dr. Silvia Quarteroni, Manager

Zurich, April 6th, 2016



AGENDA

What analytics can do in the health domain

1. About ELCA
2. What is analytics
 - I. From data to decisions
 - II. Deep content analytics
3. Analytics in the health domain
 - I. What can be done
 - II. Use cases: practitioners, patients, insurance
4. Conclusions

ABOUT ELCA

ELCA on one page

ELCA is one of the biggest independent Swiss full-service providers for business and technology solutions, and a leader in the fields of IT business consulting, software development and maintenance, and IT systems integration.

- **Founded** 1968 as electro-calcul for the software navigation of the Grand Dixence Dam
- **Employees** Over 700 employees, mostly highly skilled (IT) professionals with university degrees
- **Offices** Lausanne (headquarters), Zurich, Geneva, Bern, Madrid, Paris, Ho Chi Minh City (Vietnam)
- **Turnover** CHF 101,8 million in 2014 (+5% growth)
- **Quality Standards** ISO 9001 (since 1993), ISO 14001 (since 2011), CMMI level 3 (since 2007)
- **Project Experience** > 1'000 customer projects in ten years

- **Awards**



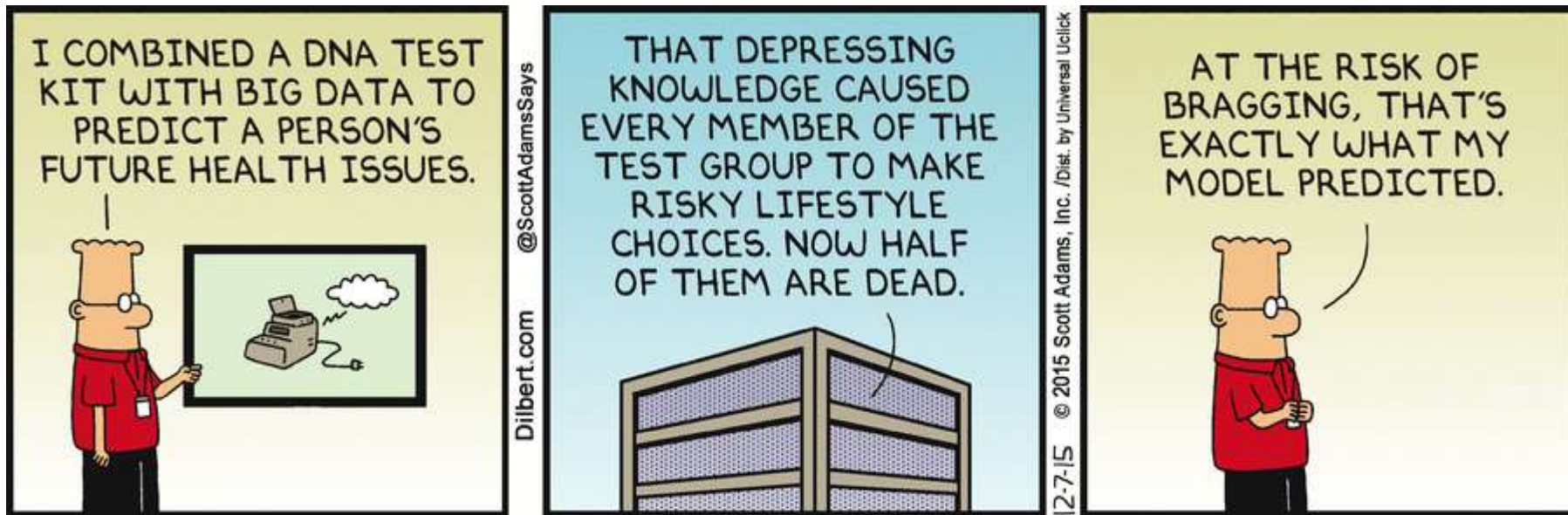
WHAT IS ANALYTICS

From data to decisions

Predictive Medicine Depends on Analytics

Healthcare

Seven ways predictive analytics can improve healthcare



Forbes / Tech

**Improve Healthcare,
Win \$3,000,000.**

APR 21, 2015 @ 10:50 AM 61,833 VIEWS

How Big Data Is Changing Healthcare

WHAT IS ANALYTICS

Deep content analytics: a key technology

Automatic categorisation

Automatically categorise or cluster your textual content according to your choice of classes thanks to machine learning.

Spoken language understanding

Recognise the language of audio files, then spot important keywords or fully transcribe it to text to make it searchable.

Image analysis

Recognise text even in poorly scanned documents, identify faces in pictures and more.

Automatic translation

Translate content into your favourite language; search for information in multiple languages.

Named entity recognition

Identify people, locations, organisations mentioned in text – no lists needed!

GIS and geocoding

Find coordinates for places, then show them on a map.

Relationship discovery

Detect and visualise relationships between entities within and across documents; use the Semantic Web to get even more knowledge

Social network analysis

Monitor what social networks are talking about – and in what terms.

Question answering

Obtain concise and relevant answers to questions without the need to browse through all the supporting documents.

Enterprise search

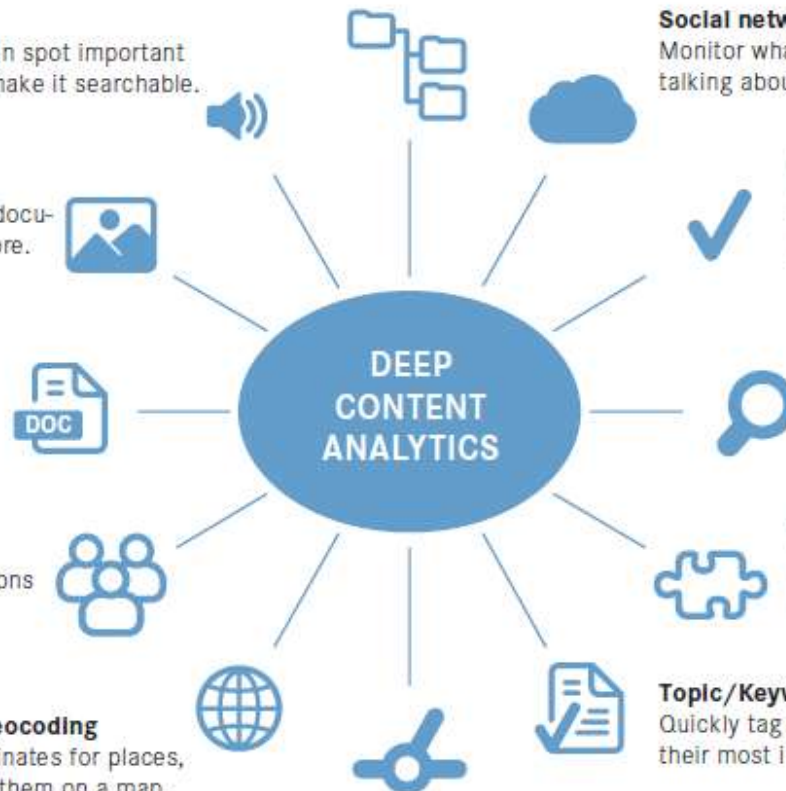
Efficient and effective search is the first step to unlocking the content of unstructured documents.

All-source connectors

Connect to file systems, web sites, social networks, databases and more to collect their unstructured content.

Topic/Keyword extraction

Quickly tag and summarise documents by identifying their most important words, expressions and topics.



ANALYTICS IN THE HEALTH DOMAIN

What can be done

„Low-hanging fruits“ of analytics

- **Deep(er) content analysis**
 - From scanned paper to information extraction, discovery, categorization
- **Geo-spatial intelligence**
 - Combine textual, numerical and geographical information



WHAT CAN BE DONE

Deep(er) content analysis

Automatic categorisation

Automatically categorise or cluster your textual content according to your choice of classes thanks to machine learning.

Spoken language understanding

Recognise the language of audio files, then spot important keywords or fully transcribe it to text to make it searchable.

Social network analysis

Monitor what social networks are talking about – and in what terms.

Image analysis

Recognise text even in poorly scanned documents, identify faces in pictures and more.

Question answering

Obtain concise and relevant answers to questions without the need to browse through all the supporting documents.

Automatic translation

Translate content into your favourite language; search for information in multiple languages.

Enterprise search

Efficient and effective search is the first step to unlocking the content of unstructured documents.

Information extraction: entities, relationships, sentiment, topics

Named entity recognition

Identify people, locations, organisations mentioned in text – no lists needed!

All-source connectors

Connect to file systems, web sites, social networks, databases and more to collect their unstructured content.

GIS and geocoding

Find coordinates for places, then show them on a map.

Topic/Keyword extraction

Quickly tag and summarise documents by identifying their most important words, expressions and topics.

Relationship discovery

Detect and visualise relationships between entities within and across documents; use the Semantic Web to get even more knowledge

DEEP(ER) CONTENT ANALYSIS

Combining OCR, NLP and enterprise search to analyze documentation

KIWI SEARCH

LOCATIONS (41)

- INDIA 4
- AUSTRIA 2
- BEIRUT 2
- CORINTH 2
- CARGO 2
- DELHI 2
- GREAT BRITAIN 2
- GREECE 2
- GUENADINES 2

ORGANIZATIONS (29)

- ISAN BANK CORP 2
- CEAM SHIPPING AGENCY 2
- COMMUNIT COMMUNITY CENTER CENTER 2
- DAMAN INDUSTRIAL ESTATE 2
- FINANCIAL SERVICES CENTER 2
- Q.E. 2
- Q.M.P REFINERIES SA 2
- MOR ATATUBIC OIL Q.M.P. 2

VESSELS (8)

- ARION 2
- MT AZOV MARINER 2
- CMA CGM 2
- MED I 2
- SILVER ROAD 2

PEOPLE (1)


- OO 2

14 DOCUMENTS

Filters:

00000002.PDF	Scanned by jcx on 2015-05-01	<input type="button" value="PDF"/> <input type="button" value="CSV"/>
00000002.PDF	Scanned by saq on 2015-05-01	<input type="button" value="PDF"/> <input type="button" value="CSV"/>
00000008.PDF	Scanned by saq on 2015-05-01	<input type="button" value="PDF"/> <input type="button" value="CSV"/>
00000007.PDF	Scanned by saq on 2015-05-01	<input type="button" value="PDF"/> <input type="button" value="CSV"/>
00000004.PDF	Scanned by saq on 2015-05-01	<input type="button" value="PDF"/> <input type="button" value="CSV"/>
00000004.PDF	Scanned by jcx on 2015-05-01	<input type="button" value="PDF"/> <input type="button" value="CSV"/>
00000005.PDF	Scanned by jcx on 2015-05-01	<input type="button" value="PDF"/> <input type="button" value="CSV"/>
00000007.PDF	Scanned by jcx on 2015-05-01	<input type="button" value="PDF"/> <input type="button" value="CSV"/>
00000008.PDF	Scanned by jcx on 2015-05-01	<input type="button" value="PDF"/> <input type="button" value="CSV"/>

HIGHLIGHTED PDF



TUKITNACHRA STEEL & COMMODITIES TRADING LTD.

The Financial Services Center, Saq Ground Kingroad VC65-098
St Vincent and The Greenland

To: ATG NOUR STEEL LTD.
529 Silver Road, Manglam
Place, behind tall mata, temple
Sector-2 Roohini, Delhi 75
India

COMMERCIAL INVOICE
No: 2014-A25689
Date: 17-July-2014

Shipping Information Form

Quantity	Description	Price	Total
132 tons	Roller steel deformed reinforcing bars As per bony Biliary's sales/purchase contract No: ATG(2013-A25698) Dated 07-June-2014	US 6225 per ton	US 83,104 \$
Invoice amount			US 82,284 \$

Country of Origin : Turkey
 Port of loading : Marmara Sea, 15 Turkey
 Place of Delivery : Tuzluca Port, Marmara Sea
 Terms of delivery : CF Tuzluca Port, Marmara Sea, Incoterms 2010
 Drawn under the documentary credit no. 3695875221C2015 of Alfa Bank Ltd, 5 Kingroad,
 DeHi 110 085, India

Authorized Signatures

Bank account : Lakshmi Bank SA, Toronto
IBAN : CA15 2315 1905 0558

WHAT CAN BE DONE

Geospatial intelligence

Automatic categorisation

Automatically categorise or cluster your textual content according to your choice of classes thanks to machine learning.

Spoken language understanding

Recognise the language of audio files, then spot important keywords or fully transcribe it to text to make it searchable.

Image analysis

Recognise text even in poorly scanned documents, identify faces in pictures and more.

Automatic translation

Translate content into your favourite language; search for information in multiple languages.

Social network analysis

Monitor what social networks are talking about – and in what terms.

Question answering

Obtain concise and relevant answers to questions without the need to browse through all the supporting documents.

Enterprise search

Efficient and effective search is the first step to unlocking the content of unstructured documents.

All-source connectors

Connect to file systems, web sites, social networks, databases and more to collect their unstructured content.

Topic/Keyword extraction

Quickly tag and summarise documents by identifying their most important words, expressions and topics.

Geospatial intelligence

Named entity recognition

Identify people, locations, organisations mentioned in text – no lists needed!

GIS and geocoding

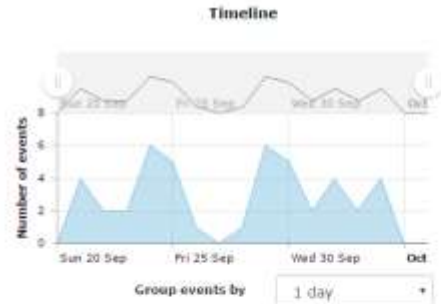
Find coordinates for places, then show them on a map.

Relationship discovery

Detect and visualise relationships between entities within and across documents; use the Semantic Web to get even more knowledge

GEOSPATIAL INTELLIGENCE

Activity-based Intelligence Dashboard



Filter events

Start date: 09/21/2015 07:59

End date: 10/05/2015 02:39

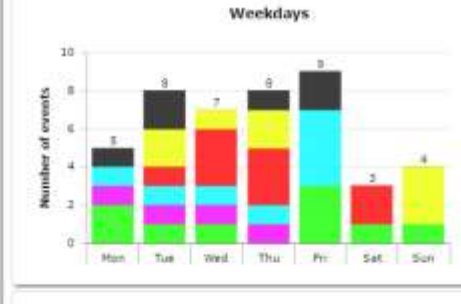
Sector: [Dropdown]

Start week: 39 2015

End week: 41 2015

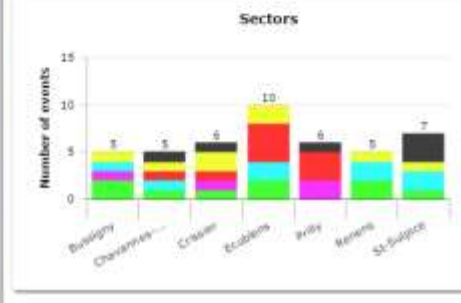
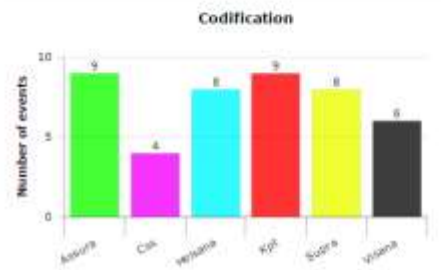
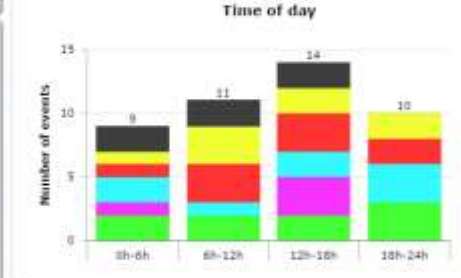
Keep only codified events

Filter events Cancel filters Export data Showing 44 events



Event types

Kidney disease Hepatitis Diabetes
Pancreatitis



ANALYTICS IN THE HEALTH DOMAIN

A non-exhaustive list of needs

- **Analyze medical documentation**
 - Collect & summarize anamneses
 - Find similar cases
 - Identify promising treatments
- **Communicate with patients**
 - Speak their language
 - Manage emotions
- **Prevent fraud**
 - Insurance
 - Legal analysis



ANALYZE PATIENT CASES

COBEDIAS®: an “electronic status sheet”



Motivations:

- Use data from daily clinical examination to improve the treatment of patients

Goal of COBEDIAS®:

- Provide a snapshot of the patient’s health status & history (‘elektronisches Statusblatt’)
- Compare it with anonymized reference data to make a diagnosis

How it works:

- Each practitioner collects patient info using a structured representation, then stores it in a secured, anonymized repository
- Data from other practitioners’ patients can be
 - mined for similar cases via an algorithm to *query cases by profile*, and
 - used to *estimate a risk profile*

www.cobedix.ch

ANALYZE PATIENT CASES (2)

COBEDIAS®: how it works

Krankheitsprofil > Arteriosklerose > Instanz 1
 Arteriosklerose; Testgruppe: wfm, enthält 125; Kontrollgruppe: wfm, enthält nicht 125 - 01.09.2015 - Instanz 1

Testgruppe
 Einträge: 40

Statisches Feld	P-Wert	Median	Feld-Vollst.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 Anzahl Diagnosen aktiv	0.000	4.00	98%	3	1	2	2	2	2	3	3	3	2	2	2	2	2	2	2	2	2	2	2
2 Anzahl Diagnosen total	0.000	4.00	98%	3	1	2	2	2	2	3	3	3	2	2	2	2	2	2	2	2	2	2	2
3 Alter bei der Untersuchung	0.000	73.50	100%	39	76	75	75	94	67	58	59	70	70	70	69	69	68	68	68	68	68	68	68
4 WHR	0.001	1.02	78%	0.95	0.98	0.98	0.8	0.8	0.89	1	0.98	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	
5 Bauchumfang (cm)	0.003	105.00	78%	99	98	92	100	73	88	92	90	105	105	105	105	105	105	105	105	105	105	105	
6 Blutdruck systolisch liegend rechts	0.009	140.00	100%	110	100	105	110	120	140	130	140	130	140	140	140	140	140	140	140	140	140	140	
7 Durchschnittlicher Blutdruck systolisch	0.016	140.00	100%	105	100	102.5	100	120	140	127.5	140	127.5	135	135	135	140	125	140	152.5	150	150	150	
8 Blutdruck systolisch liegend links	0.033	140.00	98%	110	120	110	120	120	130	140	120	130	130	130	130	130	130	130	130	130	130	130	
Krankheitsaktivitäts-Score				0.25	0.57	0.63	0.75	1.13	1.13	1.12	1.13	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.25	

Kontrollgruppe
 Einträge: 78

Statisches Feld	P-Wert	Median	Feld-Vollst.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
1 Anzahl Diagnosen aktiv	0.000	1.00	99%	2	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
2 Anzahl Diagnosen total	0.000	1.00	99%	2	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
3 Alter bei der Untersuchung	0.000	60.50	100%	62	38	10	30	23	53	50	47	35	43	79	77	71	66	48	67	60	74	60
4 WHR	0.001	0.94	83%	0.89	0.91	0.76	0.78	0.82	0.82	0.9	0.9	0.75	0.75	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94
5 Bauchumfang (cm)	0.003	97.00	83%	91	60	74	36	75	67	82	67	76	76	104	104	97	110	110	101	111	111	111
6 Blutdruck systolisch liegend rechts	0.009	130.00	99%	110	30	80	110	90	100	100	110	110	110	110	110	110	110	110	110	110	110	110
7 Durchschnittlicher Blutdruck systolisch	0.016	125.00	100%	110	97.5	65	110	90	97.5	100	92.5	110	102.5	102.5	102.5	102.5	102.5	102.5	102.5	102.5	102.5	102.5
8 Blutdruck systolisch liegend links	0.033	130.00	100%	110	100	80	80	80	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Krankheitsaktivitäts-Score				0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.67	1.75	1.75	1.75	1.75	1.80	2.00	2.10	2.10

[Als PDF öffnen](#)

COMMUNICATE WITH PATIENTS

Sumex CMI: transformation service for comprehensive medical invoices

Motivations:

- General invoice templates are optimized for OCR, not patients
- Only experts can really understand invoices

Goal of Sumex CMI:

- Make invoices readable by laymen
 - Simplify technical terms
 - Categorize/aggregate information
 - Enrich invoices with web hyperlinks

suva

Mehr als eine Versicherung

Versicherter Paul Mustermann Beispielgasse 3 6262 LANGNAU B. REIDEN	Behandelt von 6687abf08665fbd5a00958bfb 6004 LUZERN	Rechnungsdatum 07.02.2014	Rechnungsnummer 741046186
		AHV-Nummer 49080c3104c8	Unfallnummer 49080c3104c8
		Betrag CHF 231.94	

Übersicht

Leistungen	Kosten
Untersuchungsgespräch	CHF 73.51
Medikamente	CHF 5.65
Pauschale	CHF 113.30
Material	CHF 23.15
Leistungen in Abwesenheit des Patienten	CHF 16.33

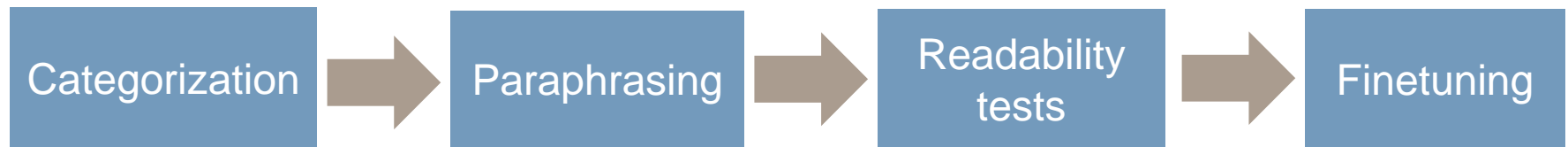
Behandlung vom 15.01.2014 - Sitzung 1

Leistungen	Dauer/Anzahl	Kosten
Untersuchungsgespräch		
Dauer des Untersuchungsgesprächs 00.0010/00.0020/00.0030	20 min	CHF 57.18
Beratung, pro 5 Min. 00.0510	5 min	CHF 16.33
Medikamente		
Einfaches entzündungshemmendes Mittel: MEPHADOLOR Neo Filmtabl 500 mg 10 Stk 3258383	1	CHF 5.65

Leistungen in Abwesenheit des Patienten		
z.B. Aktenstudium, pro 5 Min. ¹ 00.0140	5 min	16.33
Medikamente		
Einfaches entzündungshemmendes Mittel: Mephadolor Neo, Filmtabletten ¹ 3258383 durch Grundversicherung abgedeckt (10% Selbstbehalt)	1	5.65
z.B. Aktenstudium, pro 5 Min. 00.0140	5 min	CHF 16.33

COMMUNICATE WITH PATIENTS (2)

Sumex CMI: how it works



Tarmed (original)

- «Versorgung einer Schulterluxation, geschlossene Reposition»
- « Réduction fermée d'une luxation de l'épaule »
- «Cura di una lussazione della spalla, riduzione incruenta»

Tarmed (paraphrased)

- **Behandlung:**
 - «Versorgung eines ausgelenkten Schultergelenks ohne Operation»
- **Acte/soin:**
 - « Réduction d'une luxation de l'épaule sans opération »
- **Trattamento:**
 - «Trattamento di una lussazione della spalla, senza operazione»

PREVENT FRAUD

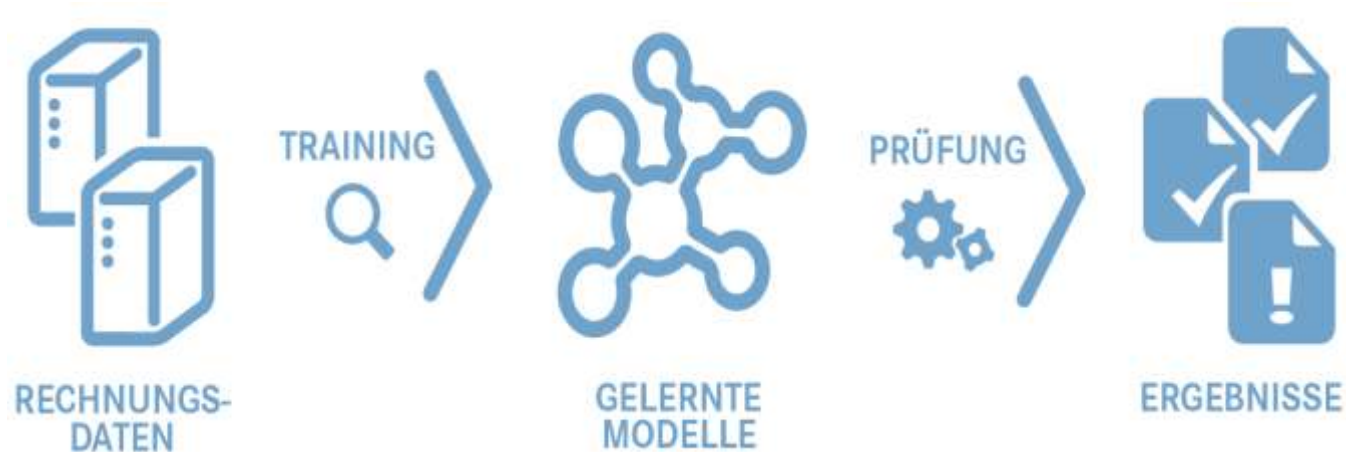
Sumex FDS

Motivations:

- According to the Swiss Insurance Association, 10% of medical refunds are based on fraudulent claims → important losses to insurance providers
- It is difficult to detect fraud based on rules as fraudulent behavior evolves

Goal of Sumex FDS:

- Post-hoc analysis to identify providers of unusual invoicing over a long period

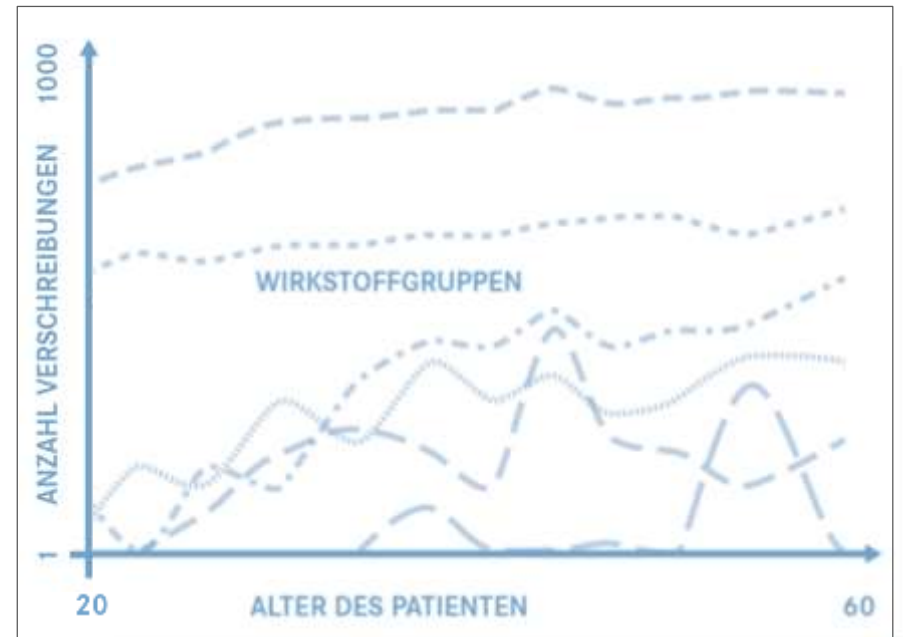


PREVENT FRAUD (2)

Sumex FDS

How it works:

- Machine learning used to learn patterns/find outliers in large invoice volumes
 - Is the MD issuing the invoice an outlier in terms of invoice cost/frequency?
 - Does invoice prescription «agree» with patient age range?



TAKE AWAY

- Analytics allows different actors in the health domain to identify patterns and receive suggestions based on „raw“ data
- Key tools: machine learning and NLP
- Use cases: patient case analysis, communication, fraud detection

- These technologies make us dream of a future where a 1000 AIs power our MD...
- ...but let's stick to the present: deeper NLP, geographical intelligence are „low-hanging fruits“ of analytics

Thank you.

Contact

Silvia Quarteroni

Manager

Silvia.quarteroni@elca.ch

ELCA Informatique SA | Lausanne 021 613 21 11 | Genève 022 307 15 11

ELCA Informatik AG | Zürich 044 456 32 11 | Bern 031 556 63 11

www.elca.ch