



Kairntech for “Smart Innovation”  
Online Workshop, 15.2.2022

Teile der Aufgaben bei “Smart Innovation” Projekten können durch automatische Verfahren aus Natürlicher Sprachverarbeitung (NLP) unterstützt werden.

Die Partner TecIntelli / Uni Stuttgart setzen hierzu Kairntech ein

→ Überblick zu Kairntech und den Analysemöglichkeiten.



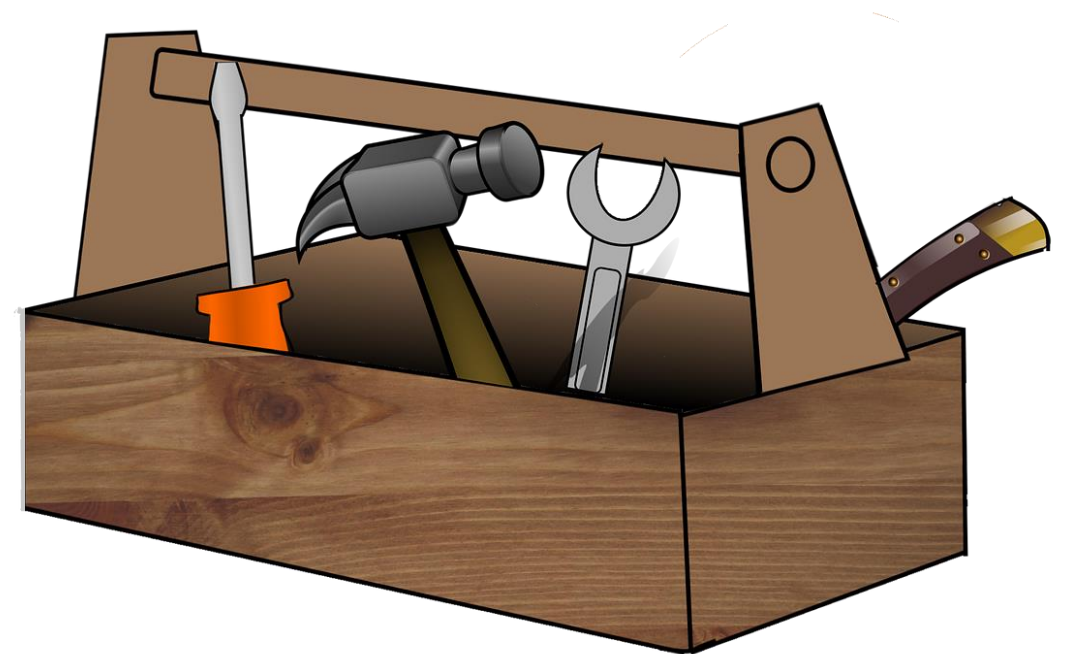
- Französisch/deutscher Spezialist in KI/NLP
- Hintergrund, Erfahrungen : Xerox, IBM, TEMIS, ...
- “No code” Plattform: NLP für Domänenexpert:innen, nicht (nur) für Programmierer:innen, ermöglicht Einsatz von KI/NLP –Verfahren ohne eigene Codeentwicklung.
- Breite Anwendbarkeit. Beispiels-Anwendungsfälle:
  - AFP: Analyse von Presstexten in versch. Sprachen
  - BVA: Analyse von Kundenzuschriften (“Sentiment Analysis”)
  - Boehringer Ingelheim: Analyse von Webinhalten
  - Sopra Steria: Strategische Analysen
  - Fraunhofer SCAI: Autom. Generierung von Wissensgraphen (“Knowledge Graph”)
- Infos unter [www.Kairntech.com](http://www.Kairntech.com)



*“Kairn”, deutsch: “Steinmännchen”*

## Analyse von Dokumenten mit KI/NLP

- Schlüsselbegriffe / Entitäten (“Weltwissen”):
  - Namen, Orte, Daten, Konzepte
- Verwendung kundenspezifischer Vokabulare
- Einsatz maschineller Lernverfahren zum Training eigener, neuer Konzepte
- Dokumentenkategorisierung
- Relationenerkennung: Zusammenhänge zwischen Konzepten/Entitäten
- Zusammenfassungen (“Summaries”)
- Komplexe Verarbeitungsketten: OCR, Speech-Recognition, Crawling → Analyse → Reporting
- REST-API, Cloud-Installation oder lokal



- Über 90 mio Konzepte, praktisch alle Themenbereiche
- Multilingual
- Ständig aktualisiert
- Terme erhalten Konfidenzwert, Typ (Überbegriffe), Verlinkung zu Hintergrundinformationen
- Viele fachspezifische Thesauri: MeSH, NCBI, Getty Geographic Names, ZBW-STW, Gemeinsame Normdatei der DNB und viele mehr.
- Desambiguierung: Passende Lesart bei Mehrdeutigkeiten
- Anpassbar auf eigene Anforderungen (“nur chem. Substanzen”, “nur niederländische Maler”, “nur bildgebende Verfahren in der Medizin”, ...)
- Online:
  - <https://sherpa-entityfishing.kairntech.com/>
  - <https://sherpa-sandbox.kairntech.com/sherpa/> (requires login) Project PressCategorization

WAYNE DOUGLAS GRETZKY CC is a CANADIAN former PROFESSIONAL ICE HOCKEY PLAYER and former HEAD COACH. He played 20 seasons in the NATIONAL HOCKEY LEAGUE (NHL) for four teams FROM 1979 TO 1999. Nicknamed the Great One, he has been called the greatest hockey player ever by many SPORTSWRITERS, players, the NHL itself, and by THE HOCKEY NEWS, based on extensive surveys of HOCKEY WRITERS, ex-players, GENERAL MANAGERS and coaches. GRETZKY is the LEADING GOAL SCORER, assist producer and POINT SCORER in NHL history, and has more ASSISTS in his career than any other player scored total points. He is the only NHL player to total over 200 points in one season, a feat he accomplished four times. In addition, GRETZKY tallied over 100 points in 16 professional seasons, 14 of them consecutive. At the time of his retirement in 1999, he held 61 NHL records: 40 REGULAR SEASON records, 15 playoff records, and 6 All-STAR RECORDS.

## NHL

Normalized: National Hockey League

Domains: Sociology

conf: 0.7078

The National Hockey League (NHL; ) is a professional ice hockey league in North America comprising 32 teams—25 in the United States and 7 in Canada. It is considered to be the premier professional ice hockey league in the world, and is one of the major professional sports leagues in the United States and Canada. The Stanley Cup, the oldest professional

NON-HODGKIN LYMPHOMA (NHL) is a group of BLOOD CANCERS that includes all types of LYMPHOMAS except HODGKIN LYMPHOMAS. [1] SYMPTOMS include ENLARGED LYMPH NODES, FEVER, NIGHT SWEATS, WEIGHT LOSS and TIREDNESS. Other SYMPTOMS may include BONE PAIN, CHEST PAIN or ITCHINESS. Some forms are SLOW-GROWING, while others are fast-growing. LYMPHOMAS are types of cancer that develop from LYMPHOCYTES, a type of WHITE BLOOD CELL. RISK FACTORS include POOR IMMUNE FUNCTION, AUTOIMMUNE DISEASES, HELICOBACTER PYLORI INFECTION, HEPATITIS C, OBESITY, and EPSTEIN-BARR VIRUS INFECTION.

## NHL

Normalized: Non-Hodgkin lymphoma

Domains: Medicine

conf: 0.8508



Non-Hodgkin lymphoma (NHL) is a group of blood cancers that includes all types of lymphoma except Hodgkin lymphoma. Symptoms include enlarged lymph nodes, fever, night sweats, weight loss and tiredness.

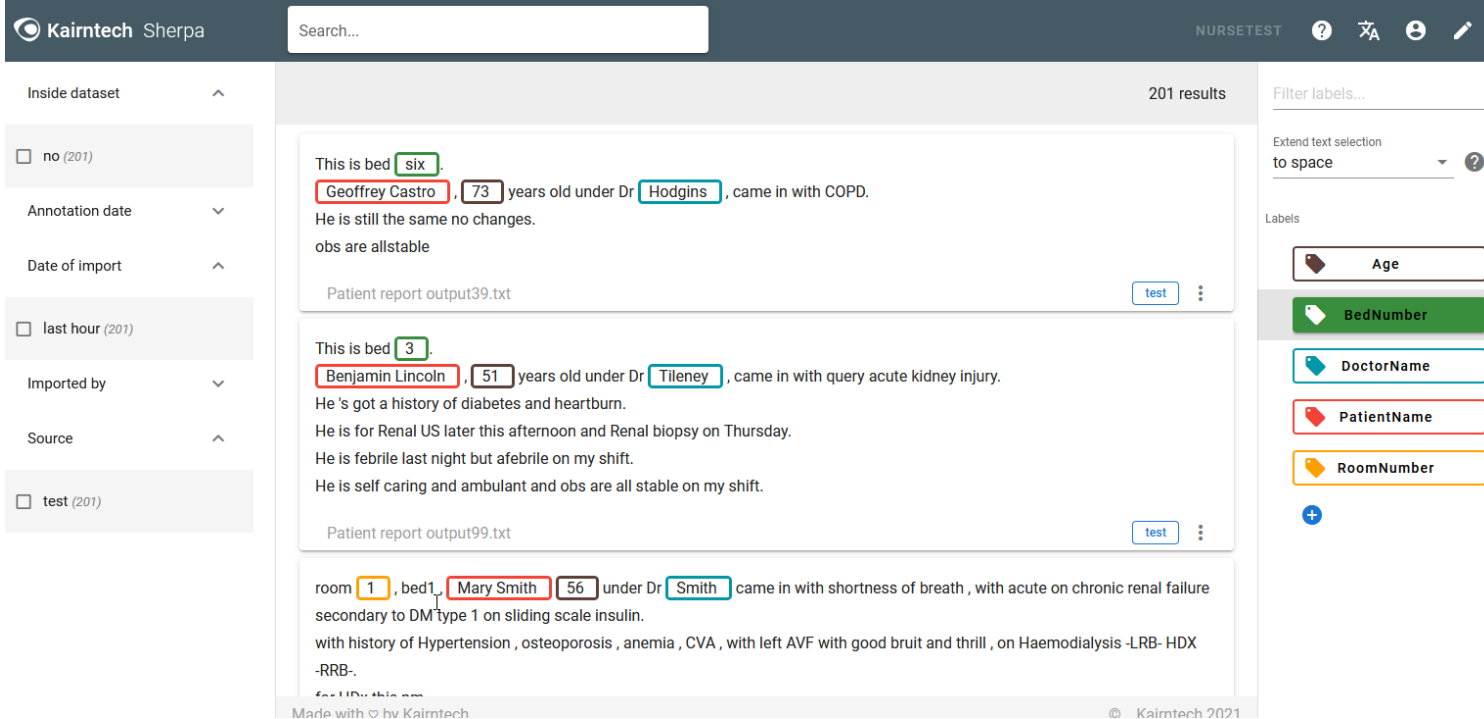
# Annotation mit eigenen Vokabularen

- Eigene Wortlisten, Thesauri importieren und in neuen Erkennen übersetzen: Fallbeispiel Batterie-Technologien
- Dokumente damit annotieren
- Training eines neuen Machine Learning Models und Anwendung auf Dokumente: Welche neuen Ausdrücke findet das Modell?
- Thesauruspflege und –erweiterung automatisch unterstützen lassen: Modell schlägt mögliche neue Terme vor.
- Online:
  - <https://cube.kairntech.com:9071/sherpa/> (requires login) Project BatterySystems

The screenshot shows the Kairntech Sherpa interface with a search bar at the top containing 'BATTERYSYSTEMS'. On the left, there is a 'Lexicon' sidebar with checkboxes for 'SulfurBased (14)', 'RedoxFlow (11)', 'AirBased (8)', and 'SolidState (4)'. Below this is an 'Imported by' section with 'stefan (37)' and an 'Import date' dropdown. The main area displays four panels for different lexicons: 'sodium-sulfur', 'li-sulfur', 'magnesium-sulfur', and 'Na-S'. Each panel shows a JSON object with fields like 'createdAt', 'createdBy', 'identifier', 'lexicon', 'preferredForm', and 'properties'. At the bottom, there are navigation arrows and page numbers 1 through 5.

The screenshot shows the Kairntech Sherpa interface with a search bar at the top containing 'Unknown Redoxflow'. The main area displays search results with 1036 results. The first result is a text snippet: 'The excellent performance of the AFB was ascribed to a zero gap design, deploying a low-resistant, inexpensive anion exchange membrane (AEM), and implementing a copper ion selective ionomer coating on the copper mesh electrodes.' The second result is: 'In this work, we present an experimental characterization of the system Br2-HBr-H2O at chemical equilibrium, featuring isothermal measurements of Open Circuit Voltage (OCV), density and conductivity in solutions with concentrations up to 2 M Br2 and 4 M HBr.' The third result is: 'Among different types of batteries, Vanadium Redox Battery is the most suitable battery.' The fourth result is: 'Here, we show that graphene foams (GFs) are ideally suited as 3D lightweight current collectors for LSBs and outperform the conventional carbon-coated Al (Al/C) foils at the electrode level.' On the left, there is a sidebar with 'Inside dataset' and 'Label' sections. On the right, there is a 'Filter labels...' section with a list of labels: 'AirBased', 'RedoxFlow', 'SolidState', 'SulfurBased', 'Unknown Airbased', 'Unknown Redoxflow', 'Unknown Solidstate', and 'Unknown Sulfurbased'.

- Ihr Projekt erfordert völlig neue Entitäten? Training eines neuen Erkenners.
- Kairntech unterstützt und beschleunigt Erstellung von Trainingskorpora (“No code”)
- Leistungsfähige Lernalgorithmen per Mausclick (schnell und einfach: CRF. Oder mächtig und state of the art: Deep Learning, Transformers)
- “Active Learning” beschleunigt Annotationsprozess
- Online:
  - <https://sherpa-sandbox.kairntech.com/sherpa/> (requires login) Project NurseReportTest

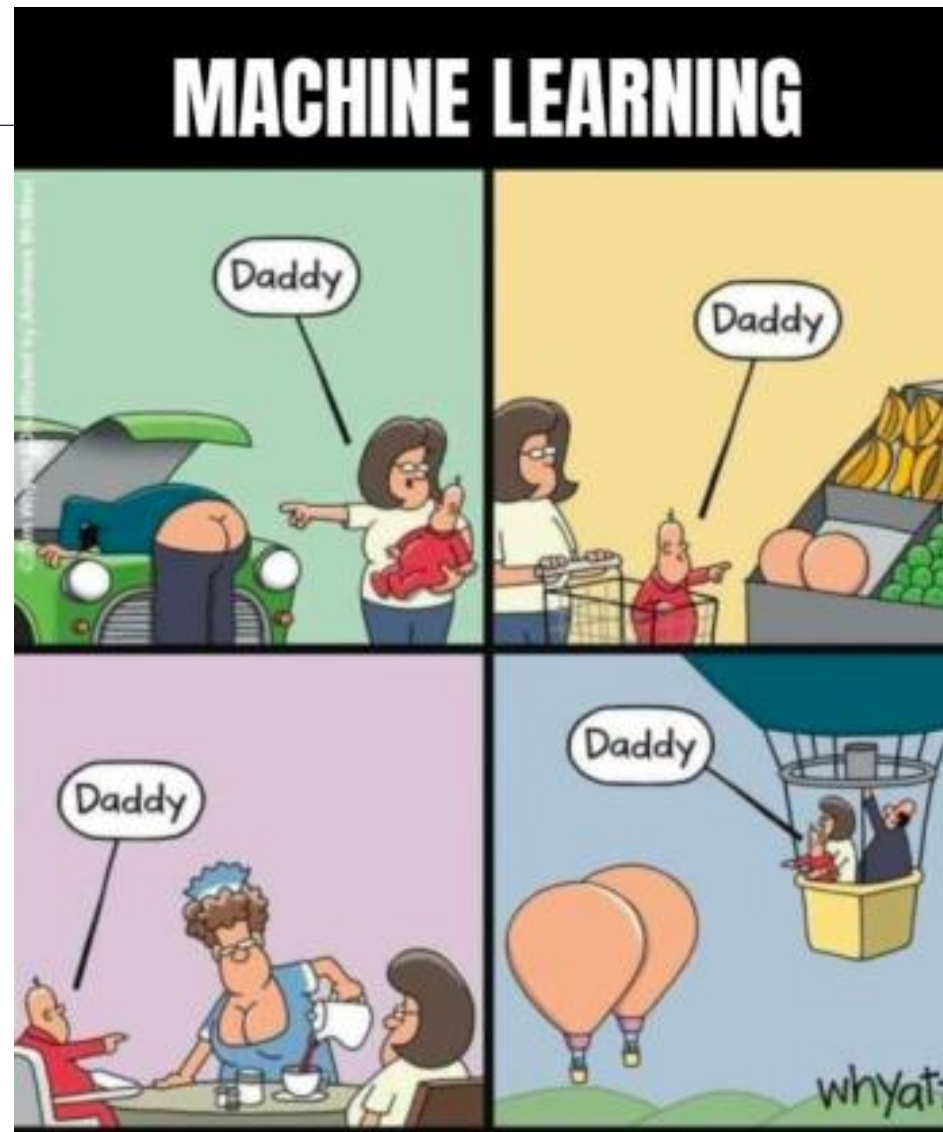


The screenshot displays the Kairntech Sherpa interface for training a model. The main window shows a patient report with several entities highlighted and labeled. The entities are: 'six' (bed number), 'Geoffrey Castro' (patient name), '73' (age), 'Hodgins' (doctor name), '3' (bed number), 'Benjamin Lincoln' (patient name), '51' (age), 'Tilney' (doctor name), '1' (room number), 'Mary Smith' (patient name), '56' (age), and 'Smith' (doctor name). The interface includes a search bar, a filter sidebar on the left, and a label selection panel on the right. The label selection panel shows 'BedNumber' as the selected label, with other labels like 'Age', 'DoctorName', 'PatientName', and 'RoomNumber' also visible. The patient report text is: "This is bed six. Geoffrey Castro, 73 years old under Dr Hodgins, came in with COPD. He is still the same no changes. obs are allstable. Patient report output39.txt". Below this, another report snippet is shown: "This is bed 3. Benjamin Lincoln, 51 years old under Dr Tilney, came in with query acute kidney injury. He's got a history of diabetes and heartburn. He is for Renal US later this afternoon and Renal biopsy on Thursday. He is febrile last night but afebrile on my shift. He is self caring and ambulant and obs are all stable on my shift. Patient report output99.txt".

- Bedarf an Unterstützung bei “Smart Innovation” Prozessen?  
→ TechIntelli, Fraunhofer, Uni Stuttgart
- Bedarf oder Interesse an NLP allgemein?  
→ Kairntech
  - [www.kairntech.com](http://www.kairntech.com)
  - Für unverbindlichen, kostenfreien Testzugang zur gezeigten Software: [info@kairntech.com](mailto:info@kairntech.com)
  - Kairntech Youtube Channel: <https://www.youtube.com/channel/UCgoylH8E6-8nopmY79FdS5A>







Vielen Dank für die Aufmerksamkeit  
[Stefan.geissler@kairntech.com](mailto:Stefan.geissler@kairntech.com)