

 **KIT**
200
JAHRE

look **KIT**

DAS MAGAZIN FÜR FORSCHUNG, LEHRE, INNOVATION
THE MAGAZINE FOR RESEARCH, TEACHING, INNOVATION
AUSGABE/ISSUE #04/2025
ISSN 1869 2311

KÜNSTLICHE INTELLIGENZ

VERLÄSSLICH: METHODEN FÜR VERTRAUENSWÜRDIGE KI
RELIABLE: METHODS FOR TRUSTED AI

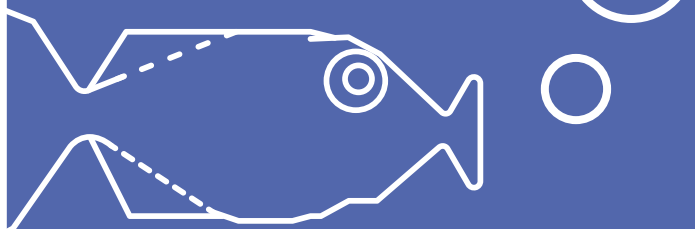
VISIONÄR: ROBOTER, DIE MENSCHEN VERSTEHEN
VISIONARY: ROBOTS THAT UNDERSTAND HUMANS

VERÄNDERT: MIT KI-TUTOR DURCHS STUDIUM
CHANGED: STUDIES WITH AN AI TUTOR

FROM
BABEL FISH
TO
**DEEP
LEARNING**

HOW MACHINES LEARN
TO UNDERSTAND US

BY MARTIN GROLMS // TRANSLATION:
FACHÜBERSETZUNGEN HUNGER/ALTMANN GBR





Artificial Intelligence is already able to translate rather well what humans are saying. Nevertheless, there remains much to do until machines can truly convey meaning, Professor Alexander Waibel and Professor Jan Niehues agree. The two KIT researchers tell the tale of how the idea of a talking machine transformed into AI real-time translations.

In Douglas Adams' novel "The Hitchhiker's Guide to the Galaxy," a little fish solved one of the oldest problems of humankind: Whoever puts the Babel fish into their ear is able to understand any language. This nice idea was considered utopian for a long time, until researchers such as Alexander Waibel started to make it happen.

Waibel heads the Interactive Systems Lab (ISL) at KIT and at Carnegie Mellon University in the US. As early as in the 1970s, he explored the idea of a machine that translates spoken language – at a time when computers had so little computing power that many considered his vision absurd. "I thought of building a machine that you could talk to, and the translation would come out at the other end," he recalls.

At that time, translations were limited to fixed, strict structures: Subject, verb, and object. Waibel thought this was the wrong way to go. He wanted to make machines learn language, instead of programming them. "Children do not learn grammar, they learn by listening and imitating," he says. This insight led him to devising statistical and early neural models that would improve by experience and that, for the first time, recognized patterns in language. But the computers of that era were not powerful enough. "We had the right ideas and algorithms, but it took 20 years until the computing power was sufficient to leverage the full potential of learning machines."

The breakthrough came with the Internet. Billions of texts became available, forming the basis for training language models. "You could almost say that humankind has written down its entire knowledge in the web," says Jan Niehues from KIT's Institute for Anthropomatics and Robotics (IAR), who earned his doctoral degree under Waibel. He is head of the Artificial Intelligence for Language Technologies lab, which ties in with the groundbreaking work of his colleague.

In earlier times, there were separate systems for each task: One for translation, one for

Professor Alexander Waibel from KIT's Institute for Anthropomatics and Robotics (IAR)

Professor Alexander Waibel vom Institut für Anthropomatik und Robotik (IAR) des KIT



FOTO: MARKUS BREIG

Professor Alexander Waibel: Language AI Pioneer

Alexander Waibel is one of the enablers of modern Artificial Intelligence. Since the 1980s, he has been investigating the interface between language, humans, and machines. As a professor at KIT and at Carnegie Mellon University in Pittsburgh, he leads international teams that work on how to make computers speak, hear, see, and translate.

In the 1980s, Waibel developed the so-called time-delay neural network that counts among the core algorithms of modern AI today. With systems such as JANUS, he laid the foundation for automated simultaneous translation and later at KIT, he realized the Lecture Translator, which transmits lectures in real time.

With more than ten companies founded, he was a pioneer of technology applications as well: His company Mobile Technologies marketed Jibbiggo, the first mobile translation app. It was taken over in 2013 by Facebook. KITES GmbH, of which he was a co-founder, was bought by Zoom in 2021. Today, it is the foundation of automated subtitled and translation features on this platform. Other companies he established are active in the healthcare or charity sectors. Waibel was honored numerous times: For example, he received the IEEE James L. Flanagan Award, the Antonio Zampolli Prize, and the ACM ICMI Sustained Accomplishment Award. He is a member of the German National Academy of Sciences Leopoldina. ■

*Professor Jan Niehues
from KIT's Institute
for Anthropomatics
and Robotics (IAR)*

*Professor Jan Niehues
vom Institut für
Anthropomatik und
Robotik (IAR) des KIT*



FOTO: KIRA HEID



MACHINE LEARNING CAN HELP US PRESERVE ENDANGERED LANGUAGES.

Professor Jan Niehues

speech recognition, and one for text analysis. "Neural models are versatile. They can process text, sound, and even gestures and learn different tasks simultaneously," explains Waibel. The basis of all this is so-called Deep Learning, where neural networks independently acquire patterns even from different sources and relate them to each other.

From the Lab to Deployment

At KIT, this research has already become part of everyday life. The most well-known project is Lecture Translator. Since 2012, it has been used to translate lectures into various languages in real time. Students can read or listen to the contents on their smartphones – an invaluable help for international guests or people with special needs. "Our idea is not to replace teachers, but to rather reduce barriers," says Niehues.

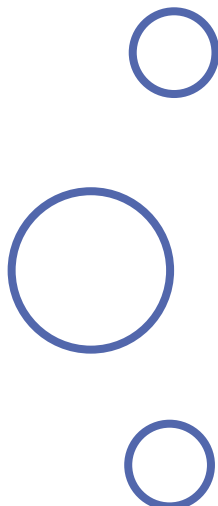
In the meantime, Waibel's team has started to advance these technologies. For example, the new Face Dubber not only translates spoken language, but also synchronizes lip movements and intonation. It thus appears as if the person actually spoke in a different language. But do we need these systems at all when teaching is already done in English at many

universities? Niehues argues that while English is convenient, not everybody perceives it as easy. "Translation can help make contents accessible while at the same time making it easier to get started with a new language."

Indeed, studies have shown that machine translation does not necessarily lead to people learning less, but that it opens up new horizons for them that were previously inaccessible. Waibel sees great opportunities here: "If somebody on a different continent can follow a lecture and communicate freely with other persons, we can speak of practiced inclusion. The technology enables human contacts that ultimately lead to human learning."

More Than Just Words: Language Reflects Culture and Identity

While automated translation systems work exceptionally with big languages such as English and German, it becomes difficult with less common languages. "The fewer data we have, the harder it is to train the machine," explains Niehues. In his research, he therefore also deals with the so-called low-resource languages, i.e., languages with a very small volume of digital data.



*Developed and used
at KIT: The Lecture
Translator translates
lectures live into
different languages*

*Am KIT entwickelt und
genutzt: Der Lecture
Translator übersetzt
Vorlesungen live in
verschiedene Sprachen*

FOTO: MARKUS BREIG

From Lecture Translator to Live Dubbing

Since 2012, the Lecture Translator has been translating lectures in real time at KIT. The system automatically detects spoken language, converts it to text, and then transmits it live as subtitles in multiple languages to the laptops or smartphones of the audience. The system was developed by the research teams of Professor Alexander Waibel and Professor Jan Niehues at the Institute for Anthropomatics and Robotics (IAR). Its purpose is to reduce language barriers during the studies and facilitate access to students with special needs.

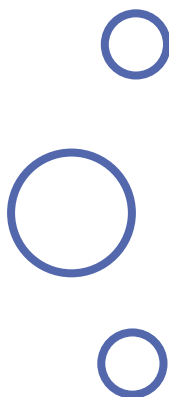
The latest development is called Live Dubbing: It not only translates but also speaks synchronously. An AI system adapts the voice, intonation, and even the lip movements, making it appear as if the person spoke in a different language. Currently, this is still technologically complex, but the objective is clear: Teaching that is understandable for everyone, independent of language. ■

“Machine learning can help us preserve endangered languages, but to do so, we must find ways for models to transfer knowledge from one language to another,” says Niehues. He states that this should work fine because all languages feature common structures. The aim is to preserve diversity. “Language reflects identity,” Niehues stresses. “If it disappears, we lose more than just words.”

The fact that machines learn languages does not at all mean that they understand them. Humor, irony, and gestures remain challenges. “A human understands immediately if something is meant as a joke, but a machine doesn’t,” says Waibel. For him, language is much more than just information. “Language means relationship – this is precisely what makes translating such a demanding task.”

Restricting the Power of AI

With his approach, Waibel also addresses the ethical side of his research. As a member of an international delegation, the professor discussed the ethics of AI with Pope Leo XIV in the Vatican. “We must make sure that it serves us humans, and not the other way round,” he says. Otherwise, wrongly translated contents



Vom Babelfish zum Deep Learning

Wie Maschinen lernen, uns zu verstehen

Künstliche Intelligenz übersetzt heute bereits ziemlich gut, was Menschen sagen. Doch bis Maschinen Sprache wirklich zuverlässig vermitteln, bleibt viel zu tun. Einer der Pioniere der Sprach-KI ist Professor Alexander Waibel. Er leitet das Interactive Systems Lab (ISL) des KIT und arbeitete schon in den 1970er-Jahren an der Idee, gesprochene Sprache maschinell zu übersetzen. „Ich hatte die Idee, eine Maschine zu bauen, in die man hineinsprechen kann und auf der anderen Seite kommt die Übersetzung heraus“, erinnert er sich. Waibel entwickelte statistische und neuronale Modelle, die sich selbst durch Erfahrung verbessern und erstmals Muster in Sprache erkennen konnten. „Wir hatten die richtigen Ideen und Algorithmen, aber es dauerte 20 Jahre, bis die Rechenpower reichte, um die volle Leistung lernender Maschinen zu demonstrieren“, so der Informatiker.

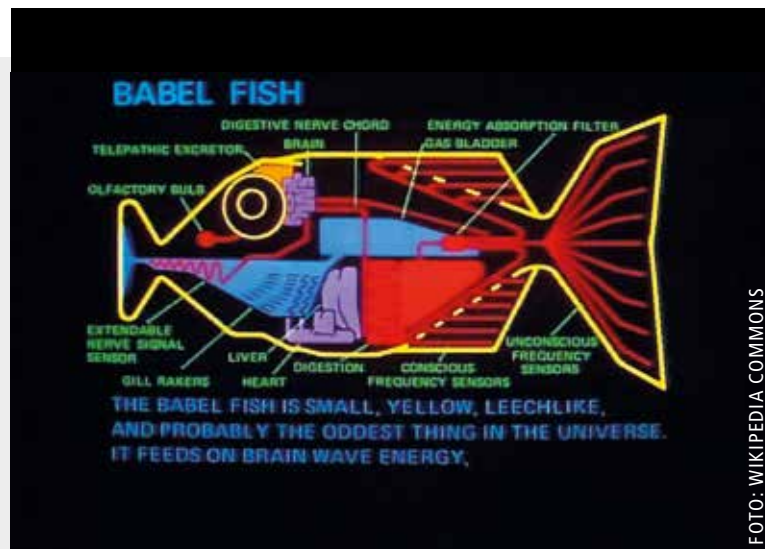
Mit dem Internet kam der Durchbruch. Milliarden von Texten wurden verfügbar und damit zur Trainingsgrundlage für Sprachmodelle. „Neuronale Modelle sind vielseitig. Sie können Text, Ton und sogar Gestik verarbeiten und verschiedene Aufgaben gleichzeitig lernen“, beschreibt Waibel. Seit 2012 übersetzt der am IAR entwickelte Lecture Translator Vorlesungen live in verschiedene Sprachen. Studierende können die Inhalte über ihre Smartphones mitlesen oder anhören – eine große Hilfe für internationale Studierende oder Menschen mit Beeinträchtigung. „Wir wollen Barrieren reduzieren“, sagt Waibel. Das Team arbeitet derzeit an einem Face Dubber, der nicht nur Sprache, sondern auch Lippenbewegung und Tonfall synchronisiert.

Mit großen Sprachfamilien wie Englisch oder Deutsch arbeiten Übersetzungssysteme problemlos, doch bei kleineren Sprachen wird es schwierig. „Je weniger Daten es gibt, desto schwerer lässt sich die Maschine trainieren“, erklärt Professor Jan Niehues vom Institut für Anthropomatik und Robotik (IAR) des KIT. In seiner Forschung geht es daher auch um Sprachen mit sehr kleinem digitalem Umfang. „Wir suchen nach Wegen, wie Modelle Wissen von einer Sprache auf eine andere übertragen können“, so der Informatiker. Das funktioniert, weil alle Sprachen gemeinsame Strukturen hätten. „Maschinelles Lernen kann helfen, bedrohte Sprachen zu bewahren.“

Auch die ethische Seite ihrer Forschung, beschäftigt die beiden Wissenschaftler. Als Mitglied einer internationalen Delegation hat Waibel im Vatikan mit Papst Leo XIV über die Verantwortung von KI diskutiert. „Wir müssen dafür sorgen, dass KI dem Menschen dient, nicht umgekehrt“, sagt Waibel. Falsch übersetzte oder manipulierte Inhalte könnten sonst leicht zur Waffe werden. Niehues sieht die Lösung in Aufklärung und Transparenz: „Technologie ist nie neutral. Deshalb müssen wir verstehen, wie sie funktioniert und wo ihre Grenzen liegen.“ ■

The Babel Fish from the novel "The Hitchhiker's Guide to the Galaxy" is no longer pure fiction, but closer to reality than ever before

Der Babelfish aus dem Roman „Per Anhalter durch die Galaxis“ ist heute keine reine Fiktion mehr, sondern der Realität näher denn je



or manipulated voice recordings might easily become weapons.

The panel discussion addressed central risks inherent in AI: Disinformation and deep fakes, loss of control when using autonomous systems, concentration of power, and missing responsibility. The panel calls for clear human responsibility, transparency, protection from fakes, and prohibition of autonomous AI decisions on life and death. Waibel wants AI to be secure, opening up communication without incapacitating humans.

For Niehues, the solution is education and transparency. "Since technology is never neutral, we must understand how it works and where its limits are."

Nevertheless, the outlook remains positive: Language technology may simplify communication, facilitate education, and make knowledge accessible. Fifty years after Waibel's first vision, Babel Fish has not arrived yet, but it is closer than ever before. ■