

European Media  
Laboratory GmbH



# USE YOUR VOICE!

Automatic Transcription

Call Center Analysis

Automatic Translation

Media Transcription

Voice Control

Real Time

Neural Networks

Machine Learning

**BEST PRACTICE EXAMPLES FOR THE USE OF  
EML SPEECH TECHNOLOGY**

# SPEECH ANALYTICS

Business Sector: Call Center, Customer Communications  
Partner: ASC Technologies AG, Hösbach



## Ensuring the key points are addressed

### Financial institute optimizes customer communications with speech analysis, staff training and quality management

Communications in today's call centers involve countless telephone conversations and interaction between customers and agents. Analyzing the interaction with customers has thus become increasingly important. This is why Italian company GERI HDP decided years ago to put its trust in a solution made in Germany: the Workforce Optimization (WFO) Suite produced by ASC Technologies AG. Based in Milan with more than 600 employees, GERI HDP is a finance company specializing in credit and risk management. By employing ASC's quality management and speech analysis solution, the Milanese company derives important information from customer communications that enables it to train its staff more efficiently and improve internal processes. The solution for 140 workplaces was integrated in the existing IT infrastructure at GERI HDP's own Contact Center.

"Introducing our highly-developed WFO solutions has worked wonders at GERI HDP because it delivers relevant feedback to the entire company," explains Marco Müller, COO of ASC. "The speech analytics help to identify the most interesting conversations in order to train employees accordingly and improve internal processes. The whole company benefits and – most important – so do the customers who get better service." ASC bases its speech analytics on EML speech technology using Italian language components. With EML's Language Model Workplace it was possible to tailor the speech analysis to GERI HPD's requirements. "Particularly in contact centers with high call volumes and an almost inconceivable number of interactions, speech analysis is a big advantage," says Marco Müller. "We are delighted to have the speech technology tools of our partner EML at our disposal so that we can offer our customers tailor-made solutions which deliver the necessary decision criteria in real time."

ASC is a global leader in providing software for multi-channel recording, quality management and analytics. The company addresses all enterprises with recording needs, especially contact centers, financial institutions and public safety organizations. ASC offers solutions for recording, analyzing and evaluating interactions across all media, including both fixed-line and cloud solutions. With its world headquarters in Germany and subsidiary companies in the United Kingdom, France, Switzerland, the United States, Brazil, Japan, Singapore, Hong Kong and Dubai as well as a world-spanning service network, ASC is one of the global players in the field.

#### *Marco Müller, COO, ASC Technologies*

"Particularly in contact centers with a high volume of calls, speech analysis is an advantage. We are delighted to be able to offer our customers tailor-made solutions with EML's speech technology."



# VOICE CONTROL

Business Sector: Smart Homes  
Partner: EU Project "Listen", 2015-2019 („Horizon 2020")



## Smart homes will "LISTEN" to your voice

### The EU project "LISTEN" creates a robust, hands-free speech control interface for smart home systems.

Wouldn't it be convenient to control heating or cooling, lighting, and media anywhere in your home by just using your voice, without any device in your hand? For the elderly and people with disabilities, it would be even more than that: It can provide assistive control of important everyday activities and even save lives in case of emergencies.

To develop such a novel smart home functionality, speech technology and signal processing experts from Germany, Greece, and Italy collaborate in the EU project "LISTEN" – "Hands-free Voice-enabled Interface to Web Applications for Smart Home Environments". They are designing a system enabling robust hands-free large-vocabulary voice-based access to Internet applications in smart homes. An important function is the use of a microphone array, which is a matchbox-size unit with 8 microphones. EML European Media Laboratory GmbH provides its speech recognition system and refines its tools for the project. For example, the speech

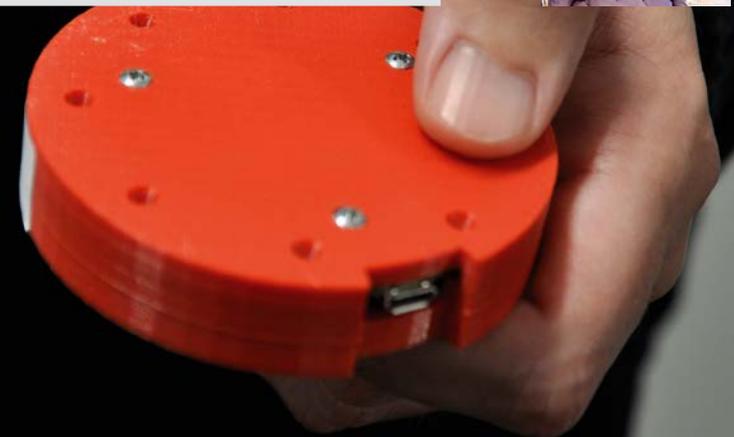
recognition can also – in contrast to the current approach of large providers – be installed locally, which is a prerequisite for enforcing strong privacy rules. In the mid-term project review meeting the project partners demonstrated the system in four languages (English, German, Greek, and Italian). The EU project officer Alina Suhetzki experienced "LISTEN" by herself. Speaking her commands in English, she switched the lights on and off in a blink. "This was an impressive demonstration", she said. "A small and smart system that works in real-time with four languages and can also be adapted to further languages."



"LISTEN" is part of the Marie Skłodowska-Curie Research and Innovation Staff Exchange (RISE) program, in the HORIZON 2020 framework. The other project partners are the **RWTH Aachen University**, Germany, **Cedat 85**, Italy, and the **Greek Foundation for Research and Technology Hellas (FORTH)**, who coordinates the project.

*Alina Suhetzki, EU project officer*

"A small and smart system that works in real-time with four languages and can also be adapted to further languages."



# VOICEMAIL TRANSCRIPTION

Business Sector: Telephony  
Partner: sipgate GmbH, Düsseldorf



## Communications at your service

### Intelligent answering machine makes realtors always reachable

Martin B. is a realtor in Berlin and has just started his own business. He has already organized his internet access, now he just needs a telephone. Of course, he wants to be reachable in the city and have a landline that he can charge to the business. He decides for the sipgate team, the VoIP business package with landline, mobile and fax – all freely scalable with complete cost control. This package offers the new entrepreneur a number of advantages: just a few clicks and he has an account he can use to phone from his notebook or take calls on his cell phone without redirection.

If Martin happens to be talking to a client and can't answer the phone, the intelligent answering machine helps him out: he has programmed the account settings so that this sipgate feature records all his voicemails, automatically transcribes them and emails them to his smartphone. Martin does not have to listen to his messages, he can just read them and reply by email when he has time to do so.

This function gives Martin a level of reachability that is usually exclusive to big real estate agencies.

The intelligent answering machine is based on automatic speech recognition developed by EML European Media Laboratory GmbH. Cooperation between sipgate and the Heidelberg company benefits all customers alike – entrepreneurs like Martin as well as families scattered all over the Republic who want to stay in touch all the time.



*Tim Mois and Thilo Salmon, CEOs, sipgate GmbH*

“We build innovative fixed-network and mobile products for hundreds of thousands of clients, for the home, the office and out and about. Thanks to EML speech technology, we are the only provider in Germany that can offer its clients a special voicemail service.”

# AUTOMATIC TRANSLATION

Business Sector: Finance  
Partner: Lexifone, Haifa, Israel

**EUREKA**   
innovation across borders

## Finding the nuggets in real time

**Lexifone (Haifa/Israel) and EML have developed the first realtime video translation system for the financial market.**

Today's world is a global marketplace, driven by the multimedia internet. Especially in the financial sector, stock market news contains valuable information that financial analysts need to know about as quickly as possible. Financial news frequently includes audio data in various different languages from the most diverse sources. Consequently, there is a growing demand for on-the-fly media transcription and translation, but up to now, there has been no such thing as real-time video translation.

Through pioneering new speech recognition, sentence identification and dynamic vocabulary technologies, the Eureka-funded "Mediatranslator" project has developed the first real-time video translator, linguistically optimised for the financial sector.

This success is the result of a collaboration between Lexifone Communication Systems in Israel and EML. "The technical challenge here was not so much the issue

of translation as much as the issue of speech recognition," explains project coordinator Dr. Ike Sagie, Managing Director of Lexifone. "Current technology does not handle two people speaking at the same time very well. Our breakthrough was achieving the ability to differentiate between two speakers, increasing the accuracy of the transcription."

"Mediatranslator" was part of the EUREKA research initiative, funded by the German Federal Ministry for Economic Affairs and Energy (BMWi) and the Israel Ministry of Economy and Industry. The 18-months' project started in late 2015 and ended in 2017. "We have been integrating transcription and translation technologies using neural networks and machine learning tools," says EML's research manager Dr. Volker Fischer.

Both companies have a distinguished record in the field of speech technology. EML develops speech recognition technologies to convert speech into text; Lexifone offers automatic voice translation of telephone conversations to connect people not speaking the same language. Lexifone supports 15 languages, 24 dialects and calls to over 100 countries.

EUREKA is a European research initiative seeking to promote cross-border cooperation between industry and research institutions conducting technological research and development. The initiative aims to efficiently connect Europe's existing potential of specialists, know-how, research facilities, and financial resources. In this way, EUREKA contributes to the competitiveness of European industry in the world marketplace.

[www.eurekanetwork.org](http://www.eurekanetwork.org)

"We are combining neural network based transcription and machine translation technologies."

**Dr.-Ing. Volker Fischer, Manager Research,  
EML European Media Laboratory GmbH**



# VOICE CONTROL

Business Sector: Automotive

Partner: Mobilcenter Zawatzky, Meckesheim



## Mobile without barriers

### How speech technology makes life easier for drivers with disabilities

When Josef Fleischmann is cruising on the interstate, he, as many young people, relishes the freedom of virtually boundless mobility provided by his motor vehicle. Yet the 25-year-old industrial engineer who works as a project manager for a telecommunications company in Munich drives a little differently from others: He steers his vehicle with a control stick, because he has been suffering from muscular atrophy since birth and uses an electric wheelchair. His car, which vehicle converter Mobilcenter Zawatzky adapted to Fleischmann's needs, ensures his mobility. He accelerates, brakes and steers with a special control stick developed by Zawatzky. But to control the secondary functions such as turn signal, horn, windshield wipers and lights Josef Fleischmann uses voice commands, powered by EML speech recognition technology.

The technology partners EML and Zawatzky developed this system for barrier-free driving a few years ago. After the project was successfully completed under the

Federal Ministry for Economic Affairs and Energy's Central Innovation Program for Mid-Sized Businesses (ZIM), Zawatzky now markets the joint product under the name "CenterVoice" in Germany and some other European countries.

Safety was paramount in the development of the system. Above all, the voice recognition and the response of the various functions had to be fast. With a relatively small vocabulary of 200 to 300 terms and optimum microphone positioning, they attained a recognition rate of almost 100 percent.

In addition, EML speech technologists also embedded what they call "comfort functions", such as controlling the navigation system or dictating messages. "I have been using voice control for almost three years and it works very well for me in my daily life," a satisfied Josef Fleischmann reports. The industrial engineer has an almost 30-mile daily commute to work, but he also drives longer distances, for example, when he goes to visit his parents in Nuremberg.

"Driving with voice control is a lot more convenient because I can fully focus on driving. It is also safer, because I only have to operate one button - unlike a remote control where I would have to take my eyes off the road whenever I want to activate a function, like the rear fog lights, which I do not need very often." Thanks to modern technology, Josef Fleischmann is mobile without barriers – in his professional and his private life.

#### Mobilcenter Zawatzky GmbH

Mechatronic systems enable drivers with disabilities to control a vehicle independently. Depending on the degree of their disability, drivers used to be dependent on expensive special equipment to activate the turn signal or switch on the AC. Voice control is the ideal solution to operate secondary and comfort functions. EML and Mobilcenter Zawatzky GmbH, Meckesheim/Baden have developed a system that is now available in several languages.

#### Andreas Zawatzky, Managing Director, Mobilcenter Zawatzky GmbH

"For over 50 years, we have been building cars for people with disabilities. We were looking for a voice control system for the safe operation of turn signals, horns, windshield wipers, the car radio or window regulators. Thanks to the collaboration with EML we have been able to achieve our goal."



# MEDIA TRANSCRIPTION

Business Sector: Administration, Government  
Partner: Cedat85 srl, Rom, Italy



## A Clear Policy

### Using speech technology to increase government transparency

Technology has become an integral part of our daily lives. Today, citizens expect from their local and national governments that they offer innovative ways to meet the growing demand for access and transparency. But government transparency cannot be achieved without the citizens' participation.

In Italy, the Italian Chamber of Deputies has partnered with the private sector to implement "DIGITAL4DEMOCRACY", a powerful set of tools. It incorporates automatic speech recognition technologies into government meetings. Citizens have free and immediate access to indexed and searchable audio and video content. They can search easily by speaker, keyword, or phrase.

"Prior to DIGITAL4DEMOCRACY, transcription was a time-consuming and resource-heavy process", explains Enrico Giannotti, Vice President of the Italian software company Cedat85 who has implemented the system. "Audio from parliamentary sessions would be recorded and dictated again by someone at a computer

programmed to recognize just his or her voice. The computer would then transcribe the text." The new system avoids the need for repeated dictation, directly transcribing the voice of each speaker and halving the transcription time. It also automatically identifies trends and hot topics, such as when votes are being cast, or legislation is being debated. "This technology improves government efficiency and transparency, as well as improving citizens' trust in government", says Enrico Giannotti.

For ten years, Cedat85 has been cooperating very closely with EML European Media Laboratory and uses EML's speech recognition technology in their solutions. The Italian company sent their speech specialists to Germany: In an EML workshop they learned to develop a language model for Italian on the basis of the knowhow acquired at EML. Back to Italy, they integrated it into Cedat's existing transcription environment.

*Enrico Giannotti, General Manager, Cedat 85 srl, Rome/Italy*

"We are a leading supplier of speech and language services based on innovative technologies. We create systems for the automatic transcription of television and radio programs, parliamentary debates, court proceedings, call center, mobile applications. Thanks to our partnership with EML we have been able to develop highly accurate automatic speech recognition components for Italian and other languages and integrate them into our transcription environment and innovative solutions."





## CALL-CENTER

Business Sector: Customer Dialog, Call Center  
Partner: voiXen GmbH, Berlin

### Interview with Dennis Schottler, CEO at voiXen GmbH, Berlin

***Your company offers a web application to archive, analyze and search voice files. What sectors do you serve and what motivates your clients to use voiXen solutions?***

Most of our clients are call center operators who record thousands of telephone calls every day for a variety of reasons. Their motivation clearly is quality assurance. VoiXen eliminates the need for time-consuming randomized manual quality checks of individual voice files by detecting conversations that are in need of improvement while they are happening, which enables us to provide flexible and effective coaching. But our solutions also support clients in the media/print media industries. Just think of the many interviews that are being conducted for magazines or radio shows. VoiXen clients can retrieve any recording that contains a specific key word or name.

***As a twenty-year veteran of the call center business, you offer tailor-made solutions for this specific industry. In your opinion, what is the added value of speech technology and automated transcription in a call center?***

In a word, I would reply: "Automation!" Customer dialog can and will no longer be conducted exclusively by low-wage employees over the phone. I believe in asynchronized communication. My detractors claim that I'm advocating a revival of the answering machine, and there is something in that. But the thing is that with self-service modules and the permanent availability of the World Wide Web encouraging people to self-help, customers seeking support are unlikely to phone some place if they want to make a service request. People have grown far too frustrated with the incompetence of call center agents (and by that, I mean not only in terms of knowledge, but also their lack of decision-making competence). In the future, you will leave a voice message on websites or even on devices that are connected to the internet. Then, transcription and real business intelligence will kick in and route the support request to an associate who is actually in a position to help and solve the problem. The spoken word, converted into text, supplemented with information on the relevant hardware (age, warranty information, place of

purchase, nearest service technician etc.) will fundamentally transform client communication over the next five years. Ask a 14-year-old how they submit service requests, and you will understand the future of the service industry. In terms of Big Data, we won't go anywhere if we don't embrace Speech2Text. Those who fail to do so will quietly disappear from the market.

***EML is voiXen's technical partner. Which EML products and services do you use?***

We use the fully automated and speaker-independent "EML Transcription Server" that immediately converts incoming calls into searchable text. We also use the "EML Language Model Workplace" which allows us to adapt voice recognition to our clients' different application domains.

***What has been your experience with this so far? How do EML solutions differ from those of other providers?***

EML develops its products and solutions based on research. That makes EML a real partner, because unlike other providers, EML does not just want to harness voiXen's market access to channel its own products. And its research pays off: Thanks to their use of "Deep Learning" methods and, not least, their profound experience with automated speech recognition EML technologists have attained a recognition rate which no other competitor on the market can beat at this point.

***You believe in the future of "asynchronized communication" at the call center, which means: leave a voicemail rather than being out on hold. What benefits does speech technology offer in this context?***

The benefits are obvious: The callers can state their issue right away, without having to endure the nerve-racking ordeal of wrestling with an interactive voice system or languishing on hold. You just start talking! Speech technology can search voicemails for key words, order numbers, names etc. and assign them to the right project or associate who will then take care of your issue. Call centers spend far too much on availability. That hurts their First Contact Resolution rates. A voice mail system with speech recognition can do more for less money.

***Which areas do you think will benefit from the use of automated transcription of the spoken word in the future?***

Anyone who publishes audio contents (spoken, not sung) can make contents retrievable with Speech2Text and determine exactly what kind of advertising insert makes sense at what point. We leave the rest up to Google, Facebook, and the others. Try searching YouTube for all the interviews in which Chancellor Merkel mentions the word "privacy" and jump directly to that spot in the video. We're not there yet, but we will be there soon!

# TELEPHONY

Business Sector: Telephone systems/Customer communications  
Partner: ADDIX Software GmbH, Kiel

## Greater efficiency in the company with transcribed voice messages

### How a telephone systems function transformed a company

It is 5:00 p.m. on Friday and employees at a medical supplies dispatcher in Berlin are finishing work. The answering machine is switched on and everyone heads home. The following Monday morning, a familiar scene unfolds: Someone sits down at the answering machine with paper and pen and listens to the messages, writes down the customers' orders and then passes on the notes to the colleagues in dispatch. Depending on how busy they are, it can take half a day before the note arrives on the right desk. The contents are then recorded and the necessary orders are fed into the system for the relevant customer. A protracted and flawed procedure that ties up human resources.

For some years, the medical supplies dispatcher has been a customer of ADDIX in Kiel. During an in-house workshop on call flow, it emerged, almost by coincidence, how the company had been using the voice mail box integrated in its Astimax telephone system. In order to expedite the process of listening to voice mail boxes, the Kiel company suggested automatically transcribing incoming calls on the answering machine.

Why? Because reading is quicker than listening. It takes about six seconds to read a voice message lasting a minute – an obvious advantage for the company. Since introducing the transcribed mail box, the orders left on it are automatically transcribed and emailed at the click of a mouse to the relevant member of staff. He or she merely has to copy and paste the details required for the order process from the email instead of laboriously typing them in. The customer's telephone order, including the voice file, can be archived in the system.

It is thanks to cooperation between ADDIX and EML European Media Laboratory GmbH that the transcription of mailboxes in Astimax telephone systems functions so successfully. The Heidelberg company encrypts the voice data and processes it exclusively on servers based in Germany – thus observing Germany's strict data privacy laws. The high recognition rates mean trouble-free transcription. The Astimax Team expanded the system to include special technical terminology frequently used by the company. The vocabulary can thus be adapted to any changes such as new product names.

Mailbox transcription offers users genuine added value with measurable saving potential. Examples like this show just what transcription technology can do for digitization. Transcription is also an optional component of the Astimax UC solution which is already a market product.

### ADDIX Software GmbH, Kiel

The software developer ADDIX Software GmbH has specialized in developing and manufacturing telecommunication products. ADDIX's professional Voice over IP telephone system Astimax has been used successfully worldwide since 2005. With its many functions, it offers companies maximum flexibility. The Astimax conference server with WebRTC integration has since extended the portfolio.

### *Björn Schwarze, Managing Partner ADDIX Software GmbH*

"We use EML voice recognition for transcribing voice mail boxes in order to read messages quickly and discreetly as emails and to document conference calls. Because we target business clients, the reliability of voice recognition is a major focus."



# AUTOMOTIVE

Business Sector: Automotive  
Partner: Robert Bosch GmbH, Daimler AG



## The cooperative alternative

### The German research project “KoFFI” uses EML language technology for partially automated automobile driving

Autonomous driving is currently considered the cornerstone of the transportation of tomorrow. Driving without a driver, however, poses major technical, legal, and ethical challenges for automobile manufacturers. The joint project “Cooperative Driver-Vehicle Interaction” (KoFFI) presents an alternative approach: It deals with semi-automated driving – that is, with situations in which it is still generally possible for the driver to intervene. “In this project, we also consider the important ethical, legal, and social issues,” Project Coordinator Dr. Rainer Erbach (Robert Bosch GmbH) explains. “These issues are then taken into account in the design and development of the prototypes.”

The project partners come from the fields of industry and university research. In addition to Robert Bosch GmbH (Car Multimedia), Daimler AG (Research and Advanced Engineering), and EML European Media Laboratory GmbH, other involved parties include the University of Ulm, Heilbronn University, and the Institute of Digital Ethics at Stuttgart Media University. The “KoFFI” project runs for a period of three years (until June 2019) and was funded by the Federal Ministry of Education

and Research (BMBF). It combines speech recognition, natural language processing, and a speech dialog system with graphic interfaces designed to meet the special requirements of semi-automated driving.

The EML provides the required speech recognition methods that are used in the project for the local- and server-based conversion of speech to text. “We develop language components for the KoFFI speech dialog system and adapt it to fit the required context,” EML Manager Research Dr. Volker Fischer explains. Just like all other components of the KoFFI system, the language technology components are also evaluated and optimized via user studies because one of the project’s goals is to increase people’s trust in the technology they need to rely on in case of an emergency.

***Dr. Rainer Erbach, Car Multimedia, Robert Bosch GmbH,  
Coordinator of KoFFI***

“We are developing a multimodal user interface for semi- and highly automated vehicles. It is particularly important to us to permanently improve drivers’ and passengers’ acceptance of and trust in this new technology. To that end, Daimler and EML’s innovative speech recognition and output system plays a critical role.”



## EML: Wir bringen Technik zur Sprache!

EML European Media Laboratory GmbH wurde im Jahr 1997 von SAP-Mitbegründer Klaus Tschira (1940-2015) als privates IT-Unternehmen gegründet. Das EML entwickelt Software und Technologien in der automatischen Sprachverarbeitung. Schwerpunkt ist die automatische Umwandlung von Sprache in Text (Transkription), die in Telefonie (Speech Analytics, Voice Mail) und mobilen Anwendungen (Voice Texting, Voice Search) eingesetzt wird.

### Kontakt

EML European Media Laboratory GmbH  
Berliner Str. 45  
69120 Heidelberg  
Germany

info@eml.org

www.eml.org

© EML 2018

### Ansprechpartner

Allgemeine Anfragen

Anja Varga  
anja.varga@eml.org  
Tel.: +49-6221-533 323  
Fax: +49-6221-533 282

### Geschäftsführer

Prof. Dr.-Ing. Dr. h.c.  
Andreas Reuter  
reuter@eml.org  
Tel. +49-6221-533 201

### Abteilung Forschung

Dr.-Ing. Volker Fischer  
volkerfischer@eml.org  
Tel. +49-6221-533 256

### Abteilung Entwicklung

Markus Klehr  
markus.klehr@eml.org  
Tel. +49-6221-533 261

### Presse und Kommunikation

Dr. Peter Saueressig  
saueressig@eml.org  
Tel.: +49-6221-533 245

@EMLVoice  
EMLEuropeanMediaLaboratory  
EMLVoiceMessaging

## EML: We make technology talk!

EML European Media Laboratory GmbH was established as a private IT enterprise in 1997 by Klaus Tschira (1940-2015), co-founder of SAP. EML develops software and technologies for automatic speech processing. The main focus is on the automatic conversion of speech into text (transcription), which is used in telephony (speech analytics, voicemail) and mobile applications (voice texting, voice search).

### Contact

EML European Media Laboratory GmbH  
Berliner Str. 45  
69120 Heidelberg  
Germany

info@eml.org

www.eml.org

© EML 2018

### Contact Persons

#### General Enquiries

Anja Varga  
anja.varga@eml.org  
Tel.: +49-6221-533 323  
Fax: +49-6221-533 282

#### Managing Director

Prof. Dr.-Ing. Dr. h.c.  
Andreas Reuter  
reuter@eml.org  
Tel. +49-6221-533 201

#### Manager Research

Dr.-Ing. Volker Fischer  
volker.fischer@eml.org  
Tel. +49-6221-533 256

#### Manager Development

Markus Klehr  
markus.klehr@eml.org  
Tel. +49-6221-533 261

#### Press and Communications

Dr. Peter Saueressig  
saueressig@eml.org  
Tel.: +49-6221-533 245

@EMLVoice  
EMLEuropeanMediaLaboratory  
EMLVoiceMessaging

45

BERLINER STRASSE

BERLINER STRASSE

45