

欧州における自動翻訳・音声認識・音声合成技術に係る  
研究開発の動向

平成 28 年 1 月

国立研究開発法人 情報通信研究機構  
(欧州連携センター)

# 目次

はじめに.....	1
全体の要約.....	3
General summary.....	6
第一部 欧州連合の自動翻訳・音声認識・音声合成技術の研究開発助成の基本的方針と機械翻訳サービスの動向.....	9
第一章 欧州連合の第7次枠組計画及びホライゾン 2020 における自動翻訳技術、音声認識技術、音声合成技術の研究開発の助成動向.....	9
第二章 欧州連合の機械翻訳サービス MT@EC の動向.....	9
第二部 欧州における自動翻訳・音声認識・音声合成技術に係る研究開発機関の概要と研究開発の動向.....	11
第一章 英国・エジンバラ大学統計機械翻訳グループ.....	11
組織の概要及び研究トピック.....	11
研究事例.....	11
第二章 アイルランド・ADAPT (旧 CNGL).....	13
組織概要及び研究トピック.....	13
スピノフ企業の概要.....	14
ヒアリング調査メモ / アイルランド・ADAPT (旧 CNGL).....	15
Interview report / Ireland・ADAPT Centre (formerly CNGL) at Dublin City University (DCU).....	18
第三章 ドイツ・カールスルーエ技術研究院インタラクティブシステム研究所 (KIT).....	20
組織概要及び研究トピック.....	20
研究事例.....	21
第四章 ドイツ人工知能研究センター言語技術研究所 (DFKI).....	22
組織概要及び研究トピック.....	22
研究プロジェクト事例.....	22
ヒアリング調査メモ / DFKI 言語技術研究所テキスト分析グループ.....	24
第五章 フランス・メヌ大学情報学研究所 (LIUM).....	28
組織概要及び研究トピック.....	28
スピノフ企業.....	28
ヒアリング調査メモ / 仏 LIUM.....	28

## はじめに

欧州は多民族、多言語が共存する地域であり、コミュニケーションの問題と常に直面してきており、言語技術に対する関心が高く、自動翻訳、音声認識、音声合成技術の研究開発を積極的に実施している。本報告書では、1) 欧州連合の自動翻訳・音声認識・音声合成技術の研究開発助成の基本的方針、2) 欧州における自動翻訳・音声認識・音声合成技術に係る研究機関の概要と研究開発の動向(第7次枠組計画及びホライズン 2020 で助成されている研究プロジェクトを含む) の2点について記す。

### 調査内容

#### 1) 欧州連合の自動翻訳・音声認識・音声合成技術の研究開発助成の基本的方針

- ・ 欧州連合の第7次枠組計画及びホライズン 2020 における同技術の研究開発への助成動向
- ・ 欧州連合諸機関の文書の機械翻訳サービスである MT@EC の動向

#### 2) 欧州における自動翻訳・音声認識・音声合成技術に係る研究開発機関の概要と研究開発の動向

- ・ 英エジンバラ大学統計機械翻訳グループ
- ・ アイルランド・ADAPT (旧 CNGL)
- ・ 独カールスルーエ技術研究院インタラクティブシステム研究所 (KIT)
- ・ ドイツ人工知能研究センター言語技術研究所 (DFKI)
- ・ 仏メヌ大学情報学研究所 (LIUM)

### 調査方法

本調査では、インターネット及び公刊物、報道記事等を精査するとともに、インタビュー調査を行った。2) の調査内容に関して、アイルランドの ADAPT、ドイツ人工知能センター言語技術研究所、フランスのメヌ大学情報学研究所の研究者にインタビュー調査を実施した。

なお、本報告書では、情報を入手したウェブサイトの URL を参考のため注に載せているが、これらの記事はウェブサイト管理運営者の判断で随時移動、修正、削除される可能性がある。従って、本報告書の発表後、注に記された URL から情報源となった記事にアクセスできないことがありうることを、ここで前もって注記しておきたい。



## 全体の要約

以下に、本報告書全体の要約を記す（より詳しい情報については報告書本文をご覧ください）。

### 第一部 欧州連合の自動翻訳・音声認識・音声合成技術の研究開発助成の基本的方針

#### **欧州連合の第七次枠組計画及びホライゾン2020における自動翻訳及び音声技術研究開発への助成動向**

- ・ 特定の言語（英語や仏語等）の自動翻訳技術の開発は進んでおり、翻訳の品質が高くなってきているものの、幾つかの欧州言語に関しては言語技術のサポートが進んでいないことが指摘されており、ホライゾン2020では、EUの全言語間の組み合わせの翻訳品質を向上させることが特に目標とされている。
- ・ 多数の言語が利用されている欧州では、言葉の壁がEUデジタル単一市場形成の障害となっているという認識がある一方で、EUの言語技術への研究開発助成の基本的方針に若干の変更が見られる。ホライゾン2020の2014-2015年度ICT作業プログラムでは、「ICT-17: 言語の壁の打破」という枠組みで、自動翻訳技術の開発支援に対して、1500万ユーロの予算が組まれているものの、2016-2017年度ICT作業プログラムでは「ICT-19: メディアとコンテンツの融合」という枠組みで、メディアとコンテンツにアクセスすることを可能にする技術として、音声転写技術や自動翻訳技術が助成される見込みであり、言語技術へ焦点は置かれていない。このような研究開発助成の基本的方針の変更に対しては、欧州の言語技術分野の研究者から不満の声が上がっている。

#### **欧州連合の機械翻訳サービスMT@ECの動向**

- ・ MT@ECは、ISAプログラムの下で開発されたEU諸機関及びEU加盟国の行政機関向けのオンライン共通機械翻訳サービスである。
- ・ MT@ECは、EU関連の文書を最も正確に翻訳する。
- ・ MT@ECは、EU全公用語間での翻訳を直訳で提供する。したがって、より正確な翻訳が必要な場合は、翻訳の専門家が文書を修正する必要がある。
- ・ MT@EC以前は、ECMTという自動翻訳サービスが利用されていた（2010年12月まで）。ECMTでは、規則ベースの機械翻訳技術が利用されていたが、MT@ECでは統計ベースの技術が使われている。
- ・ 欧州のインフラストラクチャの接続を支援するEUの「CEF (Connecting European Facility)」というプログラムでは、公共機関向けの「CEFAT (CEF Automated Translation)」という機械翻訳プラットフォームも開発も支援している。CEFATはMT@EC上に構築される予定であり、EUの24カ国語及びノルウェー語、アイスランド語の全てのペア（両方向）の翻訳を対象とする。

### 第二部 欧州における自動翻訳・音声認識・音声合成技術に係る研究開発機関の概要と研究開発の動向

#### **英国・エジンバラ大学統計機械翻訳グループ**

- ・ エジンバラ大学統計機械翻訳グループは、機械翻訳システム・MOSESの開発で有名である。

- ・ 同研究グループは、数多くの FP7 プロジェクト及びホライゾン 2020 の研究プロジェクトに参加している。
  - FP7 プロジェクトの例: Accept、Casmacat、EU-Bridge、Matecat、MosesCore、EuroMatrixPlus、LetsMT!
  - ホライゾン 2020 の例: QT21、MMT、HimL、TraMOOC、Cracker

#### **アイルランド・ADAPT (旧 CNGL)**

- ・ アイルランドは、英国と同じく、英語使用国であり、また法人税を軽減していることから、同国に多くの外国企業が欧州本拠地を置いているので、同国では自動翻訳を含めたローカライゼーションの作業が積極的に実施されている。このような社会的背景の中、2007 年に CNGL (Centre for Next Generation Localisation) が設立された。
- ・ 2015 年 1 月から、組織名称を CNGL から ADAPT に変更した。
- ・ ADAPT は、4 つの研究機関、トリニティ・カレッジ・ダブリン、ダブリンシティ大学、ダブリンカレッジ大学、ダブリン技術研究院と産業パートナーの研究者を含む研究センターである。
- ・ ADAPT では、自動翻訳だけでなく、金融サービス、e コマース、メディア、エンターテインメント・ゲーム、デジタル文化・人文科学、e ラーニング等の開発も実施されている。
- ・ ADAPT から、KantantMT.com 社と ICONIC 社という自動翻訳技術関連のスピンオフ企業が 2 つ設立されている。

#### **ドイツ・カールスルーエ技術研究院インタラクティブシステム研究所 (KIT)**

- ・ カールスルーエ技術研究院インタラクティブシステム研究所は、音声及びテキスト機械翻訳、顔追跡、注意追跡、マルチモーダル追跡、ニューラルネットワークの研究開発を行っている。
- ・ 2012 年に、同研究所は、大学講義向けの自動同時音声翻訳システムを発表している。
- ・ テキスト機械翻訳技術に関しては、特に欧州言語に焦点を当てて、研究を実施しており、世界的な機械翻訳評価ワークショップである WMT に数年来参加している。
- ・ テキスト機械翻訳技術の開発を独仏 Quaero プロジェクト、FP7 の EU-Bridge プロジェクト、ホライゾン 2020 の QT21 プロジェクトで進めている。

#### **ドイツ人工知能研究センター言語技術研究所 (DFKI)**

- ・ ドイツ人工知能研究センターでは、機械翻訳を含む言語技術の研究開発を言語技術研究所のテキスト分析グループ (ベルリン市) とマルチ言語技術グループ (ザールブリュッケン市) で実施している。
- ・ ドイツ人工知能研究センター言語技術研究所は、FP7 の QTLeap プロジェクト、ホライゾン 2020 の QT21 プロジェクトを実施している。
- ・ ドイツ人工知能研究センター言語技術研究所は、言語技術分野の欧州研究機関の提携を強化する META-NET のコーディネーターを務めており、欧州で同分野の中心的な研究組織の一つである。また、同研究所は、機械翻訳技術分野の欧州研究機関の提携を強化するホライゾン 2020 CRACKER プロジェクトのコーディネーターを務めている。
- ・ Acrolinx (言語分析エンジンの提供) と Yocoy (機械翻訳モバイルアプリケーションの開発) というスピンオフ企業が設立されている。

#### **フランス・メヌ大学情報学研究所 (LIUM)**

- ・ フランスのル・マン市に設立されたメヌ大学情報学研究所 (LIUM) では、音声言語翻訳、ロバスト音声認識、話者ダイアライゼーション、話者認識、話者言語理解の研究が実

施されており、これら全てのトピックで、特に深層ニューラルネットワークの利用について研究している。

- LIUM から、Deeplingo 社（企業向け機械翻訳のモデルの提供）、Voxolab 社（企業向け話者特定アプリケーションや音声転写アプリケーションの提供）というスピンオフ企業 2 社が設立されている。
- LIUM は、EU の CHIST-ERA の枠組みで、M2CR プロジェクト（2015 年～2017 年）を開始したところである。同プロジェクトでは、深層ニューラルネットワークを利用した音声・画像理解、音声機械翻訳の研究を行う。その他、LIUM は FP7 の MateCat プロジェクト（2011 年～2014 年）で、コンピュータ翻訳支援ツール（CAT）の開発を行った。

## General summary

This is a general summary of the “Report on the R&D situation for automated translation, speech recognition and speech synthesis in Europe”. See the text of the report for more information.

### Part 1- The EU support policy for the R&D of automated translation, speech recognition and speech synthesis

#### *The EU’s R&D support policy for automated translation and speech technologies in FP7 and Horizon 2020*

- For some particular languages (English or French), automated translation technologies are developed and the quality of translation has been improved. But for the other languages, this progress is at a standstill. Horizon 2020 aims to improve the quality of translations into all EU languages.
- But, though some people consider language barrier is an obstacle for European digital single market, the EU’s R&D support Policy for language technologies has changed in Horizon 2020.
- The workframe budget for the R&D for automated translations has barely remained in the ICT Work Programme 2014-2015 of Horizon 2020 (the research topic called “ ICT-17 : Cracking the language barriers amounts to €15 million), but language technologies together with transcription and automated translation applications will be supported in the research topic called “Media and content convergence” as technologies enabling access to media and content in the Work Programme 2015-2016. So, Horizon 2020 doesn’t focus on language technologies any more. Some researchers have expressed their dissatisfaction with this lack of financial support for language technologies.

#### *Machine translation service of the EU : MT@EU*

- MT@EC is the online machine translation service developed under ISA programme (the Interoperability Solutions for Public Administrations programme). This is a service for officials in the EU institutions and EU member states’ public administrations.
- MT@EC provides the most precise of EU documents.
- MT@EC translates literally from and into all EU official languages. When more precision is required, translation should be improved by human translators.
- Before the introduction of MT@EC, another automated translation service, called ECMT, had been used (until December 2010). While ECMT is built on rule-based machine translation technology, MT@EC is built on statistical machine translation technology.
- The EU’s Connecting European Facility programme (CEF) supports the R&D of machine translation platform for public administrations in Europe, called CEFAT (CEF Automated Translation). CEFAT will be built on MT@EC for translating the EU 24 official languages, Norwegian and Icelandic.

### Part 2 The R&D situation regarding automated translation, speech recognition and speech synthesis technologies in Europe

#### *UK: Statistical Machine Translation Group at the University of Edinburgh*

- The Statistical Machine Translation Group at the University of Edinburgh is famous for the development of a machine translation system called MOSES.
- This research group participates in many FP7 and Horizon 2020 research projects.

- Examples of FP7 projects include: Accept, Casmacat, EU-Bridge, Matecat, MosesCore, EuroMatrixPlus, LetsMT .
- Examples of Horizon 2020: QT21, MMT, HimL, TraMOOC, Cracker

***Ireland: ADAPT (ex CNGL)***

- Like the United Kingdom, Ireland is an English-speaking country. As corporate tax is low, many foreign companies have set their European headquarters in Ireland. As a result, the localisation activities, including automated translation, are very active. In order to help with this social background, the Centre for Next Generation Localisation (CNGL) was established in 2007.
- This research center has been renamed ADAPT since January 2015,
- In ADAPT, the researchers from four universities (Trinity College Dublin, Dublin City University, University College Dublin and Dublin Institute of Technology) and the researchers from their industrial partners work together.
- In addition to automated translation, ADAPT also develops financial services, E-commerce, digital media, entertainment games, digital culture and human sciences, E-learning, etc.
- KantantMT.com and ICONIC are ADAPT's spin-off companies developing automated translation technologies.

***Germany: Interactive Systems Lab at Karlsruhe Institute of Technology (KIT)***

- Interactive Systems Lab at Karlsruhe Institute of Technology researches and develops speech and text automated translation, face tracking, attention tracking, multi-modal tracking, and neural network.
- In 2012, this laboratory delivered an automatic simultaneous speech translation system for university lectures.
- As regards text machine translation technology, this institute research focuses on European languages. And it has for several years participated in WMT, an International workshop on statistical machine translation.
- Concerning the development of text automated translation technology, it has participated into Franco-German Quairo project, FP7's EU-Bridge project, and Horizon 2020's QT21 project.

***Germany: German Research Center for Artificial Intelligence (DFKI)***

- The German Research Center for Artificial Intelligence conducts the research and development for language technology, including machine translation, in two groups : one's name is Text analytics (in Berlin) and another is Multilingual technologies (in Saarbrücken).
- The language technology lab at DFKI is a key research centre in Europe. It coordinates META-NET, which enhances an alliance of European research centers for language technology. It is also the coordinator of the CRACKER project funded through Horizon 2020, which strengthens the alliance of European research centers for machine translation.

***France: Informatics Laboratory at the University of Maine (LIUM)***

- The Informatics Laboratory at the University of Maine, established in Le Mans, France, studies speech language translation, robust speech recognition, speaker diarization, speaker recognition and speaker language understanding. It is involved in research on the applicability of deep neural networks into all these research fields.

- Deeplingo and Voxolab are the LIUM's spin-off companies. Deeplingo offers customized machine translation models for companies. Voxolob offers customized speaker identification application and customized transcription application for companies.
- LIUM has just launched the M2CR project (2015-2017) through the EU's CHIST-ERA. This project aims to establish a unified architecture, based on deep neural networks, for speech understanding and translation and image understanding. LIUM also participated into FP7's MateCat project (2011-2014) for the development of the Computer-assisted translation (CAT).

## 第一部 欧州連合の自動翻訳・音声認識・音声合成技術の研究開発助成の基本的方針と機械翻訳サービスの動向

第一部では、欧州連合（以下、EU とする）の自動翻訳技術、音声認識技術、音声合成技術に対する研究開発助成の基本的方針と EU の機械翻訳サービスである「MT@EC」の動向について記す。

### 第一章 欧州連合の第7次枠組計画及びホライゾン2020における自動翻訳技術、音声認識技術、音声合成技術の研究開発の助成動向

- ・ 2007年から2013年にかけて実施されたEUの第七次枠組計画（以下、FP7とする）と競争・イノベーション枠組計画（以下、CIPとする）<sup>1</sup>では、自動翻訳技術、音声認識技術、音声合成技術に対して積極的に支援されている（30プロジェクト以上）<sup>2</sup>。
- ・ しかし、ホライゾン2020では、言語技術分野の研究開発助成に対する基本的方針の若干の変更が見られる。ホライゾン2020の2014-2015年度ICT作業プログラムでは、2年間で自動翻訳技術の研究開発に対して、「ICT-17：言語の壁の打破」という枠組みで、1500万ユーロの予算が組まれているものの、最新の2016-2017年度ICT作業プログラム<sup>3</sup>では、「ICT-19：メディアとコンテンツの融合」という枠組みで、メディアとコンテンツにアクセスすることを可能にする技術として、音声転写技術や自動翻訳技術が助成される予定であり（ICT-19の予算は全体で3900万ユーロであり、自動翻訳と音声技術以外の技術への助成も含む）、言語技術への助成に焦点は置かれていない。このような研究開発に対する基本的方針の変更には、欧州の研究者から不満の声が上がっており、欧州の機械翻訳技術分野の研究機関を提携させることを目的とするホライゾン2020のCRACKERプロジェクトは、方針を変更するように欧州議会へ働きかけている。
- ・ ホライゾン2020の2014-2015年度ICT作業プログラムの「ICT-17：言語の壁の打破」では、特定の言語（英語や仏語など）の自動翻訳技術の開発は進んでおり、翻訳の品質が高くなってきているものの、他の欧州言語に関しては、開発事情が同様ではないことが指摘されており、EUの全言語間の組み合わせの翻訳品質を向上させることが目標とされている。

### 第二章 欧州連合の機械翻訳サービスMT@ECの動向

- ・ MT@ECは、ISAプログラム（Interoperability Solution for Public Administrations Programme）の下で開発されたEU諸機関及びEU加盟国の行政機関向けのオンライン共通機械翻訳サービスである<sup>4</sup>。

<sup>1</sup> CIPは、EUの中小企業向けの研究開発助成プログラムである。

[http://ec.europa.eu/cip/index\\_en.htm](http://ec.europa.eu/cip/index_en.htm)

<sup>2</sup> EUのウェブサイトには、プロジェクトの一覧がある。

[http://cordis.europa.eu/fp7/ict/language-technologies/projects\\_en.html](http://cordis.europa.eu/fp7/ict/language-technologies/projects_en.html)

<sup>3</sup> [http://ec.europa.eu/research/participants/data/ref/h2020/wp/2016\\_2017/main/h2020-wp1617-leit-ict\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/wp/2016_2017/main/h2020-wp1617-leit-ict_en.pdf)

<sup>4</sup> [http://ec.europa.eu/isa/documents/publications/brochure-mt@ec-a5-v3\\_en.pdf](http://ec.europa.eu/isa/documents/publications/brochure-mt@ec-a5-v3_en.pdf)

- ・ 欧州委員会の ISA プログラムは、EU 諸機関及び EU 加盟国行政機関のあいだのやり取りを円滑にするために、1 億 6000 万ユーロの予算で、電子政府サービス (eGovernment) のツール、サービス、フレームワークを開発する 40 の活動を実施している<sup>5</sup>。MT@EC はその活動のうちの一つである。
- ・ 以下、MT@EC の主な概要を記す。
  - 欧州委員会、その他の EU 諸機関、EU 加盟国行政機関により利用される文書は大量にあり、人の手によって全てを翻訳することは不可能である。MT@EC は、以上の問題を解決するためオンライン機械翻訳サービスを提供する。
  - EU 諸機関、EU 加盟国内の行政機関で勤務していれば、無料で MT@EC による機械翻訳サービスを利用できる。
  - MT@EC は、EU に関連する内容の文書の翻訳において最も正確に働く。
  - MT@EC は、全ての EU の公用語間の翻訳を直訳で提供する。したがって、より正確な翻訳が必要な場合は、翻訳の専門家が文書を修正する必要がある。
  - MT@EC は、2010 年 12 月まで利用されていた同様のサービス、ECMT (European Commission Machine Translation) に取って代わり、導入された。ECMT では、規則ベースの機械翻訳技術が利用されていたが、MT@EC では統計ベースの技術が使われている<sup>6</sup>。
  - MT@EC を利用するには専用のアカウントが必要である。EU 諸機関、そして、EU 加盟国内の行政機関に勤務する者はアカウントを持ち、パスワードを使い、機械翻訳サービスにアクセスできる。
- ・ 欧州のインフラストラクチャの接続を支援する EU の CEF (Connecting European Facility)<sup>7</sup> というプログラムは、公共機関向けの「CEFAT (CEF Automated Translation)」<sup>8</sup> という機械翻訳プラットフォームを開発している。CEFAT は、MT@EC 上に構築され、EU の 24 カ国語及びノルウェー語、アイスランド語の全てのペア (両方向) の翻訳を対象とする予定である。

<sup>5</sup> [http://ec.europa.eu/isa/index\\_en.htm](http://ec.europa.eu/isa/index_en.htm)

[http://ec.europa.eu/isa/ready-to-use-solutions/index\\_en.htm](http://ec.europa.eu/isa/ready-to-use-solutions/index_en.htm)

<sup>6</sup> [http://ec.europa.eu/isa/actions/02-interoperability-architecture/2-8action\\_en.htm](http://ec.europa.eu/isa/actions/02-interoperability-architecture/2-8action_en.htm)

<sup>7</sup> <http://ec.europa.eu/digital-agenda/en/connecting-europe-facility>

<sup>8</sup> [https://joinup.ec.europa.eu/community/cef/og\\_page/catalogue-building-blocks](https://joinup.ec.europa.eu/community/cef/og_page/catalogue-building-blocks)

[http://www.rigasummit2015.eu/sites/rigasummit2015.eu/files/cef\\_29\\_04\\_2015\\_spyridon\\_pilos\\_building\\_the\\_automated\\_translation\\_platform.pdf](http://www.rigasummit2015.eu/sites/rigasummit2015.eu/files/cef_29_04_2015_spyridon_pilos_building_the_automated_translation_platform.pdf)

## 第二部 欧州における自動翻訳・音声認識・音声合成技術に係る研究開発機関の概要と研究開発の動向

第二部では、欧州における自動翻訳・音声認識・音声合成技術に係る研究開発の動向として、英エジンバラ大学統計機械翻訳グループ、アイルランド・ADAPT、独カールスルーエ技術研究院インタラクティブシステム研究所（KIT）、ドイツ人工知能研究センター言語技術研究所（DFKI）、仏メヌ大学情報学研究所（LIUM）の組織概要及び研究トピック、研究事例等について記す。

### 第一章 英国・エジンバラ大学統計機械翻訳グループ

#### 組織の概要及び研究トピック

- ・ 英国のエジンバラ大学統計機械翻訳グループ（Statistical Machine Translation Group）<sup>9</sup>は、統計ベースの機械翻訳技術の研究開発を実施している。
- ・ 全人員は、博士課程の学生を含めて、17名である。
- ・ エジンバラ大学統計機械翻訳グループは、機械翻訳システム・MOSESの開発で有名である。だが、MOSESの開発者の一人であるフィリップ・コーエン教授<sup>10</sup>はエジンバラ大学に籍を残しているものの、2014年1月から主な研究活動の拠点を米ジョン・ホプキンス大学に移している。
- ・ 多くのFP7プロジェクト及びホライゾン2020の研究プロジェクトに参加している<sup>11</sup>。
  - FP7プロジェクトの例：ACCEPT、CASMACAT、EU-Bridge、MATECAT、MOSESCORE、EuroMatrixPlus、LetsMT!
  - ホライゾン2020の例：QT21、MMT、HimL、TraMOOC、Cracker
- ・ 軍事向け技術の研究開発支援を行うDARPA（アメリカ）から助成を受けたAGILEプロジェクトにも参加していた。AGILEプロジェクトでは、アラビア語と中国語向けの音声とテキスト機械翻訳技術の開発を目的とした。

#### 研究事例

##### ホライゾン2020：MMTプロジェクト

- ・ MMTプロジェクトは、新しいオープンソースの機械翻訳アーキテクチャに基づくオンライン商用翻訳サービスを開発する。

##### MMTプロジェクトの概要

略称	MMT
正式タイトル	Modern Machine Translation
全予算（EU拠出分）	3 695 200 ユーロ（2 994 700 ユーロ）
プロジェクト期間	2015年1月～2018年1月（36ヶ月）

<sup>9</sup> <http://www.statmt.org/ued/>

<sup>10</sup> <http://homepages.inf.ed.ac.uk/pkoehn/>

<sup>11</sup> <http://www.statmt.org/ued/?n=Public.Projects>

コーディネーター	Translated 社 (伊)
参加組織	ブルーノ・ケスラー基金 (伊) 、 <b>エジンバラ大学</b> (英) 、 TAUS (蘭)
ウェブサイト	<a href="http://www.modernmt.eu">http://www.modernmt.eu</a>

### 研究内容

MMT プロジェクトは、現在機械翻訳の受容を妨げている 4 つの技術的な課題に挑戦する<sup>12</sup>。

- MMT はトレーニング段階を要求しない。トレーニングデータが与えられれば、MMT はすぐに翻訳を開始できる状態にある。
- MMT は自動的にコンテキストに対処するので、各ドメイン向けに特殊なシステムを構築する必要がない。MMT は、コンテキストに結びついた情報とトレーニングセグメントを一緒に供給することによって、どんなトピックとドメインにも最良の翻訳品質を提供する。
- MMT システムはデータとユーザに応じて変更可能なので、高価でアドホックなハードウェアインフラストラクチャを必要としない。
- MMT は、大手ウェブ企業と機械翻訳産業にあるデータ量のギャップを埋めることを加速させるデータ収集のインフラストラクチャを築く。

### ホライゾン 2020 : HimL プロジェクト

- HimL プロジェクトは、保健情報を複数の言語で公開する機械翻訳技術を開発する。

### HimL プロジェクトの概要

略称	HimL
正式タイトル	Health in my Language
全予算 (EU 拠出分)	3 051 383, 75 ユーロ (2 949 571, 25 ユーロ)
プロジェクト期間	2015 年 2 月～2018 年 1 月 (36 ヶ月)
コーディネーター	<b>エジンバラ大学</b> (英)
参加組織	プラハ・カレル大学 (チェコ) 、 ルートヴィヒ・マクシミリアン大学ミュンヘン (独) 、 LINGEA SRO (チェコ) 、 NHS24 (英) 、 コ克蘭共同計画 (英)
ウェブサイト	<a href="http://www.himl.eu">http://www.himl.eu</a>

### 研究内容

- 統計とディープ言語技術を組み合わせ、健康・医療情報向けの自動機械翻訳システムを開発する。
- 最終的に開発された機械翻訳システムは、NHS24 (英スコットランドの国立遠隔保健・介護機関) とコ克蘭 (英健康・医療情報の提供プログラム) のウェブシステムに統合される予定である。両機関は、市民向けに健康・医療情報をオンラインで提供しており、情報を特に東欧言語で提供する必要がある。
- Himl の課題は、1) ドメインに適応すること、2) 正確な翻訳を提供すること、3) 中欧・東欧言語のような形態論的に豊かな言語をサポートすることである。
- 特に英語をチェコ語、ポーランド語、ルーマニア語、ドイツ語に翻訳するテストを行う予定である。現在、テストセットと第一次年の機械翻訳システムを開発したところである。

<sup>12</sup> [http://cordis.europa.eu/project/rcn/194327\\_en.html](http://cordis.europa.eu/project/rcn/194327_en.html)

テストセットは、NHS24 とコクランから代表的なテキストを抽出し、4つのターゲットとなる言語へと翻訳した。その際、両組織から3万言語を収集している。

#### ホライゾン 2020: TraMOOC プロジェクト

- TraMOOC プロジェクト<sup>13</sup>は、オンライン講義のための機械翻訳プラットフォームを開発する。

#### TraMOOC プロジェクトの概要

略称	TraMOOC
正式タイトル	Translation for Massive Open Online Courses
全予算 (EU 拠出分)	3 270 710 ユーロ (3 081 147, 5 ユーロ)
プロジェクト期間	2015年2月～2018年1月 (36ヶ月)
コーディネーター	フンボルト大学ベルリン (独)
参加組織	ダブリンシティ大学 (アイルランド)、 <b>エジンバラ大学 (英)</b> 、イオニアン大学 (ギリシア)、カトリック大学 (蘭)、EASN 技術イノベーションサービス (ベルギー)、Deluxe メディア・ヨーロッパ (英)、ティルブルフ大学 (蘭)、IVERSITY (独)、ナレッジ4オール基金 (英)
ウェブサイト	<a href="http://tramoc.eu">http://tramoc.eu</a>

#### 研究内容

- オンライン講義向けの大規模オープン機械翻訳プラットフォームを提供する。
- データの分野は、宿題、テスト、プレゼンテーション、講義の字幕、ブログ等である。
- 英語から欧州言語 (ブルガリア、ドイツ、ギリシア、オランダ、ポーランド、ポルトガル、クロアチア、ハンガリー、チェコ) と BRICs の 2 カ国 (中国、ロシア) の言語への翻訳を対象とし、多くの機械翻訳サービスでサポートされていない言語も含む。
- サービスの核となる部分はオープンソースで公開され、付加サービスは商用化される予定である。
- TraMOOC のプラットフォームは、オープンオンライン講義を受講したい欧州及び世界各国の市民、大規模オープンオンライン講義プロバイダ、機械翻訳開発者、言語技術分野の技術者向けに開発される。

## 第二章 アイルランド・ADAPT (旧 CNGL)

#### 組織概要及び研究トピック

- アイルランドは、英国と同じく、英語使用国であり、また法人税を軽減していることから、同国に多くの外国企業が欧州本拠地を置いているので、同国では自動翻訳を含めたローカライゼーションの作業 (ソフトウェアを現地国の言語に適応させる作業等) が積極的に実施されている。
- このような社会的背景の中、2007年に次世代ローカライゼーションセンター (Centre for Next Generation Localisation : CNGL)<sup>14</sup>が、アイルランド科学基金 (SFI : Science foundation

<sup>13</sup> <http://www.tramoc.eu>

Ireland)<sup>15</sup>と産業界から資金を供給され、設立された。CNGLは公立の大学研究機関と産業を提携させることを目的に設立され、ローカライゼーションに係る研究開発を産官学で推進する手段である。

- ・ 2015年1月からSFIの資金供給の第三期に入り（2007年から2012年が第一期、2013年から2014年が第二期にあたる）、それに伴い、組織名称をCNGLからADAPTに変更した。第一期から第二期にかけて、SFIと産業界などから1億ユーロ以上の資金を得ている。
- ・ 2014年10月、アイルランド政府はSFIを通して、2015年から2020年の6年間に渡って、1億5500万ユーロを拠出し、産業界とともに5つの研究機関を設立することを決定した<sup>16</sup>。産業界の拠出は9000万ユーロであり、合計で2億4500万ユーロが予算である。ADAPTはこれらの機関の一つであり、SFIと産業界から合わせて5000万ユーロの資金投入が決定されている（SFIは2400万ユーロの拠出）。
- ・ ADAPTでは、自動翻訳だけでなく、金融サービス、eコマース、メディア、エンターテインメント・ゲーム、デジタル文化・人文科学、eラーニング等の開発も実施されている。
- ・ ADAPTは140以上の企業と関係を持ち、産業サポーターは19社である（独 Acrolinx、米 AOL、米 CISCO、英 COMMETRIC、米 Ebay、アイルランド・FBD、中 Huawei、中 IFLYTEK、米 Intel、米 Microsoft、チェコ共和国・Moravia、スイス・Novartis、米 PayPal、米 Sajan、米 Symantec、米 Welocalize、アイルランド・VistaTec、アイルランド・KantanMT.com、アイルランド・Xanadu）。特に、アメリカ企業と中国企業との結びつきが強い。

### スピノフ企業の概要

- ・ ADAPT から、KantanMT.com 社と ICONIC 社という自動翻訳技術関連のスピノフ企業が2つ設立されている。以下に、2社の概要と提供製品・サービスについて記す。

#### KantanMT.com 社

- ・ ダブリンシティ大学がKantanMT.com社のパートナーであり、大学キャンパス内に同社は事務所を持つ。ベンチャーキャピタル（DELA パートナーズ社）と政府組織（ENTREPRISE アイルランド）から投資を受けている。
- ・ KantanMT.com 社はクラウドベースの機械翻訳サービスを提供する。
- ・ ユーザは、KantanMT.com 社のプラットフォーム上で、機械翻訳エンジンを直接カスタマイズでき、特殊なハードウェアやソフトウェアを設置することを必要としない。
- ・ KantanMT.com の統計的機械翻訳システムには、KantanAnalytics という翻訳品質評価を行う機能が付属しており、翻訳された文章の品質評価スコア（パーセンテージという表記で）を示す。同機能によって、事前に最終的に必要となる経費、期間を知ることが可能になる<sup>17</sup>。同機能は、ADAPT のダブリンシティ大学と提携して開発された。
- ・ 代表的なユーザ事例は、翻訳企業等の言語サービスプロバイダであり、その他、ソフトウェア開発企業、工業系企業、出版企業、観光業、ビッグデータ企業等である。

---

<sup>14</sup> <http://www.cngl.ie>

<sup>15</sup> <http://www.sfi.ie>

SFIは科学及び工学部門のアイルランド国立研究助成機関である。

<sup>16</sup>

<http://www.sfi.ie/news-resources/press-releases/245-million-to-be-invested-in-five-new-world-class-sfi-research-centres-in-ireland.html>

<sup>17</sup> KantanAnalytics については、ウェブサイトを参考のこと。

<https://www.kantanmt.com/overview-analytics.php>

## ICONIC 社

- ICONIC 社は、KantanMT.com 社と同じように、ダブリンシティ大学キャンパス内に事務所を持つ。
- ICONIC 社の統計的機械翻訳システムのコア技術は、「Ensemble Architecture」という名称である。この技術は翻訳の対象である言語だけでなく、翻訳されるコンテンツにも焦点を置いている。Ensemble Architecture は、単なる統計的機械翻訳システムでは品質の良い翻訳が困難な法律、薬品、医療などの複雑な文章を翻訳することを目的として開発された。Ensemble Architecture では、統計的機械翻訳技術というデータ工学に、ドメイン適応という言語工学を融合させている。Ensemble Architecture では、統計ベースの手法、規則ベースの手法、あるいは、これらのハイブリッドな手法等というように、手法を単一のものに限ることなく、言語、構文、意味の情報を含む全てのパラダイムに由来するプロセスを結合する<sup>18</sup>。
- ICONIC 社の主要製品のの一つは、特許及び法律文書向けの統計的機械翻訳システムである IPTranslator<sup>19</sup>であり、Ensemble Architecture が利用されている。欧州特許庁が英語とポルトガル語間の翻訳に IPTranslator を選択している<sup>20</sup>。

## ヒアリング調査メモ / アイルランド・ADAPT (旧 CNGL)

- 日程：平成 27 年 9 月 4 日（金）午前 10 時～
- 場所：先方事務所（アイルランド・ダブリン）
- 先方：ADAPT (旧 CNGL) 自然言語処理グループリサーチリーダー及び主任研究員 兼 ダブリンシティ大学機械計算学部教授：Qun Liu 氏<sup>21</sup>

## ヒアリング概要

### ADAPT (旧 CNGL) について

- CNGL<sup>22</sup>はアイルランド科学基金（SFI: Science foundation Ireland）から資金提供を受けて、2007 年に設立された。
- SFI の資金供給は、2007 年から 2012 年が第一期、2013 年から 2014 年が第二期にあたり、現在は第三期（2015 年～2020 年）にあたる。
- 2015 年 1 月から、新しいパートナーとともに組織再編し、CNGL から ADAPT<sup>23</sup>へと名称を変更した。
- CNGL（Centre for Next Generation Localisation）という名称には、ローカライゼーションという語が含まれている。アイルランドにとって、翻訳を含め、ローカライゼーションは非常に重要である。なぜなら、アイルランドは英国と同じく英語使用国であるとともに、法人税を軽減しているため、多くの外国企業がアイルランドに欧州本拠地を置き、ローカライゼーション、すなわち、ソフトウェアなどを欧州言語に適応させる作業を行っている。このような理由で、アイルランド政府は同研究組織を積極的に支援している。

<sup>18</sup> ICONIC 社のウェブサイトに Ensemble Architecture の図が掲載されている。

[http://iconictranslation.com/machine-translation/ensemble\\_architecture/](http://iconictranslation.com/machine-translation/ensemble_architecture/)

<sup>19</sup> <http://iconictranslation.com/products/iptranslator/>

<sup>20</sup> <http://iconictranslation.com/case-studies/epo/>

<sup>21</sup> <http://computing.dcu.ie/~qliu/>

<sup>22</sup> <http://www.cngl.ie>

<sup>23</sup> <http://adaptcentre.ie>

- ADAPT は、4つの研究機関、トリニティ・カレッジ・ダブリン、ダブリンシティ大学、ダブリンカレッジ大学、ダブリン技術研究院と産業パートナーの研究者を含む研究センターである。大まかな数字であるが、博士課程の学生や事務なども含めて、ADAPT の全人員は 200 名を超える。
- ADAPT では、ローカライゼーションに係る研究開発を実施しているため、多くの応用研究を実施しており、産業との結びつきが密接である一方で、基礎研究も行っており、多くの研究論文を発表している。
- ADAPT の全予算は、6年間で 2400 万ユーロである。
- ADAPT (トリニティ・カレッジ・ダブリン所属) の音声合成分野の研究者、Nick Campbell 氏<sup>24</sup>は過去に NICT 及び国際電気通信基礎技術研究所 (ATR) で研究を実施していたことがある。日本に帰国後、ADAPT で勤務している。

#### ADAPT における機械翻訳の研究活動について

- ダブリンシティ大学の ADAPT では、自然言語処理と機械翻訳という二つの研究が実施されている。Qun Liu 氏は自然言語処理分野のチームリーダーであるが、同時に機械翻訳分野の開発にも関わっている。機械翻訳分野のチームリーダーは Andy Way 氏である<sup>25</sup>。Way 氏は、ADAPT の副所長も務めている。
- ADAPT では、基礎研究のほか、機械翻訳分野のアプリケーション指向の研究が実施されている (Qun 氏の研究ではない)。例えば、機械翻訳の翻訳品質 (quality estimation) や後編集 (post-editing) に係る優れた研究が実施されている。
- ADAPT は、統計機械翻訳と翻訳メモリを結合する研究においても優れた成果を残している。
- ADAPT では、主に欧州言語を対象に機械翻訳の研究開発が実施されている (英語、仏語、ポーランド語など)。アジア言語の開発は少ない。どの言語を対象とするかは、プロジェクト次第である。
- 機械翻訳の対象分野についても同様であり、どの分野 (特許や医療など) を対象とするかは、プロジェクト次第である。
- ADAPT では、機械翻訳技術の開発のために、対訳コーパスを基本的に収集していない (アイルランド語以外)。研究開発の際には、機械翻訳の分野 (特許、マニュアルなど) ごとに、産業パートナーが必要な対訳コーパスを用意する。

#### Qun Liu 氏の研究関心及び研究活動について

- Qun 氏は 2012 年よりダブリンシティ大学及び CNGL (現 ADAPT) に勤務している。CNGL 以前は、中国の科学アカデミーに所属しており、中国語で初めてのオープンソース単語分割システムである ICTCLAS<sup>26</sup>を開発した。以前は、Qun 氏は規則ベースの機械翻訳の研究を実施していた。
- Qun Liu 氏は過去に国際ユニバーサルコミュニケーションシンポジウム (IUCS) に NICT 研究者とともに参加したことがある。
- Qun 氏は、テキストの機械翻訳に係る基礎研究を実施しており、特にアルゴリズム、統計翻訳モデルについて研究している。音声技術の研究は実施していない。また、産業と直接結びつく応用研究は実施していない。

<sup>24</sup> <http://www.speech-data.jp/nick/cv.html>

<sup>25</sup> <http://www.computing.dcu.ie/~away/>

<sup>26</sup> <http://sewm.pku.edu.cn/QA/reference/ICTCLAS/FreeICTCLAS/English.html>

- ・ Qun 氏は基礎研究を実施しているので、多くの言語と分野の一般的な問題に焦点を当てており、どれか特定の言語の機械翻訳、また同様に、特定の分野（特許、マニュアルなど）の機械翻訳について研究しているわけではない。
- ・ FP7 に関しては、統計機械翻訳技術分野の若手研究者を養成することを目的とする EXPERT プロジェクト<sup>27</sup>に参加している。
- ・ HORIZON2020 に関しては、DFKI がコーディネーターを務める QT21 プロジェクトに参加している。QT21 プロジェクトは、特に欧州言語を対象に統計機械翻訳の品質を向上させることを目的にしている。問題は、例えば英語や仏語などの幾つかの言語の間では機械翻訳の品質は非常に高いが、欧州では多くの言語が利用されており、一定の言語間では翻訳の品質が低いことである。同プロジェクトにおいて、Qun 氏はダブリンシティ大学側の研究責任者である。同プロジェクトの予算は中型規模であるが、参加組織数は多い。

#### ADAPT の技術移転の実績について

- ・ CNGL から、ICONIC 社と KantanMT.com 社がスピンオフ企業として生まれている。
- ・ ICONIC 社<sup>28</sup>は様々な分野に対応する機械翻訳サービスを提供している。特に特許翻訳の分野で実績がある。
- ・ KantanMT.com 社<sup>29</sup>が企業向けにクラウドベースの機械翻訳サービスを提供している。

#### 機械翻訳分野におけるアメリカと欧州の違いについて

- ・ アメリカでは、DARPA という軍事目的の技術開発を支援する機関によって、機械翻訳技術の研究助成がされており、英語とアラビア語や中国語間の機械翻訳技術が研究されている。これに対して、欧州では産業目的のために多数ある欧州言語間の機械翻訳技術が実施されており、アプリケーションの種類も多い。したがって、欧州の方が機械翻訳の需要がアメリカよりも多い。このような市場状況を踏まえて、ADAPT では機械翻訳の研究開発を盛んに実施している。
- ・ 補足的な情報になるが、2004 年から 2014 年の間、米グーグル社の機械翻訳部門の責任者であったフランツ・ジョセフ・オック (Franz Josef Och) は欧州出身者である。同氏はドイツの RWTH アーヘン大学で研究を実施し、博士号を取得している。2014 年に、同氏はグーグルを退社している。
- ・ オープンソースの統計的機械翻訳ツールで最も人気のある MOSES は、欧州の大学によって開発されている (英エジンバラ大学の主導)。

#### 参考資料

##### FP7 プロジェクト (Liu 氏が参加しているプロジェクトの概要)

略称	EXPERT
正式タイトル	EXPloting Empirical appRoaches to Translation
全予算 (EU 拠出分)	3 935 340,3 ユーロ (3 935 340,3 ユーロ)
プロジェクト期間	2012 年 10 月～2016 年 9 月 (48 ヶ月)
コーディネーター	英ウルヴァーハンプトン大学

<sup>27</sup> <http://expert-itn.eu>

<sup>28</sup> <http://iconictranslation.com>

<sup>29</sup> <https://www.kantanmt.com>

参加組織数	スペイン・PANGAONIC B I EUROPA SLU、マラガ大学（スペイン）、シーフィールド大学（英）、ザールランド大学（伊）、TRANSLATED（伊）、ダブリンシティ大学（アイルランド）、HERMES 翻訳・言語サービス（スペイン）、アムステルダム大学（蘭）
ウェブサイト	<a href="http://expert-itn.eu">http://expert-itn.eu</a>
研究内容	翻訳メモリと機械翻訳技術の分野で、若手研究者の育成を支援する。

### Interview report / Ireland · ADAPT Centre (formerly CNGL) at Dublin City University (DCU)

- Dates: September 4, 2015 (Friday) 10:00 -
- Location: ADAPT's office, Dublin City University (Dublin, Ireland)
- Interviewee: Dr. Qun Liu

Research Leader and Principal Investigator of Natural Language Processing Group, ADAPT Centre (formerly CNGL)

Professor of the School of Computing, Dublin City University

#### Summary

##### *ADAPT Centre (formerly CNGL)*

- In 2007 CNGL (Centre for Next General Localisation) was established with funding from SFI (Science Foundation Ireland).
- There are three rounds of SFI funding up to now. The first round is from 2007 to 2012, its second round is from 2013 to 2014, and its third round is from 2015 to 2020.
- From January 2015, CNGL was reorganized. Its new name is ADAPT.
- CNGL (Centre for Next Generation Localisation) contains the term “localisation” in its name. For Ireland, localisation (including translation) is very important. Ireland is an English-speaking country and its corporate tax is low. So, many foreign companies put their European headquarters in Ireland and carry out localisation activities, that is, adapt their software in Europe language. For this reason, the Irish government has been actively supporting this research organization.
- In ADAPT the researchers from four universities (Trinity College Dublin, Dublin City University, University College Dublin and Dublin Institute of Technology) and the researchers from its industry partners work together. The number of its staff, including doctoral students and management and support team, amounts to more than 200.
- ADAPT carries out not only applied research in the close relationship with the industry, but also basic research. It publishes many research papers.
- The total budget for ADAPT Centre for 6 years is 50 million euro (the part from SFI is 24 million euro).
- Dr. Nick Campbell, researcher of the speech synthesis in ADAPT (Trinity College Dublin affiliation), once worked for NICT and ATR.

#### ***Research activities of machine translation at ADAPT at DCU***

- ADAPT at DCU carries out two research activities: natural language processing and machine translation. Dr. Qun Liu is the research leader of the natural language processing. He is also engaged in the research of machine translation. Dr. Andy Way is the research leader of machine translation. He is also the deputy director of ADAPT.
- ADAPT has carried out not only the basic research, but also the application-oriented research of machine translation (which is not the research field of Dr. Qun), for example, quality estimation and post-editing of the machine translation.
- ADAPT achieves excellent results for the research combining statistical machine translation and translation memories.
- ADAPT focuses primarily on the European languages for machine translation (English, French, Polish, etc.). The Asian languages are a small part of its research. The selection of language depends on demands of the specific projects.
- The research field of machine translation (such as patents and medical care) also depends on demands of the projects.
- ADAPT does not basically keep translation corpus necessary for the innovation of machine translation technology (except the Irish). Its industry partners prepare the translation corpus for the particular field of machine translation in which they are interested (patents, manuals, etc.).

#### ***Research of Dr. Qun Liu***

- Dr. Qun works for Dublin City University and CNGL (currently ADAPT) since 2012. Before working for CNGL, he belonged to the Academy of Sciences of China, and he developed ICTCLAS (Institute of Computing Technology, Chinese Lexical Analysis System), that is, the first open source lexical analysis system for Chinese. Previously, he had conducted a research of the rule-based machine translation.
- Previously, Dr. Qun Liu has participated with NICT researchers into the International Universal Communication Symposium (IUCS). He also has visited NICT for one month in 2004.
- Dr. Qun conducts basic research related to machine translation of text, especially, algorithm and statistical language model. He doesn't study speech technology. His research is not applied research directly connected to the industries.
- Dr. Qun conducts basic research, so his research is focusing on general problems among many languages and fields rather than particular languages or particular fields (patents, manuals, etc.) of machine translation.
- For the FP7, he is engaged in the EXPERT project that aims to train young researchers in the field of statistical machine translation technology.
- For the HORIZON2020, he is engaged in the QT21 project (DFKI is its coordinator). QT21 project aims to improve the quality of statistical machine translation, especially for the European languages. The quality of machine translation is very high between several languages, for example between English and French. But, in Europe, a lot of languages are used and the quality of translation is low between certain languages. In this project, Dr. Qun represents the researchers of the Dublin City University. Although its budget is a medium-sized scale, many organizations participate into this project.

#### ***ADAPT's achievement of the technology transfer***

- ICONIC and KantanMT.com are spin-off companies of CNGL.

- ICONIC provides a machine translation service in various fields, in particular, patent translation.
- KantanMT.com provides a cloud-based machine translation services for enterprises<sup>1</sup>.

#### *Differences between the United States and Europe in the machine translation field*

- In the United States, a government department, called DARPA, which promotes the development of military technologies, supports the research of machine translation technology. DARPA focus on the studies of machine translation between English and Arabic and between English and Chinese. In Europe, many European languages are studied for machine translation, and there are many types of application for industrial purposes. So, there are more demands for machine translation in Europe than in the United States. Based on these market conditions, ADAPT actively carries out the research and development of machine translation.
- For supplementary information, Franz Josef Och, the person responsible for the machine translation department of the US Google during 2004 and 2014 is from Europe. He conducted the research and obtained his PhD at RWTH Aachen University in Germany. He has left Google in 2014.
- Moses, the most popular open source SMT tool, is developed by Europe universities (lead by University of Edinburg).

### 第三章 ドイツ・カールスルーエ技術研究院インタラクティブシステム研究所 (KIT)

#### 組織概要及び研究トピック

- ドイツのカールスルーエ技術研究院インタラクティブシステム研究所（以下、KIT と略す）<sup>30</sup>では、音声及びテキスト機械翻訳、顔追跡（Face Tracking）、注意追跡（Attention Tracking）、マルチモーダル追跡、ニューラルネットワークの研究開発を行っている。
- KIT の所長は Alex Waibel 氏である。
- KIT は、米カーネギー・メロン大学のインタラクティブシステム研究所と提携して、音声認識システムである Janus Recognition Toolkit を開発している<sup>31</sup>。
- 2012 年に、KIT の Waibel 氏の研究チームは、大学講義向けの自動同時音声翻訳システムを発表している<sup>32</sup>。KIT 研究チームは、自動音声認識技術と統計的機械翻訳技術を組み合わせて、音声翻訳システムを開発している。このシステムは、大学の講義をリアルタイムで自動的に、記録、転写、翻訳する。外国人の学生は、講義後、自分の PC あるいはモバイル上で再び講義を母国語で聞くことができる。
- テキスト機械翻訳技術に関しては、特に欧州言語に焦点を当てて、研究を実施しており、世界的な機械翻訳評価ワークショップである WMT に数年来参加している<sup>33</sup>。また、テキスト機械翻訳技術の開発を、独仏 QUAERO プロジェクト、FP7 の EU-Bridge プロジェクト、ホライゾン 2020 の QT21 プロジェクトで進めている。
- KIT は多くの国際プロジェクト、EU プロジェクト、ドイツ国内のプロジェクトに参加している。
  - EU プロジェクト：QT21（ホライゾン 2020）、EU-Bridge（FP7）、EVEIL-3D プロジェクト（欧州地域開発基金）

<sup>30</sup> [http://isl.anthropomatik.kit.edu/english/21\\_74.php](http://isl.anthropomatik.kit.edu/english/21_74.php)

<sup>31</sup> <http://isl.anthropomatik.kit.edu/cmu-kit/english/1406.php>

<sup>32</sup> <http://isl.anthropomatik.kit.edu/english/1520.php>

<sup>33</sup> <http://www.statmt.org/wmt15/>

- 独仏プロジェクト:QUAERO
- 国内プロジェクト:CLICS

## 研究事例

### FP7 の EU-Bridge プロジェクト<sup>34</sup>

- ・ 音声転写及び音声翻訳サービスを開発する。TV 放送の字幕翻訳サービス、大学の講義音声翻訳サービス、欧州議会における同時通訳支援サービス（通訳に特殊な語彙などの訳語を事前に準備する）、ウェブ上のセミナーや講義の音声翻訳

略称	EU Bridge
正式名称	Bridges across the language divide
研究期間	2012 年 2 月 - 2015 年 1 月 (36 カ月)
全予算 (EU 拠出金)	1051 万ユーロ (787 万ユーロ)
プロジェクトコーディネーター	KIT (独)
参加組織	エジンバラ大学 (英)、アルカテル (仏)、RWTH アーヘン大学 (独)、レッド・ビーメディア (英)、香港科学技術大学 (中)、Pervoice (伊)、モバイル技術 (独)、ポーランド日本情報工科大学 (ポーランド)、ブルーノ・ケスラー基金 (伊)、ACCIPIO プロジェクト (独)
ウェブサイト	<a href="http://www.eu-bridge.eu">http://www.eu-bridge.eu</a>

### EVEIL-3D プロジェクト

- ・ EVEIL-3D プロジェクト (2012 年～2014 年) では、外国語の習得プログラムのために複数の仮想現実の技術を結合する。外国語習得者 (ドイツ語とフランス語) は、専用のゴーグルを装着し、3D の仮想現実空間の中で展開されるゲームにチャレンジする。このゲームの中で、外国語習得者は、外国語を話し、様々な状況に対応せねばならない。同プロジェクトには、ドイツとフランスの教育機関と研究機関が参加しており、KIT は、同プロジェクトに音声認識技術を提供している。

### QUAERO (クアエロ) プログラムにおける View4You プロジェクト

- ・ クアエロプログラムは、仏独共同研究開発プログラム (2008 年～2013 年) であり、特に多言語・マルチメディアコンテンツ自動分析に関わるアプリケーションを開発することを目的とする。特に、音声、文字、イメージ (写真等)、動画、音楽の自動処理アプリケーション (機械翻訳を含む) を開発している。仏独から全 32 組織が参加し、産学を提携させて研究を実施させた。
- ・ クアエロプログラムの助成の下、KIT は View4You プロジェクトを実施した。同プロジェクトで開発した View4You システムは、ドイツ語のテレビ放送報道番組を自動で録画し、放送された番組をニュース毎に分割し、音声によって検索することを可能にする<sup>35</sup>。

<sup>34</sup> <http://www.eu-bridge.eu/index.php>

## 第四章 ドイツ人工知能研究センター言語技術研究所 (DFKI)

### 組織概要及び研究トピック

- ・ ドイツ人工知能研究センター (DFKI) は、1988年に設立され、連邦政府や州政府等の公的機関と産業パートナーから助成されている。
- ・ 産業パートナーには、欧州企業のアエバス、BMW やドイツテレコム、SAP などの独企業その他、マイクロソフト、インテル等の米企業もいる (日本からはリコー社)。2015年10月には、グーグルが産業パートナーに加わっている。
- ・ DFKI は、ドイツ国内に4つの拠点を持ち、ソフトウェア技術の研究開発 (アプリケーション指向の基礎研究) を実施しており、特に人工知能分野の研究で有名である。
- ・ 機械翻訳を含む言語技術に関しては、言語技術研究所 (LT lab) に設置されたテキスト分析グループ (ベルリン市) とマルチ言語技術グループ (ザールブリュッケン市) で実施されている。
- ・ DFKI の言語技術研究所は、数多くの EU プロジェクトと国内プロジェクトに参加している。同研究所は、言語技術分野の欧州研究機関の提携を強化する連合体「META-NET」<sup>36</sup> のコーディネーターを務めており、欧州で同分野の中心的研究組織の一つであると言える。また、同研究所は、機械翻訳技術分野の欧州研究機関の提携を強化するホライゾン2020のCRACKERプロジェクトのコーディネーターを務めている。
- ・ DFKI からは60以上のスピンオフ企業が設立され、2500人以上の雇用を生んでいる。現在は、そのうち45の企業が活動している。機械翻訳及び音声技術に関しては、Yocoy<sup>37</sup> (機械翻訳及びトラベルガイド向けのモバイルアプリケーション開発)、CELI<sup>38</sup> (音声合成技術やテキスト分析技術の開発)、SemanticEdge<sup>39</sup> (スマートTV やスマートホーム、音声対話モバイルアプリケーション向けの音声認識、音声合成、自然言語理解技術の開発)、SEMVOX<sup>40</sup> (音声対話技術) がある。

### 研究プロジェクト事例

#### FP プロジェクト QTLEAP

略称	QTLEAP
正式タイトル	Quality Translation by Deep Language Engineering Approaches
全予算 (EU 拠出分)	3 972 243 ユーロ (3 003 000 ユーロ)
プロジェクト期間	2013年11月～2016年10月 (36ヶ月)
コーディネーター	リスボン大学 (ポルトガル)

<sup>35</sup> View4You プロジェクトの成果は、インターネット上でビデオで紹介されている。

<sup>36</sup> [http://www.dailymotion.com/video/xwngiy\\_kit-view4you\\_tech](http://www.dailymotion.com/video/xwngiy_kit-view4you_tech)

<sup>37</sup> [http://www.dfki.de/lt/project.php?id=Project\\_599&l=en](http://www.dfki.de/lt/project.php?id=Project_599&l=en)

<sup>38</sup> <http://www.yocoy.com/en/home.html>

<sup>39</sup> <http://www.celi.it/en/sophia-semantic-engine.shtml>

<sup>40</sup> [https://www.semanticedge.de/page\\_en/index.html](https://www.semanticedge.de/page_en/index.html)

<sup>40</sup> <http://www.semvox.de/en/start.html>

参加組織数	情報通信技術研究院（ブルガリア）、プラハ・カレル大学（チェコ）、DFKI（独）、フンボルト大学ベルリン（独）、バスク大学（スペイン）、フローニンゲン大学（蘭）、HIGHER FUNCTIONS（ポルトガル）
ウェブサイト	<a href="http://qtLeap.eu">http://qtLeap.eu</a>
研究内容	自動翻訳技術のために、ディープ・アンダースタンディング（Deep understanding）の研究を実施している。

#### **HORIZON2020 CRACKER プロジェクト**

略称	Cracker
正式タイトル	Cracking the Language Barrier : Coordination, Evaluation and Resources for European MT Research
全予算（EU 拠出分）	999 995 ユーロ（999 995 ユーロ）
プロジェクト期間	2015年1月～2018年1月（36ヶ月）
コーディネーター	DFKI（独）
参加組織数	チャールズ大学プラハ（チェコ）、評価・言語資源分配庁（仏）、ブルーノ・ケスラー基金（伊）、アテナ情報通信・ナレッジ技術研究・イノベーションセンター（ギリシア）、エジンバラ大学（英）、シーフィールド大学（英）
ウェブサイト	<a href="http://www.meta-net.eu/projects/cracker">http://www.meta-net.eu/projects/cracker</a>
活動内容	欧州の研究機関の提携を強化するとともに、FP7においてすでに実施されたプロジェクトを延長して支援する。

#### **HORIZON2020 プロジェクト QT21**

略称	QT21
正式タイトル	Quality Translation 21
全予算（EU 拠出分）	3 997 428 ユーロ（3 997 428 ユーロ）
プロジェクト期間	2015年2月～2018年1月（36ヶ月）
コーディネーター	DFKI（独）

参加組織数	RWTHアーンヘン大学（独）、アムステルダム大学（蘭）、ダブリンシティ大学（アイルランド）、エジンバラ大学（英）、KIT（独）、CNRS（仏）、プラハ・カレル大学（チェコ）、ブルーノ・ケスラー基金（伊）、シーフィールド大学（英）、TAUS（蘭）、TEXT & FORM（独）、TILDE（ラトビア）、香港科学技術大学（中）
ウェブサイト	<a href="http://www.gt21.eu">http://www.gt21.eu</a>
研究内容	欧州言語を対象に機械翻訳技術の質を向上させる（すでに機械翻訳の品質がある程度高い英語や仏語などのペアではなく、特に利用者が少ない欧州言語が対象）。

### ヒアリング調査メモ / DFKI 言語技術研究所テキスト分析グループ

- ・ 日程：平成 27 年 12 月 15 日（火）午前 11 時半～
- ・ 場所：NICT 欧州連携センター事務所（スカイプを利用したインタビュー）
- ・ 先方：DFKI 言語技術研究所テキスト分析グループ<sup>41</sup> 上席研究員 ゲオルグ・ラン氏<sup>42</sup>

### ヒアリングの概要

#### DFKI 言語技術研究所の概要について

- ・ DFKI（ドイツ人工知能研究所）は 1988 年に設立された。
- ・ 言語技術の研究開発は、DFKI 言語技術研究所に設置されたテキスト分析グループ（ベルリン）とマルチ言語技術グループ（ザールブリュッケン）で実施されている。テキスト分析グループは、機械翻訳を含む言語技術に関する様々な研究（テキスト分析、情報抽出など）を実施するとともに、欧州の研究開発機関を提携させ、EU 機関などにロビー活動をしている。マルチ言語技術グループでは、マルチ言語と機械翻訳技術の研究を実施している。
- ・ DFKI は、欧州における言語技術研究分野の研究機関の集合体である META-NET<sup>43</sup>のコーディネーターであり、様々な研究機関とつながりがある。

#### ゲオルグ・ラン氏の研究関心について

- ・ ラン氏は、2010 年から DFKI に勤務している。人工知能の研究とコンピュータ科学が専門であり、現在、特に計算言語学に関心を持っている。
- ・ 現在、ラン氏はデジタルキュレーション（デジタル資産の管理など）の研究プロジェクトを実施している。この分野は将来的に発展していく分野である。

<sup>41</sup> <http://www.dfki.de/it/>

<sup>42</sup> ゲオルグ・ラン氏については以下のウェブサイトを参考のこと。  
<http://www.dfki.de/~grehm/>

<sup>43</sup> META-NET については、メモ後半部でより詳しく記す。

#### DFKI 言語技術研究所の研究活動について

- DFKI 言語技術研究所（以下、DFKI とする）では、機械翻訳技術に関しては、テキスト機械翻訳技術の研究開発を実施している。過去に、音声機械翻訳（Speech to Speech）の研究開発も実施していたが、現在は行っていない。
- DFKI は、機械翻訳の品質を飛躍的に向上させ、次世代レベルの翻訳品質を実現することを研究課題としている。
- DFKI は、翻訳企業など、多くの言語技術サービスプロバイダと提携している。例えば、翻訳者のワークフローを改善する方法の開発を実施している。
- 機械翻訳技術の開発に関しては、一般的な機械翻訳システムと各ドメインに合わせた機械翻訳システムの双方を開発している。
- デジタルキュレーションは、ユーザが情報を探すために機械翻訳技術を必要とするが、その際、一般的な機械翻訳システムを開発している。
- 特殊なドメインの機械翻訳技術に関して、FP7 の QTLeap プロジェクト<sup>44</sup>は、機械翻訳のユーザ事例として、一般企業のオンライン顧客向けサービスの機械翻訳システムを開発している。他には、EU の CEF（Connecting European Facility）<sup>45</sup>というプログラムでは、欧州のインフラストラクチャの接続が支援されているが、公共機関向けの「CEFAT（CEF Automated Translation）」<sup>46</sup>という機械翻訳プラットフォームも開発しており、DFKI はこのプラットフォームの開発に関わっている。CEFAT は、現在 EU 諸機関と EU 加盟国の公的機関向けに利用されている「MT@EC」の上に構築され、EU の 24 カ国語及びノルウェー語、アイスランド語の全てのペア（両方向）の翻訳を対象とする。
- DFKI では、対訳コーパスを収集するよりもむしろ、機械翻訳の品質を向上させることに焦点を置いている。

#### DFKI 言語技術研究所のスピンオフ企業について

- DFKI 言語技術研究所の機械翻訳技術関連のスピンオフ企業には、Acrolinx<sup>47</sup>と Yocoy<sup>48</sup>がある<sup>49</sup>。
- 現在、ドイツではシリアからの難民を多く受け入れているので、ドイツ人と難民のコミュニケーションを支援するために、DFKI のスピンオフ企業である Yocoy<sup>50</sup>に、テンプレートベースの機械翻訳アプリケーションを開発する注文があった。

#### FP7 における機械翻訳技術の研究プロジェクトについて

- DFKI は、機械翻訳システムの誤りを修正するために、人間の翻訳者と共同して研究している。翻訳者は機械翻訳システムがどのような誤りをしたか指摘する。DFKI は、同研究を FP7 の QT LAUNCH PAD プロジェクト<sup>51</sup>で実施し、MQM（Multidimensional Quality Metrics）システム<sup>52</sup>を開発した（QT LAUNCH PAD プロジェクト自体はより幅広い研究

<sup>44</sup> <http://qt leap.eu>

<sup>45</sup> <http://ec.europa.eu/digital-agenda/en/connecting-europe-facility>

<sup>46</sup> [https://joinup.ec.europa.eu/community/cef/og\\_page/catalogue-building-blocks](https://joinup.ec.europa.eu/community/cef/og_page/catalogue-building-blocks)

[http://www.rigasummit2015.eu/sites/rigasummit2015.eu/files/cef\\_29\\_04\\_2015\\_spyridon\\_pilos\\_building\\_the\\_automated\\_translation\\_platform.pdf](http://www.rigasummit2015.eu/sites/rigasummit2015.eu/files/cef_29_04_2015_spyridon_pilos_building_the_automated_translation_platform.pdf)

<sup>47</sup> <http://www.acrolinx.de>

<sup>48</sup> <http://www.yocoy.com/en/home.html>

<sup>49</sup> DFKI 言語技術研究所のスピンオフ企業のリストがウェブサイト上で公表されている。

<http://www.dfki.de/web/ueber/spin-offs>

<sup>50</sup> <http://www.yocoy.com/en/home.html>

<sup>51</sup> <http://www.qt21.eu/launchpad/>

<sup>52</sup> [http://www.dfki.de/lt/publication\\_show.php?id=7717](http://www.dfki.de/lt/publication_show.php?id=7717)

で、MQMは同プロジェクトの成果の一つである)。MQMは、システムというよりは、むしろアノテーションスキームであり、単に意味の違いを指摘することに留まらず、機械翻訳システムがおかした誤りにアノテーションを与えていく。

- ・ FP7のQTLeapプロジェクトでは、ディープ・アンダースタンディング(Deep understanding)の研究を実施している。

#### ホライゾン 2020における機械翻訳技術の研究プロジェクトについて

- ・ 現在、ホライゾン 2020で助成支援されている機械翻訳分野のプロジェクトは合計で6つある。
- ・ DFKI言語技術研究所は、ホライゾン 2020の2つのプロジェクトに参加している(QT21プロジェクトとCRACKERプロジェクト)。
- ・ QT21プロジェクトは、技術開発を行うプロジェクトであり、CRACKERプロジェクトは、機械翻訳の研究を実施する機関の提携などを支援するプロジェクトである。
- ・ 残りの4つのプロジェクトは、LT-Observatory、HimL、MMT、Tramoocである。LT-Observatoryは、言語リソース(加盟国の公共機関の対訳コーパス)の保管場所を構築するなどしており、技術開発そのものは行っていない。HimL、MMT、Tramoocプロジェクトは、機械翻訳技術の革新的な開発を目的とする。

#### ホライゾン 2020のQT21プロジェクトについて

- ・ ザールブリュッケンに設置されたDFKI言語技術研究所マルチ言語技術グループは、QT21プロジェクトを実施している(同研究グループのジョセフ・ヴァン=ゲネップ氏がプロジェクトリーダー)<sup>53</sup>。QT21プロジェクトは、欧州で最も優秀な研究機関が集まり、研究を実施している。現在、同プロジェクトでは、MOSESの開発を進め、欧州言語の翻訳の品質を向上させる研究を実施している。
- ・ 同プロジェクトでは、言語技術サービスプロバイダも参加し、人間の翻訳者が機械翻訳技術の開発に加わっており、非常に独自である。

#### ホライゾン 2020のCRACKERプロジェクトについて

- ・ DFKI言語技術研究所テキスト分析グループが実施しているCRACKERプロジェクト(ゲオルグ・ラン氏がコーディネーター)は、技術革新を狙うプロジェクトではなく、機械翻訳の研究を実施する機関の提携、研究者のネットワーク作り、過去のFP7プロジェクトの成果の延長などを支援している。
- ・ CRACKERプロジェクトは、FP7のQT LAUNCH PADプロジェクトで開発した機械翻訳技術開発のエコシステム(翻訳プラットフォーム)の改善を支援している。このエコシステムは、異なる機械翻訳エンジンの翻訳文を比較し、評価し、分析する一種のワークベンチである(一つのソフトウェアで、複数の機械翻訳エンジンの利用、翻訳された文章(アウトプット)の閲覧、アウトプットへのアノテーション、人間の翻訳者の翻訳の閲覧が可能である。またMQMアプローチも利用される)。
- ・ Crackerプロジェクトは、META-NETを支援している。META-NETは、2010年に設立され、30カ国60の言語技術分野の研究開発組織の集合体(民間企業は少ない)である。META-NETに対するEUからの助成は2013年に終えている。DFKIは、META-NETのコーディネーターをしている。

---

<http://www.qt21.eu/mqm-definition/definition-2015-06-16.html>

<sup>53</sup> [http://www.dfki.de/lt/project.php?id=Project\\_888&l=en](http://www.dfki.de/lt/project.php?id=Project_888&l=en)

- ・ CRACKER プロジェクトの支援のもと、META-NET は、2015 年、2016 年、2017 年にカンファレンスを開催する。2015 年は、4 月末にラトビアのリガでカンファレンスを開催した。2016 年は7月にポルトガルのリスボンで、2017 年はハンガリーのブダペストで開催する予定である。
- ・ CRACKER プロジェクトは、現在、欧州委員会、欧州議会にも働きかけている（ロビー活動）。なぜなら、FP6、FP7 では、機械翻訳技術を含む言語技術（Language technology）は一つの研究トピックとして認められていたが、ホライゾン 2020 では、言語技術という研究トピックがなくなってしまうからである<sup>54</sup>。EU には、24ヶ国語の公用語があり、機械翻訳技術が非常に重要であることを考えれば、このように助成枠がなくなることはおかしいと言わざるをえない。EU 機関では、ビッグデータやセキュリティなどを重視する傾向がある。
- ・ META-NET は、言語技術による各欧州言語のサポート状況を報告するために、2012 年に META-NET 白書<sup>55</sup>（全32巻 / 31か国語）を発表している（ラン氏は共同編集者のひとり）。白書によれば、例えば英語に関しては、言語技術の研究開発が非常によく進められているが、他の言語、例えば、使用者数が少ないラトビア語などに関しては、言語技術はあまり開発が進められていない。META-NET の目的は、すべての欧州言語を対象にして言語技術を開発することである。また、白書の報告では、欧州の21か国がデジタル化から取り残される危険がある<sup>56</sup>。使用者が少ない言語を母語とする人は、インターネット上でサービスを利用するときに他の言語（例えば、英語）を利用しなければならないことが多く、インターネット上で母語を利用していない。そして、そのまま他の言語を利用し続けていけば、母語のデジタル化の必要性も最終的になくなっていくだろう。META-NET はこのような社会状況に対して戦っている。その上、現在、欧州では、国境、言語の壁を超えたオンラインマーケットを形成することを目標とする「マルチ言語デジタル単一市場」<sup>57</sup>というプログラムが策定されており、そのような市場が作られれば欧州にとって非常に価値がある<sup>58</sup>。だが、機械翻訳技術によるサポートがなくては、欧州内の24の市場が形成されてしまう。このような問題意識を共有するプロジェクトとしては、ホライゾン 2020 の LT-Observatory<sup>59</sup>がある。
- ・ META-NET を含む欧州の言語技術に係るグループや EU プロジェクトに関しては、「Cracking the language barrier」というウェブサイト<sup>60</sup>から情報を得ることができる。
- ・ META-NET は、META-SHARE<sup>61</sup>というオープン言語リソース共有システムの開発に助成していた（2010 年から 2013 年まで）。現在は、CRACKER が META-SHARE に助成している。

<sup>54</sup> ホライゾン 2020 の 2013-2014 年度 ICT 作業プログラムでは、「ICT-17: 言葉の壁の打破」という枠組みで、機械翻訳技術の研究開発が助成されるものの、最新の 2016-2017 年度 ICT 作業プログラムでは、「ICT-19: メディアとコンテンツの融合」という枠組みで、転写技術や自動翻訳技術が、メディアとコンテンツにアクセスすることを可能にする技術として助成されることにとどまる（ICT-19 の予算は全体で 3900 万ユーロであり、自動翻訳と音声技術以外の技術への助成も含む）。

[http://ec.europa.eu/research/participants/data/ref/h2020/wp/2016\\_2017/main/h2020-wp1617-leit-ict\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/wp/2016_2017/main/h2020-wp1617-leit-ict_en.pdf)

<sup>55</sup> <http://www.meta-net.eu/whitepapers/overview>

<sup>56</sup> <http://www.meta-net.eu/whitepapers/press-release>

<sup>57</sup> <http://rigasummit2015.eu>

<sup>58</sup> <http://www.meta-net.eu/projects/cracker/multimedia/mdsm-fact-sheet.pdf>

<sup>59</sup> <http://www.lt-observatory.eu/partners.html>

<sup>60</sup> <http://www.cracking-the-language-barrier.eu>

<sup>61</sup> <http://metashare.elda.org>

## 第五章 フランス・メヌ大学情報学研究所 (LIUM)

### 組織概要及び研究トピック

- ・ フランスのル・マン市に設立されたメヌ大学情報学研究所 (LIUM)<sup>62</sup>では、音声言語翻訳、ロバスト音声認識、話者ダイアライゼーション、話者認識、話者言語理解の研究が実施されており、これら全てのトピックについて、特に深層ニューラルネットワークの利用について研究している。

### スピンオフ企業

- ・ LIUMから、スピンオフ企業が2社設立されている。Deeplingo社<sup>63</sup>はパリに本拠地を持ち、企業向けに性能のよい機械翻訳のモデルを提供している。Voxolab社<sup>64</sup>はルマン市に本拠地を持ち、音声認識技術に基づく、性別等の話者特定アプリケーションや転写アプリケーションを企業向けに提供する。

### ヒアリング調査メモ / 仏 LIUM

- ・ 日程：平成27年9月28日(月)午後2時～
- ・ 場所：先方事務所 (フランス・ルマン市)
- ・ 先方：
  - メヌ大学コンピュータ科学部教授・クロード・シャップ研究院所長及びLIUM所長<sup>65</sup>：ヤニック・エステーヴ氏<sup>66</sup>
  - LIUM研究者3名：アントニ・ラルシェ氏 (音声認識・話者ダイアライゼーション分野)、ロイック・バロー氏<sup>67</sup> (音声機械翻訳)、フェティ・ブガレス氏<sup>68</sup> (音声機械翻訳)

### ヒアリング概要

#### LIUMの概要について

- ・ LIUM (Laboratoire d'Informatique de l'Université du Maine：メヌ大学情報学研究所)は1980年に設立された。メヌ大学においては、コンピュータ科学の研究及び教育活動は1975年に開始されており、40年間で同分野の活動は拡大してきている。現在、同大学コンピュータ科学の学生数は約600名である。
- ・ LIUMの人員数は、非常勤のスタッフを合わせて50名である。
- ・ メヌ大学から支給されている年間予算は5万ユーロである。その他の予算は、産業や研究助成組織など(欧州、フランス政府、仏ペイ・ド・ロワール地域圏など)との研究契約から得ている。2013年の全予算は83万4000ユーロ、2014年の全予算は136万5000ユーロであった。年間予算は締結した研究契約によって変化する。

<sup>62</sup> <http://www-lium.univ-lemans.fr/en/content/welcome>

<sup>63</sup> <http://www.deeplingo.com/about>

<sup>64</sup> <http://voxolab.com>

<sup>65</sup> <http://www-lium.univ-lemans.fr/en/content/welcome>

<sup>66</sup> <http://www-lium.univ-lemans.fr/~yannick/>

<sup>67</sup> <http://www-lium.univ-lemans.fr/~barraut/home.php>

<sup>68</sup> <http://www-lium.univ-lemans.fr/~bougares/accueil.php>

- 2015年9月25日、メーヌ大学で、ルマン市があるサルト県の複数の研究開発組織を連合させるクロード・シャップ研究院の開所式が行われた。この研究院は研究機関の役割をなすだけでなく、全ての教育プログラムを相互調整して、教育・研究活動をより目に見えるかたちにするとともに、より魅力的にする役割を持つ。つまり、同研究院は、研究、教育、産業の交流を強化し、スピノフの起業を促進し、地域経済政策を実現するためのツールである。これはルマン市をデジタル最先端技術都市として発展させる地域経済振興政策に沿い、地元産業、ひいては、フランス産業の復興を目的とする。ルマン市は、パリから近い上に（TGVで1時間程）、家賃なども安いので、企業を招致できる環境にある。こうして、LIUMの研究開発は地域経済に貢献する。開所式には、サルト県出身のステファヌ・ル・フォル農業・農産加工業・林業大臣兼政府報道官及びアクセル・ルメール経済・産業・デジタル大臣付きデジタル担当大臣が参加した。
- LIUMはクロード・シャップ研究院の中核にあり、それゆえ、LIUM所長が同研究院の所長も務めている。

#### LIUMが参加する研究プロジェクト及び国際パートナーについて

- 過去4年間で5つの欧州プロジェクトに参加している。FP7のEuromatrixPlus<sup>69</sup>（機械翻訳）、FP7のMateCat<sup>70</sup>（機械翻訳）、FP7のEUMSSI<sup>71</sup>（スピーチ及び話者認識）、CHIST-ERA<sup>72</sup>のJOKER<sup>73</sup>（音声言語理解）、CHIST-ERAのM2CR<sup>74</sup>（深層ニューラルネットワークを利用した音声・画像理解）である。これらのプロジェクトにおいて、独DFKI、伊FBK、チェコ・チャールズ大学、チェコ・CEET、アイルランド・ダブリンシティ大学、英エジンバラ大学、米ジョン・ホプキンス大学、スペイン・ブンペウファブラ大学、独L3S、スイス・IDIAP、独ドイツ・ヴェレ、独GFKAIなどと提携している。
- カナダのCRIM（モントリオール・コンピュータ研究院）と定期的に提携し、また米カーネギー・メロン大学のスタッフと本を一冊公刊している。
- 米軍事向け研究開発助成スキームであるDARPAの枠組みで、米IBM社主導のプロジェクトに参加している（GALEプログラム：2009年～2011年及びBOLTプログラム：2011年～2016年の枠組み）。
- 機械翻訳と音声処理に係る国内プロジェクトに参加している<sup>75</sup>。ASH<sup>76</sup>（プロジェクトリーダー：研究期間2009年～2012年：研究内容転写システム）、VERA<sup>77</sup>（プロジェクトリーダー：2013年～2015年：音声認識）、Cosmat（2009年～2012年：学術論文の機械翻訳）、PACTE（2011年～2015年：光学文字認識（OCR）の機械翻訳の訂正）、PORT-MEDIA<sup>78</sup>（2009年～2012年：マルチ音声言語理解システム）、TRAD（軍事向け機械翻訳ドメイン適応）、TRIAGE（2013年～2015年：災害被害者のトレーサビリティ・確認・特定・管理）、DEPART（2009年～2012年：マルチモダル・マルチ言語の文書処理）である。

<sup>69</sup> <http://www.euromatrixplus.net>

<sup>70</sup> <https://www.matecat.com>

<sup>71</sup> <http://www.eumssi.eu>

<sup>72</sup> CHIST-ERAは、FP7とは異なり、EUの支援のもと、欧州諸国とカナダのケベック州の研究助成機関を提携させ、研究開発を助成するスキームである。

<sup>73</sup> <http://www.chistera.eu/projects/joker>

<sup>74</sup> <http://www.chistera.eu/projects/m2cr>

<sup>75</sup> LIUMが参加する研究プロジェクトのリスト：<http://www-lium.univ-lemans.fr/en/content/projects-0>

オルガー・シュリンク元教授のウェブサイトも参考のこと。[http://www-lium.univ-lemans.fr/~schwensk/projects\\_fr.html](http://www-lium.univ-lemans.fr/~schwensk/projects_fr.html)

<sup>76</sup> <http://projet-ash.univ-lemans.fr/doku.php>

<sup>77</sup> <http://projet-vera.univ-lemans.fr/doku.php?id=home>

<sup>78</sup> <http://www.port-media.org/doku.php?id=start>

- ・ 音声言語機械翻訳の評価キャンペーンである IWSLT (International Workshop on Spoken Language Translation)、ASRU<sup>79</sup>、米 NIST が主催する自動音声認識評価キャンペーンなどに積極的に参加し、優秀な成績を収めている。音声認識と話者ダイアライゼーションの評価キャンペーンで、シンガポール・科学・技術・研究庁 (ASTAR)<sup>80</sup>、フィンランド・東フィランド大学、伊コレ大学、シンガポール・南洋理工大学<sup>81</sup>とコンソーシアムを形成している。また、米ジョージア工科大学と交流がある。

#### LIUM 全般の研究活動について

- ・ 研究トピックは、音声言語翻訳、ロバスト音声認識、話者ダイアライゼーション、話者認識、話者言語理解である。これら全てのトピックについて、深層ニューラルネットワークの利用に関心を持っている。
- ・ LIUM では、基礎研究と応用研究の双方を実施しているが、応用研究のほうにより力を入れている。

#### 機械翻訳技術の研究活動について

- ・ **機械翻訳の基本システム**: LIUM では、英エジンバラ大学が開発した MOSES を基本システムとして利用している。MOSES は欧州、そして、世界中で非常によく利用されている。ちなみに、MOSES の開発者であるフィリップ・コーエン氏は、エジンバラ大学から米ジョン・ホプキンス大学に移動している。
- ・ **研究開発の方針**: 音声機械翻訳の精度だけでなく、計算時間の速度を可能なかぎり増大させることにも力を入れており、数年前と比べて、格段に加速させた。昨年、LIUM は IWSLT に参加したが、機械翻訳の精度を評価の規準にしており、翻訳の速度は評価していない(昨年の IWSLT では、NICT が最上位に入っていたことを覚えている)。
- ・ **1) 研究トピック**: LIUM の統計機械翻訳の研究開発グループでは、ニューラルネットワークを利用した機械翻訳システムの研究を行っている。この研究は、LIUM で機械翻訳技術と深層学習 (Deep Learning) の研究開発を行っていたオルガー・シュリンク元教授により主導されていた。同教授<sup>82</sup>は、2015 年に退職し、同年 6 月にパリに創設された米フェイスブックの人工知能研究センターに移っている<sup>83</sup>。LIUM では、同氏が退職した後も同分野の研究開発は続けられている。LIUM は深層学習向けのライブラリとして有名な THEANO を利用している。
- ・ **2) 研究トピック**: 機械翻訳のドメイン適応に関心を持ち、機械翻訳システムを異なるリソース (報道記事、音声会話など) とテキストの異なるスタイル (ウェブフォーラム、ブログ、チャット、ショートメッセージなど) に適応させる研究を行っている。
- ・ **3) 研究トピック**: リソースが少ない言語 (low resource language) の機械翻訳の研究を行っている。ニューラルネットワークは少しずつ効率性を増加させているが、リソースが少ない言語に関してはそうではないので、この研究は段々と重要性が増している。リソースが少ない言語とは、例えば、アフガニスタンのパシュトゥーン語などである。だが、ドメインによってはリソースが少ない言語もある。例えば、英語とアラビア語の機械翻訳のリソースは一般に豊富にあるとしても、両言語の医療分野のリソースは少ないので、このケ

<sup>79</sup> <http://www.asru2015.org>

<sup>80</sup> <http://www.a-star.edu.sg>

<sup>81</sup> <http://www.ntu.edu.sg/Pages/home.aspx>

<sup>82</sup> <http://www-lium.univ-lemans.fr/~schwenk/>

<sup>83</sup>

[http://www.lemonde.fr/economie/article/2015/06/02/facebook-ouvre-un-laboratoire-d-intelligence-artificielle-a-paris\\_4645842\\_3234.html](http://www.lemonde.fr/economie/article/2015/06/02/facebook-ouvre-un-laboratoire-d-intelligence-artificielle-a-paris_4645842_3234.html)

ースはリソースが少ない言語と見なしうる。だが、パシュトゥーン語などは全般にリソースが少ないので、他の言語よりも技術が必要とされる。

- ・ **研究開発物**: オルガー・シュリンク元教授が開発した CSLM ツールキット (Continuous Space Language Model)<sup>84</sup> はオープンソース (GPL ライセンス) で提供されている。CSLM ツールキットは、主に言語モデリングのために利用されるが、Continuous Space MT のために拡張して利用している (Continuous Space はニューラルネットワークを意味する)。
- ・ **機械翻訳の対象言語**: 機械翻訳に関しては、仏語→英語、英語→仏語・独語、アラビア語 (現代標準アラビア語と方言) →英語・仏語、北京語→英語の組み合わせを開発対象としている。アラビア語等と比べて、仏語と英語の組み合わせへの関心は低くなってきているが、LIUM の音声機械翻訳は非常に高い成果を出している。
- ・ **機械翻訳の対象分野**: 主に、報道分野の機械翻訳に焦点を当てている。学术论文、特許翻訳を対象にしたことがあるが、説明書、金融、医療分野に関しては、対象にしたことがない。学术论文を機械翻訳の開発対象にした理由は、ドメイン適応を研究するにあたって、よいタスクであったからである (実験を行うのに、データの収集が容易であった)。
- ・ **対訳コーパスの収集**: LIUM 自身では、特定のプロジェクト以外では、対訳コーパスを収集していない。多くのデータが無料で手に入るからである (欧州議会、TED コーパスなど)。LIUM は多くのプロジェクト、また評価キャンペーンに参加しているので、そこで多くのコーパスが提供される。話者認識の研究分野では、コーパスの収集について、個人情報収集が問題になるが (音声データを、個人を特定可能な情報と一緒に収集するのは難しい)、音声機械翻訳の研究分野ではその問題はない。

#### 機械翻訳分野の研究プロジェクトについて

- ・ **EU プロジェクト**: EU の CHIST-ERA の枠組みで、カナダの MILA (モンリオール学習アルゴリズム研究院) とスペインの CVC (コンピュータビジョンセンター) と提携し、M2CR プロジェクト<sup>85</sup> (2015 年～2017 年) を開始したところである。同プロジェクトでは、深層ニューラルネットワークを利用した音声・画像理解、音声機械翻訳の研究を行う。
- ・ **EU プロジェクト**: FP7 の MateCat<sup>86</sup> プロジェクト (2011 年～2014 年) は、コンピュータ翻訳支援ツール (CAT) の開発を行っている。CAT は翻訳者による後編集を支援するツールでもあり、翻訳の速度と質を向上させ、翻訳作業の生産性を改善する。
- ・ **米 DARPA プロジェクト**: IBM が主導する米軍事向け研究開発助成スキームの DARPA (GALE と BOLT<sup>87</sup> の枠組み) のプロジェクトに参加し、機械翻訳システムをドメイン適応の開発を行っている。アラビア語等の特定の言語の音声会話、チャット、SMS、ブログが対象となるデータである。
- ・ **仏国内プロジェクト**: PACTE プロジェクト (2011 年～2015 年) では、光学文字認識 (OCR) を利用する機械翻訳の訂正の研究を行っている。
- ・ **仏国内プロジェクト**: COSMAT プロジェクト (2009 年～2012 年) では、学术论文を対象とする機械翻訳の研究を実施した。
- ・ **仏国内プロジェクト**: ペイ・ド・ロワール地域圏が支援するプロジェクトである DEPART プロジェクト (2009 年～2013 年) では、マルチモデル・マルチ言語の文書処理を対象とする (例: 音声認識と手書きのテキスト)。このプロジェクトでは、フランスのナント市

<sup>84</sup> <http://www-lium.univ-lemans.fr/~cslm/>

<sup>85</sup> <http://www.chistera.eu/projects/m2cr>

<sup>86</sup> <https://www.matecat.com>

<sup>87</sup> <http://www.darpa.mil/program/broad-operational-language-translation>

にあるナント・アトランティック情報学研究所 (LINA)<sup>88</sup>とナント・通信サイバネティクス研究院 (IRCCyN) と提携している。

#### 音声言語処理分野の研究活動について

- ・ 音声認識の基本システムには、Kaldi を利用している。
- ・ 音声認識の対象言語：仏語、英語、スペイン語、独語、イタリア語、現代標準アラビア語、チュニジア方言、パシュトゥーン語（アフガニスタン国の言語）を開発対象としている。
- ・ 1) 研究開発ソフトウェア：XenC<sup>89</sup>は、機械翻訳と音声認識におけるデータ選択のためのオープンソースツールキットで、大量のデータから特定のデータを抽出することを可能にする。
- ・ 2) 研究開発ソフトウェア：SIDEKIT<sup>90</sup>は、話者・言語認識のオープンソースツールキットである。
- ・ 3) 研究開発ソフトウェア：MANY<sup>91</sup>は、異なる機械翻訳のシステムを組み合わせるソフトウェアである。
- ・ 4) 研究開発ソフトウェア：LIUM\_SpkDiarization<sup>92</sup>は、話者ダイアライゼーション向けのソフトウェアで、同分野で非常に有名である。
- ・ 5) 研究開発ソフトウェア：YADTK<sup>93</sup>は、音声対話システムである。
- ・ 1) その他の研究開発物：フランス語の音声認識向けの聴覚言語モデル<sup>94</sup>を開発した。
- ・ 2) その他の研究開発物：TED-LIUM コーパス<sup>95</sup>、EPAC コーパス<sup>96</sup>を開発した。TED-LIUM コーパスは、TED ウェブサイトで手に入る聴覚情報とその転写情報から開発され、LIUM は IWSLT (2011 年) に参加するためにこの聴覚モデルをトレーニングした。EPAC コーパスは、フランス語のブロードキャスト放送のコーパスである。
- ・ 3) その他の研究開発物：パシュトゥーン語の音声辞書を開発した。
- ・ 4) その他の研究開発物：自動転写システムを開発した。

#### 話者ダイアライゼーションの研究について

- ・ 話者ダイアライゼーション (Speaker diarization) のツールキットである LIUM\_SpkDiarization を、オープンソースで提供している (GPL ライセンス)。同ツールキットは、LIUM のシルヴァン・ムニエが開発した<sup>97</sup>。同氏は、10 年以上前から話者ダイアライゼーションの研究を実施している。当初は mClust<sup>98</sup>というソフトウェアを開発していたが、米カーネギー・メロン大学の研究者と提携を開始してから、ソフトウェアを Java で書き換えることにした。それが LIUM\_SpkDiarization である。現在、このソフトウェアは世界中で普及している。

<sup>88</sup> <https://www.lina.univ-nantes.fr>

<sup>89</sup> <http://www-lium.univ-lemans.fr/fr/node/8346>

<sup>90</sup> <http://www-lium.univ-lemans.fr/sidekit/index.html>

<sup>91</sup> <http://www-lium.univ-lemans.fr/fr/content/many>

<sup>92</sup> <http://www-lium.univ-lemans.fr/fr/content/liumspkdiazation>

<sup>93</sup> <http://www-lium.univ-lemans.fr/fr/content/dialogue>

<sup>94</sup> <http://www-lium.univ-lemans.fr/fr/content/modeles-acoustic>

<sup>95</sup> <http://www-lium.univ-lemans.fr/fr/content/corpus-ted-lium>

<sup>96</sup> <http://projet-epac.univ-lemans.fr/doku.php?id=corpus>

<sup>97</sup> [http://catalog.elra.info/product\\_info.php?products\\_id=1119](http://catalog.elra.info/product_info.php?products_id=1119)

<sup>97</sup> <http://www-lium.univ-lemans.fr/diarization/doku.php/contact>

<sup>97</sup> <http://www-lium.univ-lemans.fr/~meignier/home.php>

<sup>98</sup> <http://www-lium.univ-lemans.fr/fr/content/mclust>

- ・ LIUM\_SpkDiarization は、音声データを分割し、ある話者がそのデータのどこで話しているのかを特定するのに利用されている。これにより、音声データにおいて、誰がどのようなことを話したのかを最終的に特定できるようになる。
- ・ 話者ダイアライゼーションシステムには、深層学習のアルゴリズムを含め、音声認識、音声機械翻訳、話者認識、自動転写など、多くの技術を統合する必要がある。LIUM は各音声技術を開発するだけでなく、それらを調整し、融合する研究を行っており、音声技術の包括的なシステムの開発を目指している。LIUM は、ニューラルネットワークを利用することによって、様々な技術の統合がより簡単になり、音声認識、音声翻訳、音声合成が一つのプロセスで可能になるのではないかと考えており、このような研究を中期的な研究目標としている。以上のため、LIUM は他の研究機関との提携に積極的である。
- ・ このシステムにより、例えば、ビッグデータをもとに、ある話者の発言の索引 (index) をつくるアプリケーションを構想している。ある話者がさまざまな音声データの中でどのような発言をしたのかを収集する。
- ・ ニューラルネットワークは人工知能の一種で、それ自体は非常に昔からあるが、10 年程に再び注目を集め、特に 5 年ぐらい前から、例えば GMM (Gaussian Mixture Model) に取って代わり、現在非常に支配的である。今年、ドイツで開催された音声技術の国際カンファレンスでは、音声処理分野のプレゼンテーションの半分が深層ニューラルネットワークを利用するものであった。

#### 技術移転活動について

- ・ 欧エアバスディフェンス&スペース (TRAD プロジェクト)、米 IBM (米 DARPA の GALE と BOLT の枠組み) に技術移転を行ったことがある。
- ・ 近年、LIUM から、DeepLingo 社 (機械翻訳) と Voxolab 社 (音声処理) という二つのスピンオフ企業が生まれている。
- ・ Deeplingo 社<sup>99</sup>はパリに本拠地を持ち、企業向けに性能のよい機械翻訳のモデルを提供している。同社には LIUM のオルガー・シュリンク元教授が関わっている。
- ・ Voxolab 社<sup>100</sup>は 1 年前に設立され、ルマン市に本拠地を持つ。同社は音声認識技術に基づく、性別等の話者情報の特定アプリケーションや転写アプリケーションを企業向けに提供する。LIUM の 3 名の研究者が関わっており、同社の最高経営責任者は LIUM の元学生である。LIUM には技術のプロトタイプなどがあるが、すぐに商用化できるものではないので、スピンオフ企業を通して、研究を産業へと移転することが必要であった。ルマン市があるサルトル県は同市を最先端デジタル技術都市として発展させることを望んでおり、Voxolab の創設は地域の経済振興政策に沿うものであった。

#### LIUM 及び欧州の研究機関の特徴について (米大企業と比べて)

- ・ LIUM 等の欧州の研究機関は、米グーグル社や米マイクロソフト社と比べて、非営利組織であり、研究協力のさいに信頼できるパートナーである。例えば、多額の給料を約束して、研究者を引き抜いたりしない。
- ・ LIUM 等の欧州の研究機関は教育機関でもあり、次世代の研究者を育成している。
- ・ LIUM 等の欧州の研究機関は、欧州等の企業を米大企業に依存させることがなく、企業を活気づける。また企業データを漏らすことがない。

<sup>99</sup> <http://www.deeplingo.com/about>

<sup>100</sup> <http://voxolab.com>

- ・ 北米では、アラビア語や北京語を英語に翻訳する技術が非常に開発されているが、欧州では多くの言語があるため、様々な組み合わせの翻訳技術が開発されている。
- ・ グーグル翻訳で中国語を仏語に翻訳するさいに、英語のミスが見つかる場合がある。同社は中国を直接仏語に翻訳するのではなく、中国語を英語に翻訳し、それを仏語に翻訳しているようである。
- ・ アイルランドは、英語使用国であることと法人税が低いことで、欧州で非常に多くの利益を得ている。
- ・ アメリカは英語使用国なので、多くの国の人々を惹きつけているが、フランスでは、仏語ができないと日常生活で難しい面があり、アメリカとは事情が異なる。だが、フェイスブックが人工知能分野の研究センターを設立するなど、動きもある。

参考資料:LIUM が参加する EU プロジェクトの概要

#### FP7 プロジェクト EUROMATRIXPLUS

略称	EuroMatrixPlus
正式タイトル	Bringing Machine Translation for European Language to the User
全予算(EU 拠出分)	594 万ユーロ(426 万ユーロ)
プロジェクト期間	2009 年 3 月～2012 年 2 月(36 ヶ月)
コーディネーター	ドイツ人工知能センター
参加組織	プラハカレル大学(チェコ)、ジョン・ホプキンス大学(英)、ダブリン・シティ大学(アイルランド)、ブルーノケスラー基金(伊)、ルーシーソフトウェア・サービス(独)、CEET(チェコ)、LIUM(仏)、エジンバラ大学(英)
ウェブサイト	<a href="http://www.euromatrixplus.net">http://www.euromatrixplus.net</a>

#### FP7 プロジェクト MATECAT

略称	MATECAT
正式タイトル	Machine Translation Enhanced Computer Assited Translation
全予算(EU 拠出分)	336 万ユーロ(265 万ユーロ)
プロジェクト期間	2012 年 2 月～2015 年 1 月
コーディネーター	ブルーノ・ケスラー基金(伊)
参加組織	エジンバラ大学(英)、TRANSLATED(伊)、LIUM(仏)
ウェブサイト	<a href="https://www.matecat.com">https://www.matecat.com</a>

#### FP7 プロジェクト EUMSSI

略称	EUMSSI
正式タイトル	Event Understanding through Multimodal Social Stream Interpretation
全予算(EU 拠出分)	326 万 2995 ユーロ(248 万ユーロ)
プロジェクト期間	2013 年 12 月～2016 年 11 月(36 ヶ月)
コーディネーター	バンペイファブラ大学(スペイン)
参加組織	IDIAP(スイス)、GFAI(独)、ドイチェ・ヴェレ(独)、ライプニッツ大学ハノーバー(独)、ビデオストリームネットワークス(スペイン)、LIUM

ウェブサイト	<a href="http://www.eumssi.eu">http://www.eumssi.eu</a>
--------	---

**CHIST-ERA プロジェクト JOKER**

略称	JOKER
正式タイトル	JOKe and Empathy of a Robot/ECA : Towards social and affective relations with a robot
EU 拠出分予算	85 万 4300 ユーロ
プロジェクト期間	2013 年 10 月～2016 年 9 月(36 ヶ月)
コーディネーター	CNRS-LIMSI(仏)
参加組織	ダブリン・トリニティ・カレッジ(アイルランド)、Koç 大学 (トルコ)、UMONS (ベルギー)、LIUM (仏)
ウェブサイト	<a href="http://www.chistera.eu/projects/joker">http://www.chistera.eu/projects/joker</a>

**CHIST-ERA プロジェクト M2CR**

略称	M2CR
正式タイトル	Multimodal Multilingual Continuous Representation for Human Language Understanding
EU 拠出分予算	60 万ユーロ
プロジェクト期間	2015 年 10 月～2018 年 9 月(36 ヶ月)
コーディネーター	LIUM(仏)
参加組織	モントリオール大学(カナダ)、コンピュータビジョンセンター(スペイン)
ウェブサイト	<a href="http://www.chistera.eu/projects/m2cr">http://www.chistera.eu/projects/m2cr</a>