

# The Challenge of Machine Translation of Patent Specifications and the Approach of the European Patent Office

**Georg Artelsmair**

Head of Department European Affairs/Member States  
European Patent Office

Prague, 14 May 2009

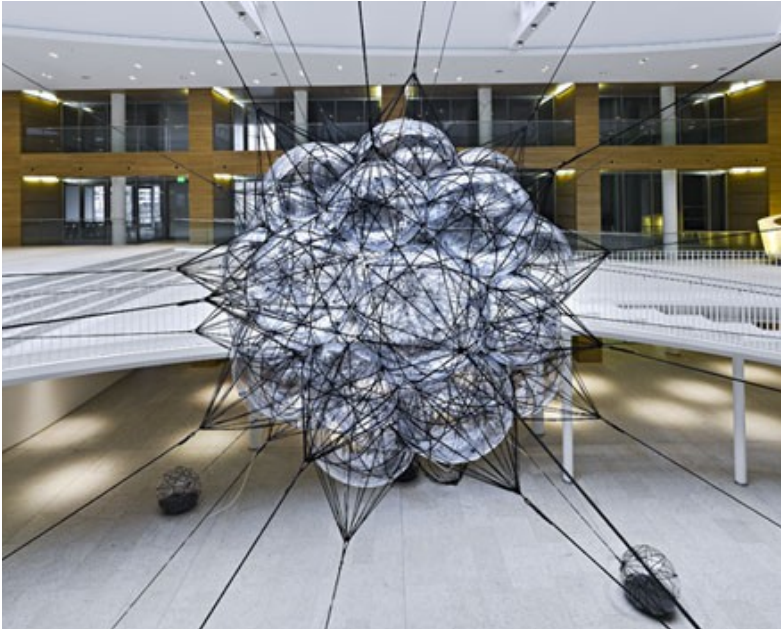




# Contents

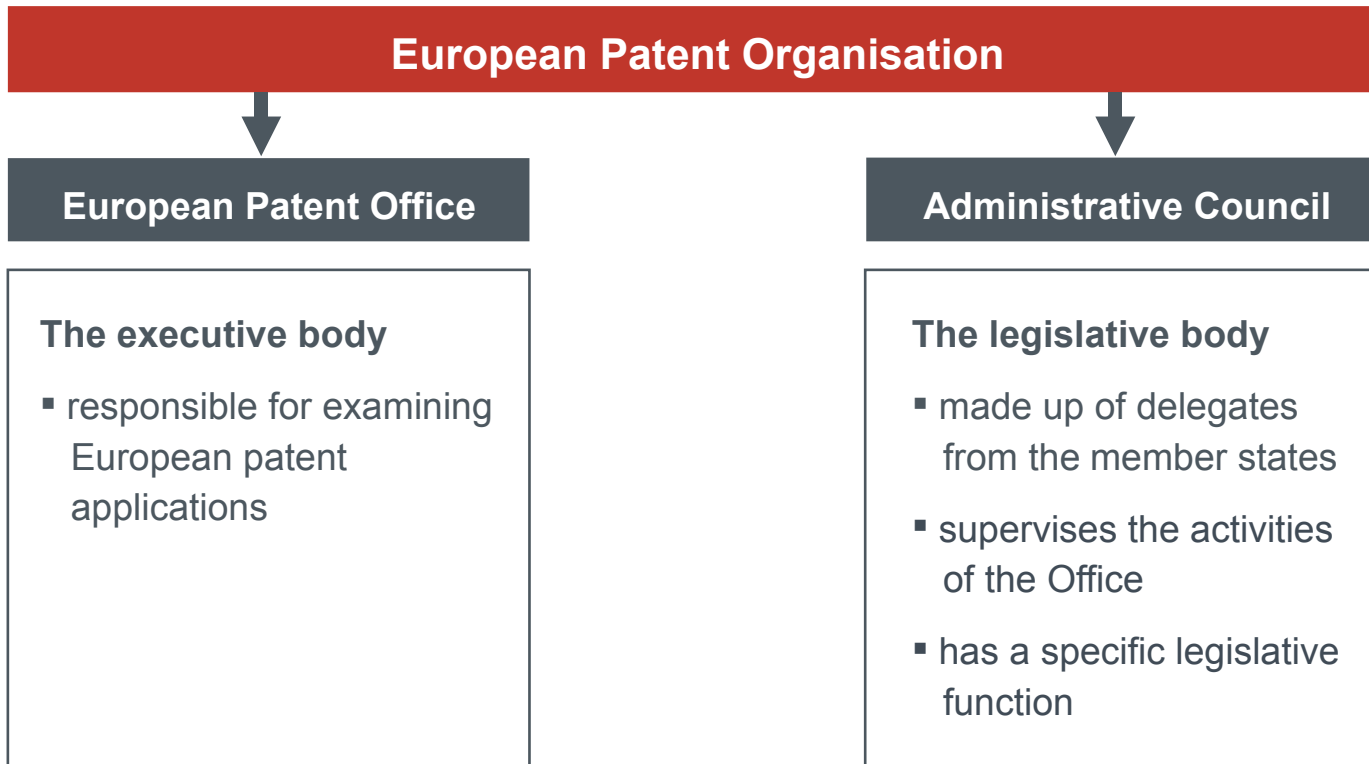
- Brief presentation of the European Patent Office
- EMTP Background
- EMTP Current approach
- EMTP Status
- Aims of the future enhanced EMTP programme
- Challenges

# The mission of the European Patent Office



As the Patent Office for Europe, we support **innovation, competitiveness and economic growth** across Europe through a **commitment to high quality and efficient services** delivered under the European Patent Convention.

# Structure of the European Patent Organisation



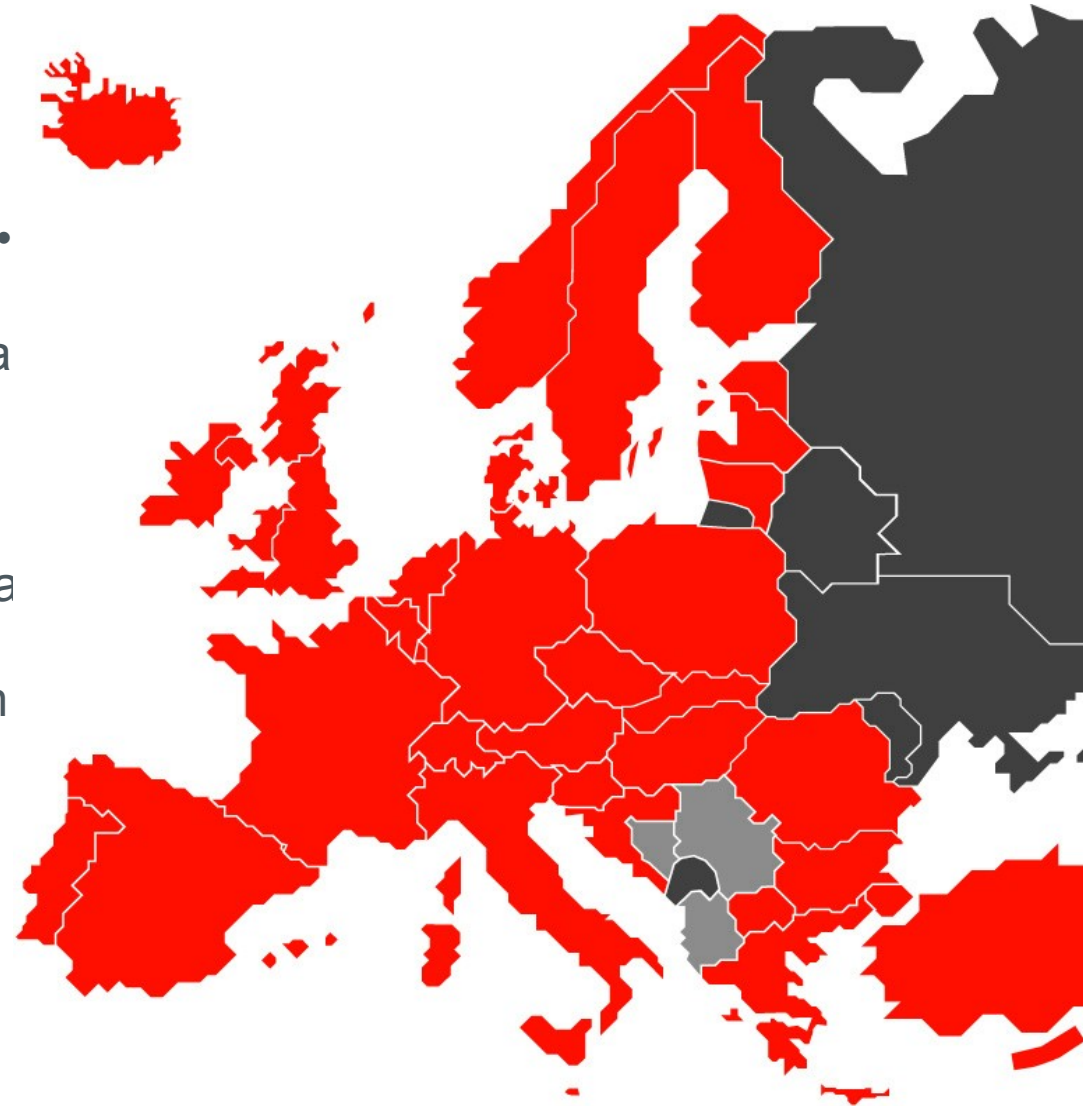
## 35 member states

Austria • Belgium • Bulgaria • Croatia •  
Cyprus • Czech Republic • Denmark •  
Estonia • Finland • France • Germany •  
Greece • Hungary • Iceland • Ireland •  
Italy • Latvia • Liechtenstein • Lithuania  
Luxembourg • Former Yugoslav  
Republic of Macedonia • Malta •  
Monaco • Netherlands • Norway •  
Poland • Portugal • Romania • Slovakia  
Slovenia • Spain • Sweden •  
Switzerland • Turkey • United Kingdom

European patent applications and patents  
can also be extended at the applicant's  
request to the following states:

Albania • Bosnia-Herzegovina • Serbia

Status: January 2009



# Autonomy



- Second largest intergovernmental institution in Europe
- Not an EU institution
- Self-financing, i.e. revenue from fees covers operating and capital expenditure

# Number of staff

Munich	3 514
The Hague	2 579
Berlin	283
Vienna	119
Brussels	4
<b>Total</b>	<b>6 499</b>



**Around 60% are  
patent examiners**

Status: December 2007

# Locations



The EPO has offices at five different locations.

Its headquarters are in Munich.

For more information visit [www.epo.org](http://www.epo.org)



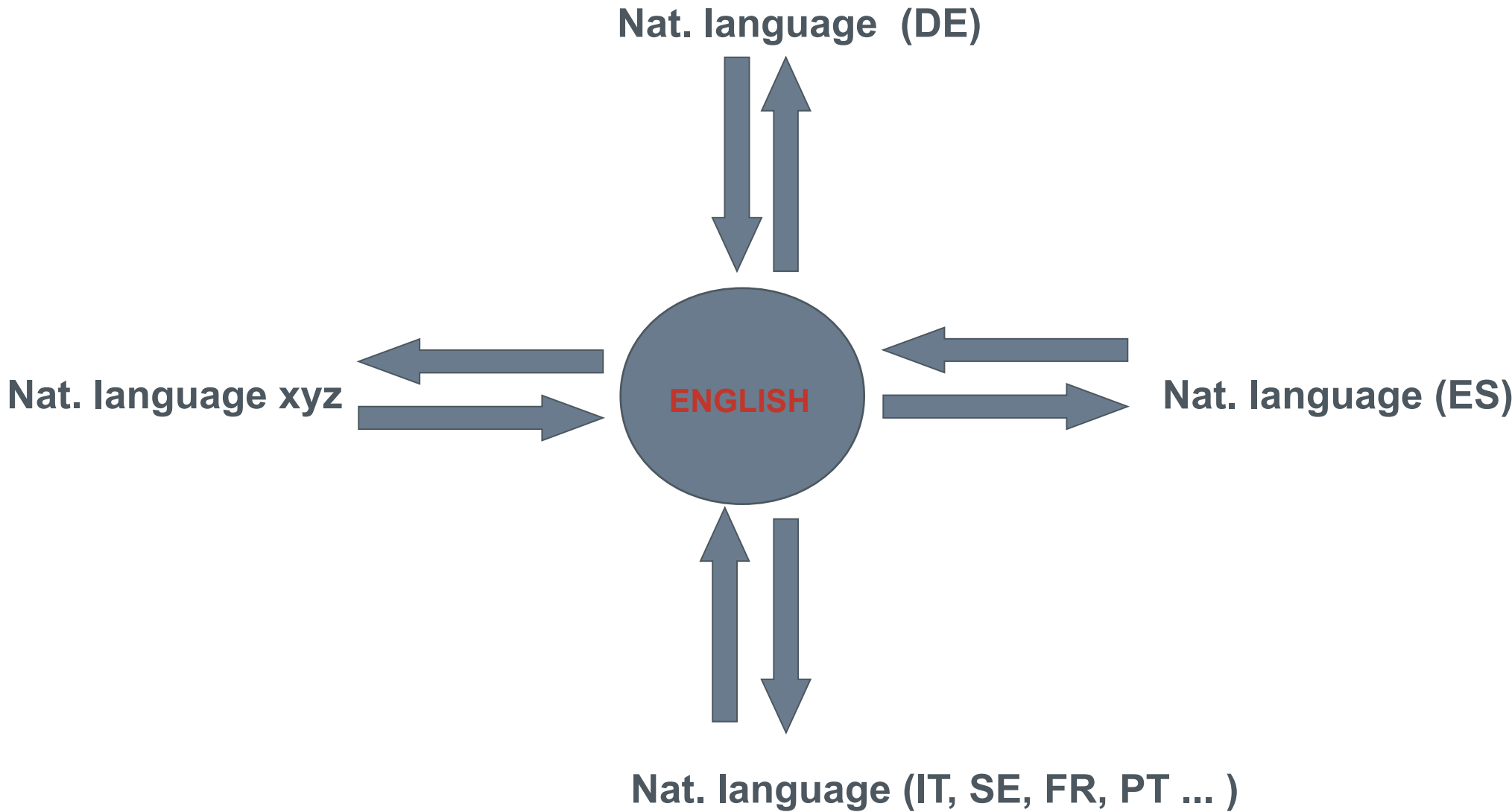
# Background

- 2004: Approval of the European Machine Translation Programme (EMTP) by the Administrative Council of the EPO
- Technical approach used: rule-based engine, dictionaries built on IPC-based patent terminology
- 2008: first language pairs (EN-ES and EN-DE) entered into production
- No suitable rule-based translation engines for certain EPO languages
- European Council document from 7 April 2009 addressing the translation issue related to the Community Patent

# Current EMTP Approach

- EPO buys a service for offering machine translations. We are not in the business of building translation engines
- Added value: Technical dictionaries based on the IPC classification scheme which the EPO builds together with external suppliers
- Current implementations are with English as a pivot language
- Current engines are rule based

# What We Have Today





# EMTP Status

- Delivery of translations on two levels:
  - Level 1: Generic dictionaries
  - Level 2: Dictionaries with technical terms based on IPC classification
- Automatic, real-time translations can be obtained via:
  - esp@cenet
  - Machine Translation Portal (<http://epo.worldlingo.com/>)
  - SEA Viewer (for EPO examiners)

# Status per Language – IPC Dictionaries Complete

- **German**

- Co-operation between the EPO and the German Patent and Trademark Office, the Austrian Patent Office and the Swiss Federal Institute of Intellectual Property
- Nearly 60,000 translation requests in December 2008
- 10-fold increase in translation requests within one year (between January 2008 and January 2009)

- **Spanish**

- Co-operation between the EPO and the Spanish Patent and Trademark Office
- Nearly 35,000 translation requests in December 2008.
- 7-fold increase in translation requests within one year (between January 2008 and January 2009)

# Status per Language – IPC Dictionaries in Progress

- **Italian**

- Co-operation between the EPO and the Italian Patent Office
- About 30,000 translation requests in December 2008
- In production with 30,000 entries per language direction (it↔en)

- **Swedish**

- Co-operation between the EPO and the Swedish Patent and Registration Office
- IPC dictionaries (level 2) available for internal testing via the MT Portal.
- Technical dictionaries complete
- Project go-live planned for late June 2009

# Status per Language – Ongoing Implementations

- **French**
  - Co-operation between EPO and the Belgian, French, Monaco and Swiss IP Offices
  - Translations currently available for internal testing via the MT Portal and the SEA Viewer
  - Operation planned for 3Q 09
- **Portuguese**
  - The English→Portuguese language direction planned for 3Q 09
- **Dutch, Finnish and Greek in preparation**
- **Romanian and Turkish: project implementation is conditional on the identification of suitable translation engines**

# Aims of the future enhanced MT programme at the EPO

- Explore available MT technologies that will enable us to introduce a machine translation system in the next 5 to 10 years, which
  - ➔ is tailor-made for translating patent specifications
  - ➔ offers scalable machine translation services from the three EPO official languages into all EPO languages

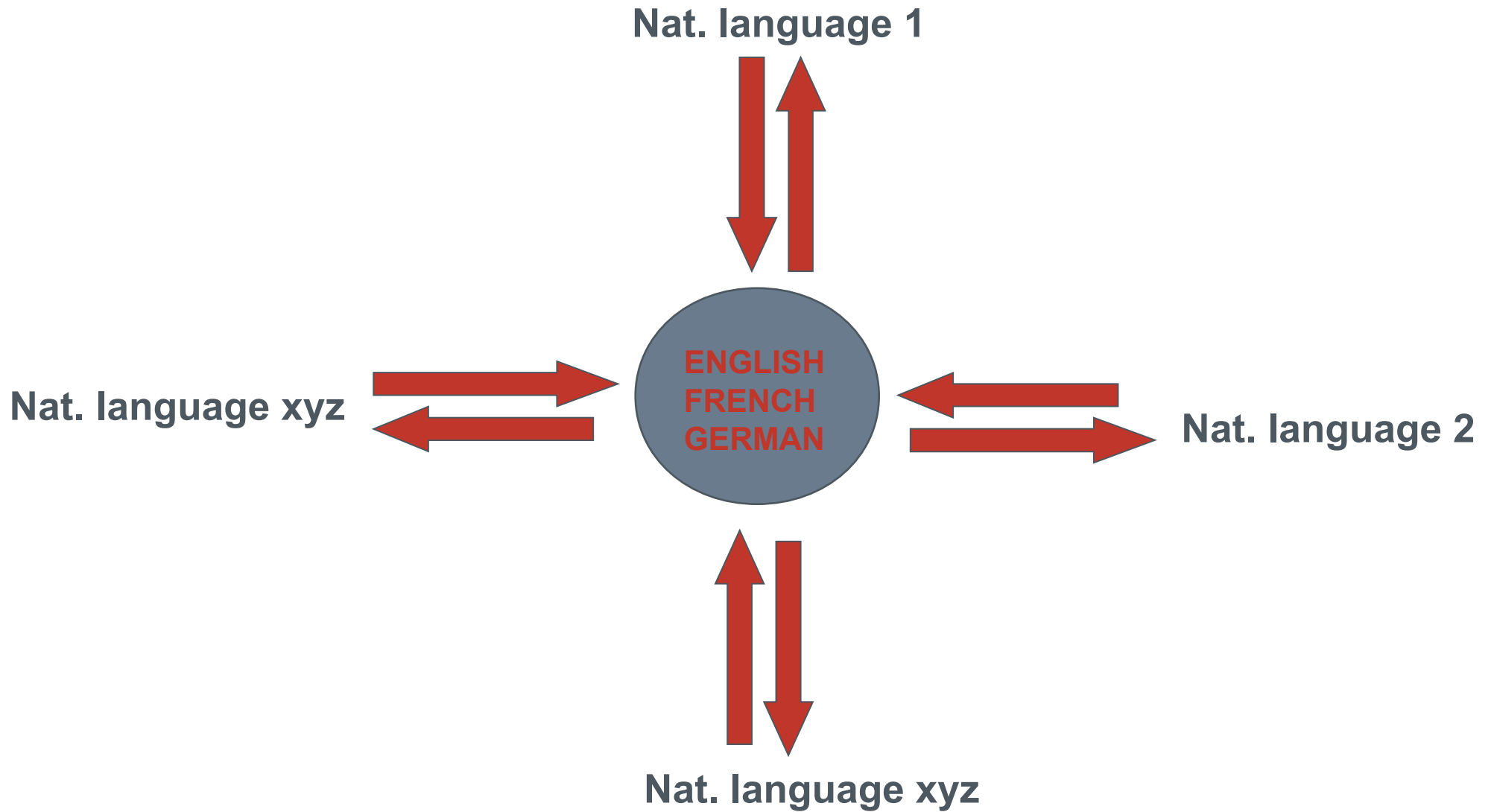




# Justification

- To enable access to patent information to enterprises, researchers and technically qualified users in Europe
- To support the London agreement
- To serve as a contribution to resolving the translation/language issue related to the Community Patent
- To make national documentation available to EPO examiners

# What We Will Need in the Future



# Which Technical Solution? What Implications?

- Rules-based vs. statistical engine? or a hybrid approach?
- Which criteria shall be applied for comparable testing of different engines?
- How can different engines be combined in one translation system?
- What would this mean for the architecture of the overall system?
- What would the selection of an engine imply for the implementation process of the MT services?



# Challenges

- Finding a suitable translation engine for languages such as Finnish, Romanian, Estonian, Turkish, etc.
- Improving the quality of translations
- Managing patent specificity
- Operating, maintaining and improving of the MT service



# Question

**What would researchers need from the EPO  
in order to make progress in this area?**



**Thank you for your attention**