



## Artificial Intelligence & Machine Learning in Intellectual Property Administration

# WIPO IP information roundtable

Bruno Pouliquen, Michal Ziemski

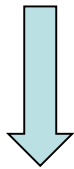
head

Advanced Technology Applications Center

World Intellectual Property Organization (WIPO)

# Team: ATAC

Advanced  
Technology  
Applications  
Center



*explore  
machine learning  
techniques in IP*



Christophe - *MT/Image*



Michal - *MT*



Daniel – *MT*



Bruno - *MT*



Akshat – *Speech*



Husaini – *Speech*



Sandrine - *partnerships*

Image?

# Summary of WIPO **AI**-based Tools and Applications

*AI=“Deep supervised machine learning”*



## **Text processing:**

- WIPO Translate
- Classification
- Similarity
- Transliteration



## **Speech processing:**

- Speech to text
- Speech to translated text
- Search in video-audio



## **Image similarity processing:**

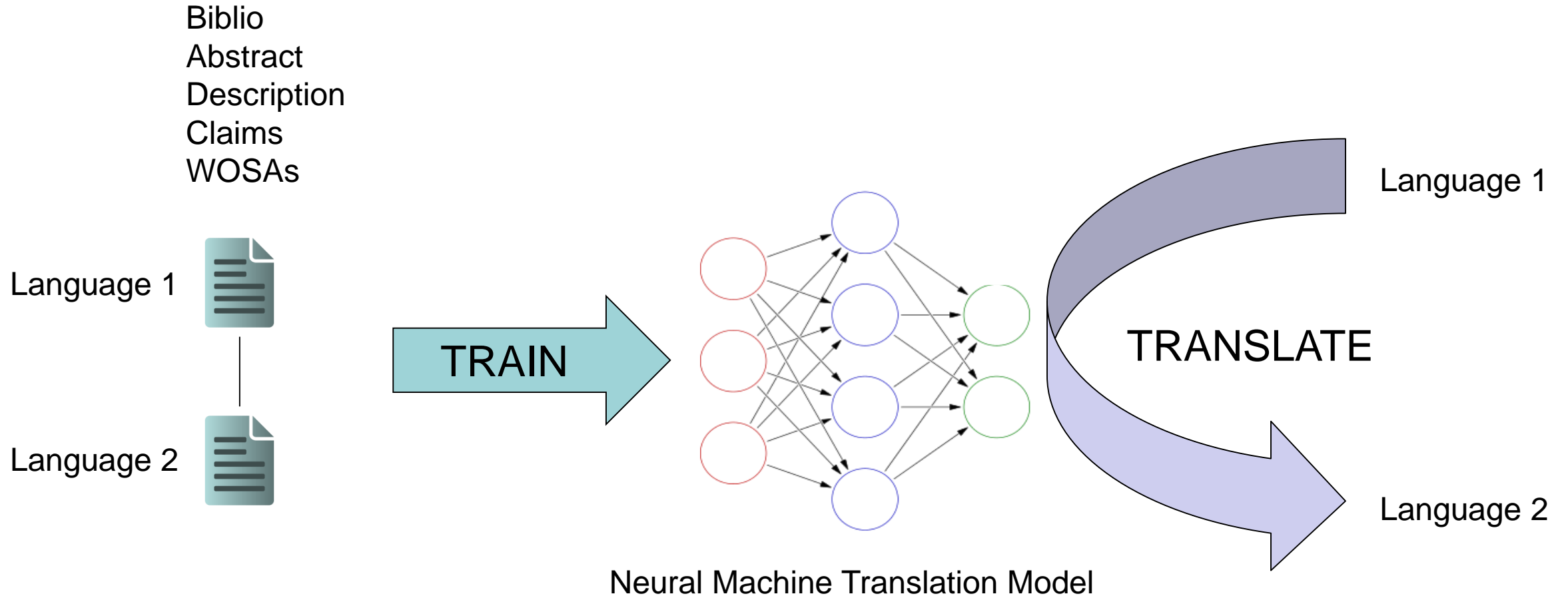
- Image classification  
(logos & industrial designs)
- Image-based search



# WIPO Translate

- In place since 2011, using neural networks since 2016
- Competitive for in-domain texts
- More than hundred different models
- Language covered for patents: Arabic, German, Spanish, French, Italian, Japanese, Korean, Portuguese, Russian, Chinese, Finish, coming: Polish
- In-house development, based on open-source:
  - Can translate confidential information
  - We can adapt it to any domain (provided that we have enough training data)
  - We can share the tool with others

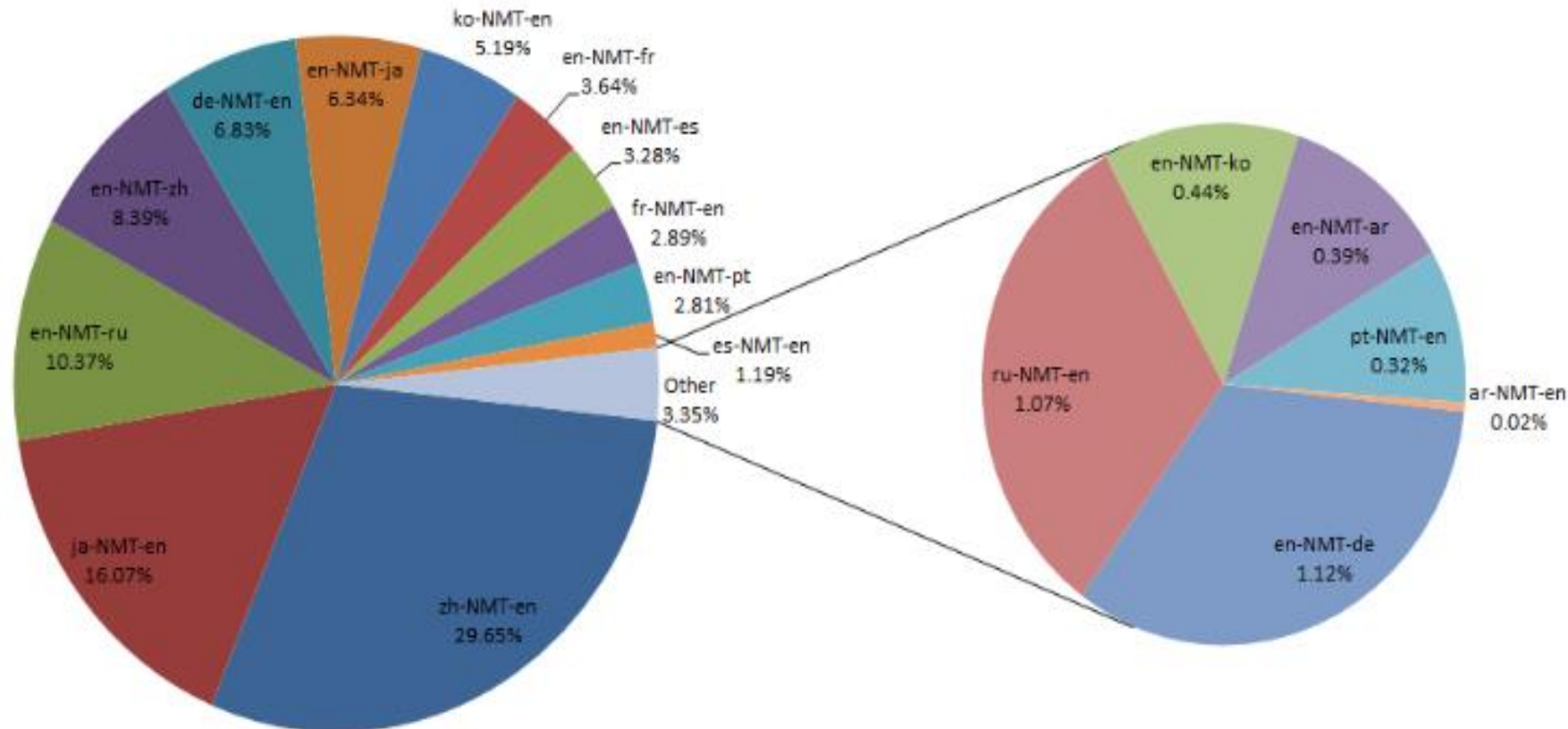
# WIPO Translate Architecture





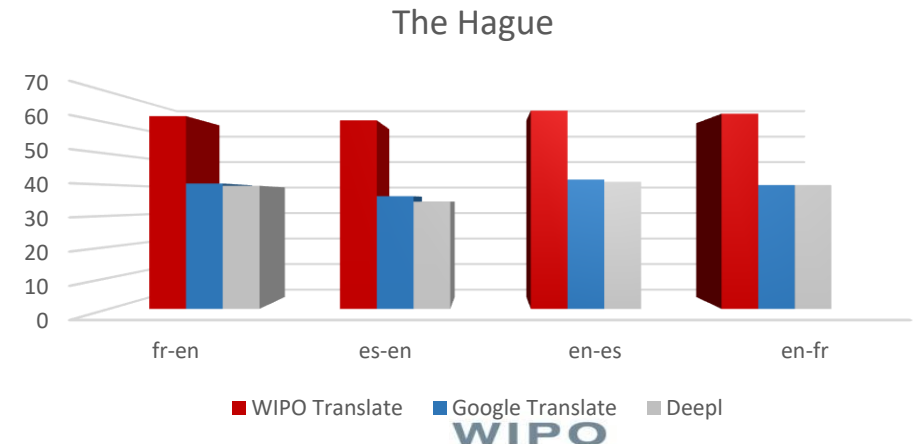
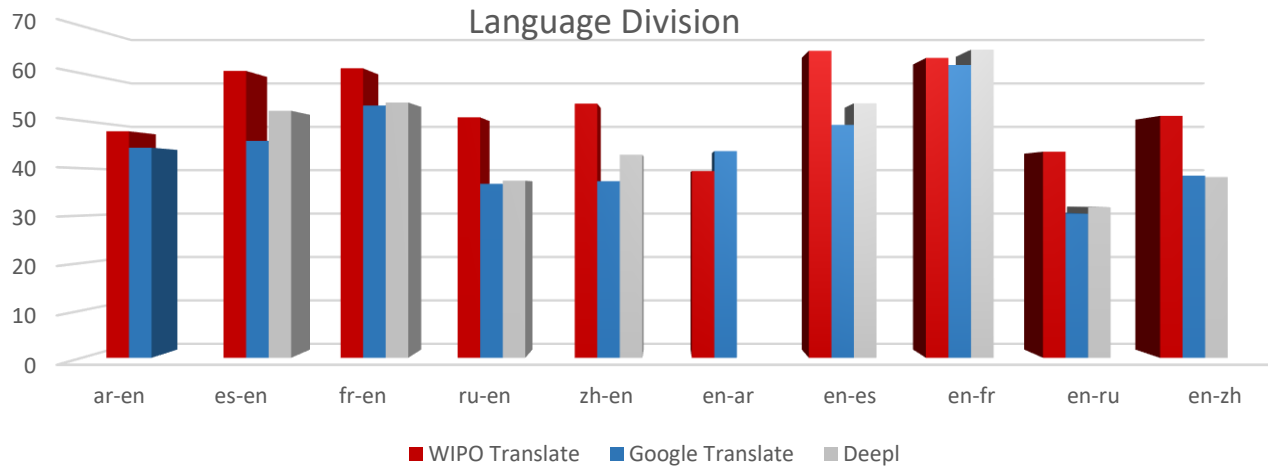
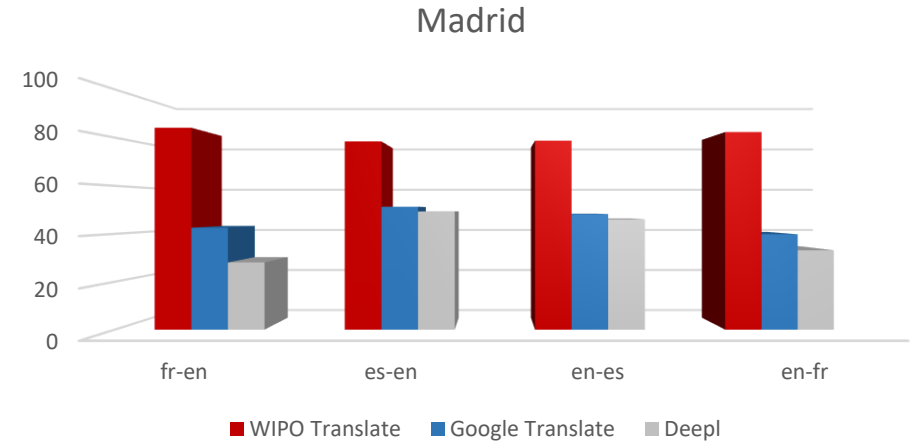
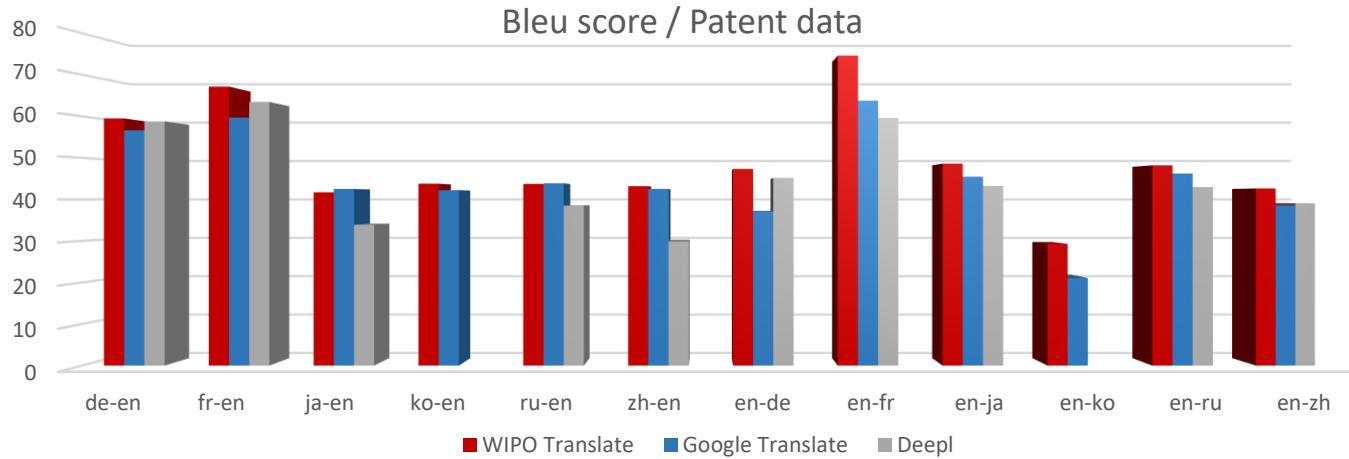
# WIPO Translate usage on PATENTSCOPE

~ 3.5 Million words translated every day  
65 % Chinese, Japanese and Korean





# Competitive: WIPO Translate usually better compared to commercial tools using automatic metric



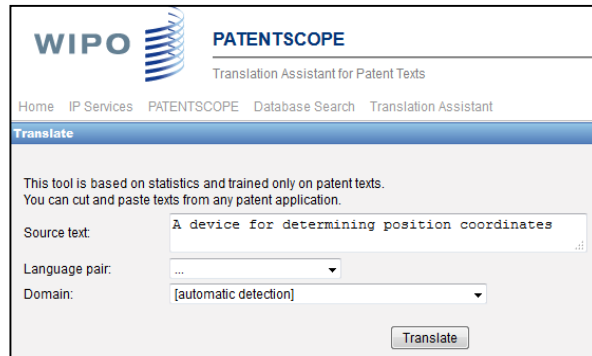
BLEU: measures similarity with a human translation reference, higher is better.

We use recent in-domain translations, never used in training \*(for Patents: only on bibliographic data)\*\*

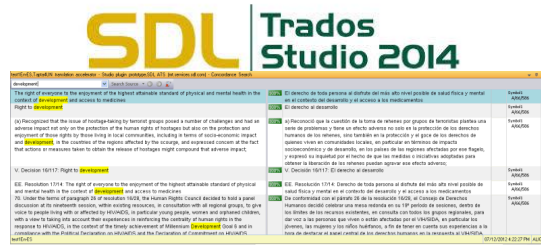
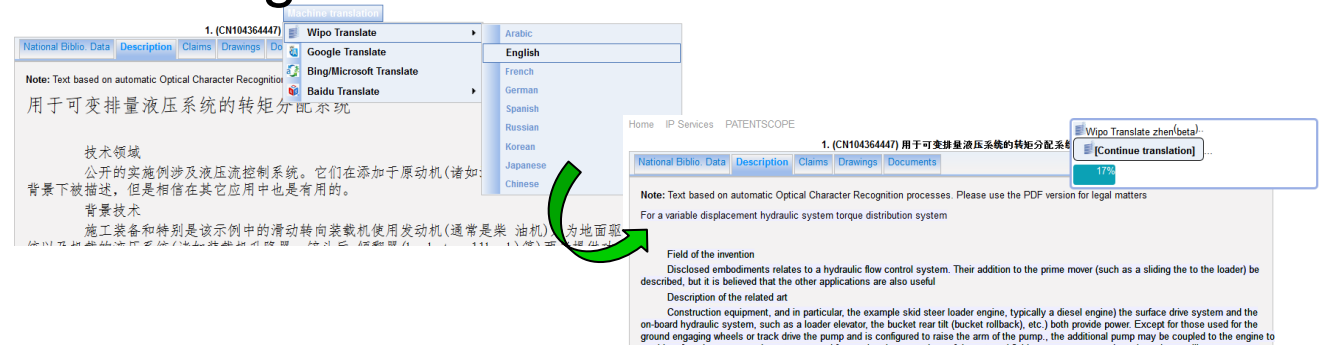


# Various Use Cases

## Web



## Html widget



DOC/XML translate



# WIPO Speech-to-Text



Publically available at:  
<https://www.wipo.int/s2t>

← → ↻ 🔒 wipo.int/s2t/welcome.html

WIPO

Home › Knowledge › Meetings › Speech-To-Text

## WIPO Speech-to-Text

Search WIPO meeting transcripts

Search WIPO meeting transcripts

English

ALL

### Latest Speech-to-Text transcripts



Automatic transcription of the CDIP24 session that took place on 20.11.2019 (afternoon)

Wed, 20 Nov 2019 18:32:21 CET



Automatic transcription of the CDIP24 session that took place on 20.11.2019 (morning)

Wed, 20 Nov 2019 15:10:32 CET

WIPO

User Guide

Contact Us

Transcript Language

Text File **TXT**

Download **ZIP**

Search CDIP24 transcripts

Home › Knowledge › Meetings › Speech-To-Text



Set playback speed: 2x 1.5x 1x

**Disclaimer** These transcripts are generated automatically by WIPO Speech-to-Text © from the audiovisual recording. All non-English language versions of the transcripts are generated automatically from the English text by WIPO Translate ©. The accuracy of the transcripts and their translations cannot therefore be guaranteed. Only the original speech of the audiovisual recording constitutes the official record of the meeting proceedings.

Powered by WIPO Speech-to-Text

And should take into account all the comments made by my colleagues and the other delegates so as to improve the projects of the Development Agenda. Thank you.

I thank the Delegations of Jordan.

The Distinguished Delegate, we have heard views, observation and comments on the proposals.

First, there was a request by the proponents that we can discuss and take a **decision** of this project at this session.

Second, I have heard many Delegations saying that it will be beneficial to Member States and those who would like to have a project proposal in the long run.

Third, I happen, heard any clear objection for WIPO who have been speaking about these project proposals.

But last one, we have heard Uganda on behalf of the African Group.

It is not even saying objection, but they would like to look at it again.

We have to confer and stimulus undertaking with the previous proposal.

So the request from African Group that has been delivered by Uganda, but I can

I will make myself clear, and I understood fully that your request to be postponed the discussion and you will come back to us before the end of

- Language switching on the floor:

[https://www.wipo.int/s2t/CWS9/CWS\\_9\\_2021-11-05\\_PM\\_1\\_xx\\_mp4\\_en.html?start\\_time=5249.18](https://www.wipo.int/s2t/CWS9/CWS_9_2021-11-05_PM_1_xx_mp4_en.html?start_time=5249.18)

- Floor: Spanish, [https://www.wipo.int/s2t/CWS9/CWS\\_9\\_2021-11-04\\_PM\\_1\\_xx\\_mp4\\_en.html?start\\_time=6032.18](https://www.wipo.int/s2t/CWS9/CWS_9_2021-11-04_PM_1_xx_mp4_en.html?start_time=6032.18)



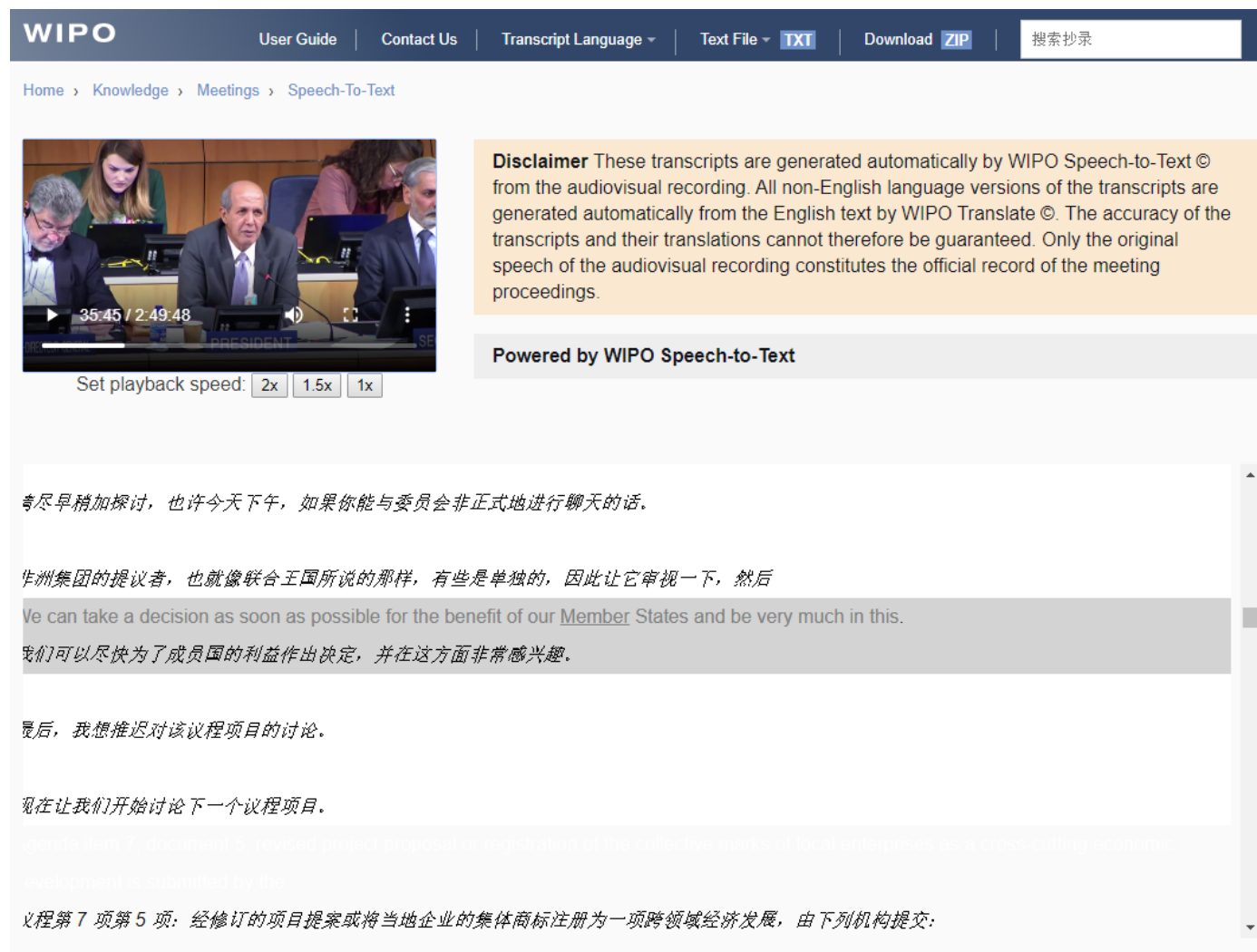
# WIPO Speech-to-Text (S2T)

- Model trained on WIPO meetings and corresponding verbatim  
=> augmented by other institutions' data
- Understands EN, FR, ES, RU, ZH, AR, PT and IT directly
- Use cases:
  - Verbatim aligned with video
  - Send quick automatic verbatim reports in 6 languages
  - Send verbatim for translation
  - Full-text search engine for video/audio content
  - Live captioning
  - Speaker recognition, Journaling

# WIPO S2T - MT

■ Using Machine Translation as an option

■ Powered by WIPO Translate



The screenshot displays the WIPO S2T - MT interface. At the top, there is a navigation bar with the WIPO logo, a search bar, and links for User Guide, Contact Us, Transcript Language, Text File (TXT), and Download (ZIP). Below the navigation bar, the breadcrumb trail reads: Home > Knowledge > Meetings > Speech-To-Text.

The main content area features a video player showing a meeting recording. The video player includes a progress bar, a play button, and a timestamp of 35:45 / 2:49:48. Below the video player, there is a playback speed control set to 2x, with options for 1.5x and 1x.

To the right of the video player, there is a disclaimer box with the following text: **Disclaimer** These transcripts are generated automatically by WIPO Speech-to-Text © from the audiovisual recording. All non-English language versions of the transcripts are generated automatically from the English text by WIPO Translate ©. The accuracy of the transcripts and their translations cannot therefore be guaranteed. Only the original speech of the audiovisual recording constitutes the official record of the meeting proceedings.

Below the disclaimer, there is a section titled "Powered by WIPO Speech-to-Text".

The transcript area shows the following text in Chinese and English:

请尽早稍加探讨，也许今天下午，如果你能与委员会非正式地进行聊天的话。

非洲集团的提议者，也就像联合王国所说的那样，有些是单独的，因此让它审视一下，然后

We can take a decision as soon as possible for the benefit of our Member States and be very much in this.

我们可以尽快为了成员国的利益作出决定，并在这方面非常感兴趣。

最后，我想推迟对该议程项目的讨论。

现在让我们开始讨论下一个议程项目。

议程第 7 项第 5 项：经修订的项目提案或将当地企业的集体商标注册为一项跨领域经济发展，由下列机构提交：

# WIPO S2T Search

## WIPO Speech-to-Text search

Japan

English

ALL

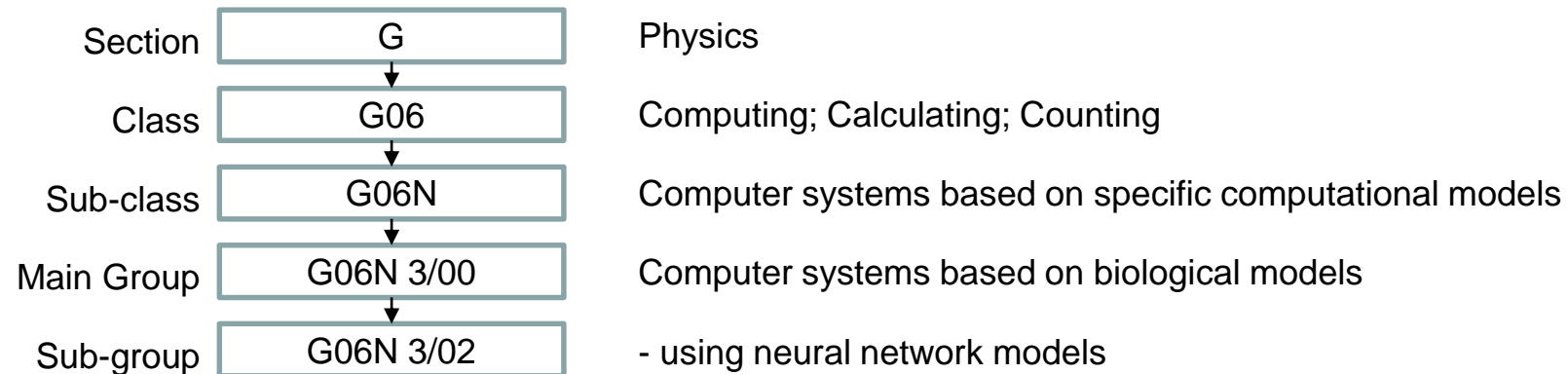
Search WIPO meeting transcripts

### Search results

Text snippet result	Conference	Date	Media timestamp
Under the <a href="#">Japan</a> funds in trust <a href="#">Japan</a> has been conducting various activities.	CDIP24	2019-11-18	2 hours, 13 minutes, 38 seconds
I recognize the Delegations of <a href="#">Japan</a> .	CDIP24	2019-11-18	1 hour, 46 minutes, 17 seconds
So through the <a href="#">Japan</a> funds in trust <a href="#">Japan</a> has been assisting WIPO to advance initiatives for enhancing	CDIP24	2019-11-18	2 hours, 12 minutes, 16 seconds
I wish to thank the delegation of <a href="#">Japan</a> for his intervention.	CDIP24	2019-11-18	1 hour, 49 minutes, 23 seconds
Going forward the government of <a href="#">Japan</a> in cooperation with WIPO is committed to further improving its,cooperative initiatives to ensure that the <a href="#">Japan</a> funds in trust are used even more efficiently and effectively	CDIP24	2019-11-18	2 hours, 14 minutes, 7 seconds

# International Patent Classification

- A “**hierarchical** system of language independent symbols for the classification of patents...”





# IPC/CPC classification

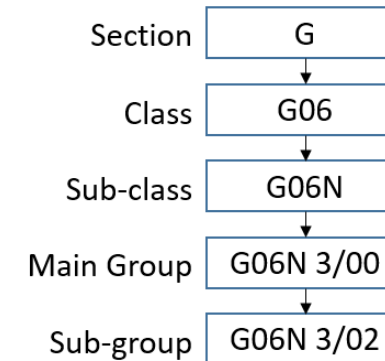
- IPC classification
  - Similar to IPC Cat, but built/trained in-house
  - Used in production to classify NPL texts
- CPC classification, trained but not yet used
- Ongoing projects e.g. with WIPO Green
  - Create a custom tool trained on specific data
  - predict the “WIPO Green” category of new incoming document (patent, NPL or user provided document)



# IPC Classification Accuracy

- Trained specifically on 27M WIPO documents (title + abstract)
- Using neural networks for classification

IPC depth	Example	Top-3*	Top-5*	Comment
Sub-class	G06N	> 95 %	> 98 %	
...	...	...	...	
Sub-group	G01N 0033/543000	< 80 %	> 80 %	Full IPC




\* Top-X Accuracy: considered correct if any 1 of the top X predictions is correct.

# Vienna classification assistant



<https://www3.wipo.int/bnd-api/vienna-classification-assistant/>

WIPO IP PORTAL MENU *Vienna Classification Assistant* Covid-19 Update X HELP ENGLISH LOGIN WIPO



added1569402813251.png Edit/crop... Clear

Quick add...

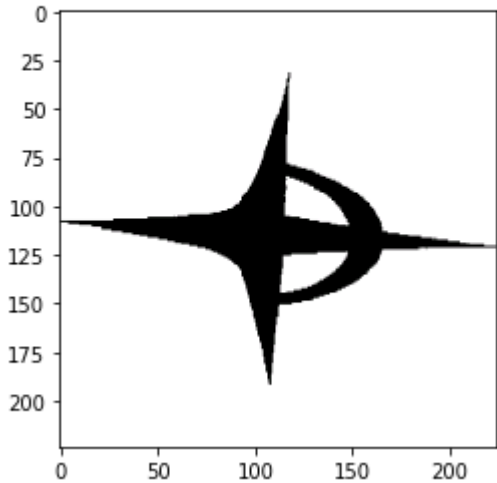
Select suggested sections  
or start a manual selection on the right.

1 CELESTIAL BODIES, NATURAL PHENOMENA, GEOGRAPHICAL MAPS	2 ★★ HUMAN BEINGS	3 ★★★★★ ANIMALS	4 SUPERNATURAL, FABULOUS, FANTASTIC OR UNIDENTIFIABLE BEINGS	5 PLANTS
6 LANDSCAPES	7 CONSTRUCTIONS, STRUCTURES FOR ADVERTISEMENTS, GAT...	8 FOODSTUFFS	9 TEXTILES, CLOTHING, SEWING ACCESSORIES, HEADWEAR, FOOTWEAR	10 TOBACCO, SMOKERS' REQUISITES, MATCHES, TRAVEL GOODS, FANS,...
11 HOUSEHOLD UTENSILS	12 FURNITURE, SANITARY INSTALLATIONS	13 LIGHTING, WIRELESS VALVES, HEATING, COOKING OR...	14 IRONMONGERY, TOOLS, LADDERS	15 MACHINERY, MOTORS, ENGINES
16 TELECOMMUNICATIONS, SOUND RECORDING OR REPRODUCTION,...	17 HOROLOGICAL INSTRUMENTS, JEWELRY, WEIGHTS AND MEASURES	18 ★ TRANSPORT, EQUIPMENT FOR ANIMALS	19 CONTAINERS AND PACKING, REPRESENTATIONS OF...	20 WRITING, DRAWING OR PAINTING MATERIALS, OFFICE REQUISITES,...
21 GAMES, TOYS, SPORTING ARTICLES, ROUNDABOUTS	22 ★ MUSICAL INSTRUMENTS AND THEIR ACCESSORIES, MUSIC ACCESSORIES,...	23 ARMS, AMMUNITION, ARMOUR	24 ★ HERALDRY, COINS, EMBLEMS, SYMBOLS	25 ★★ ORNAMENTAL MOTIFS, SURFACES OR BACKGROUNDS WITH...
26 ★★★★★ GEOMETRICAL FIGURES AND SOLIDS	27 ★★★ FORMS OF WRITING, NUMERALS	28 INSCRIPTIONS IN VARIOUS CHARACTERS	29 COLOURS	



# Challenging task

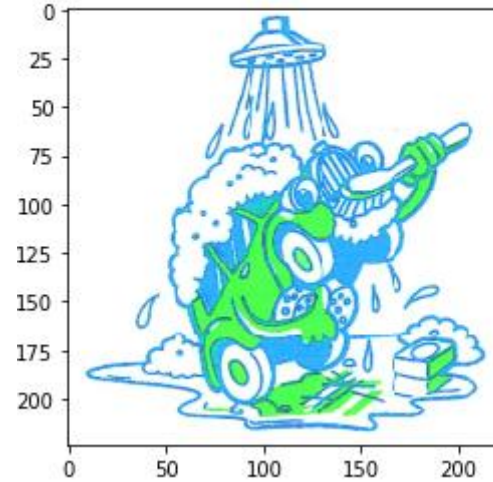
Label: 01.01.09 Stars with four points



**Inter-class confusion, latent hierarchy**

- 01.01.02, "One star"**
- 01.01.03, "Two stars"**
- 01.01.04, "Three stars"
- 01.01.05, "More than three stars"
- 01.01.06, "Concentric stars"
- 01.01.08, "Stars with three points"
- 01.01.09, "Stars with four points"**
- 01.01.10, "Stars with more than four points"
- 01.01.12, "Stars with uneven points"
- 01.01.13, "Stars with rays or radiating lines"**
- 01.01.14, "Incomplete stars"**
- 01.01.15, "Comets, stars with tail"
- 01.01.16, "Stars with points ending in spheres"

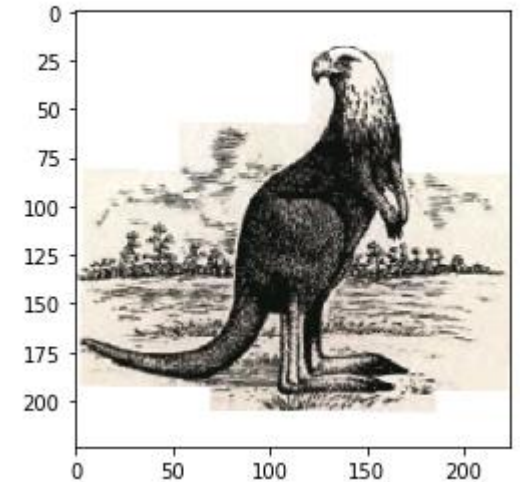
Label: 18.01.07 Motor road vehicles



**Multi-label:**

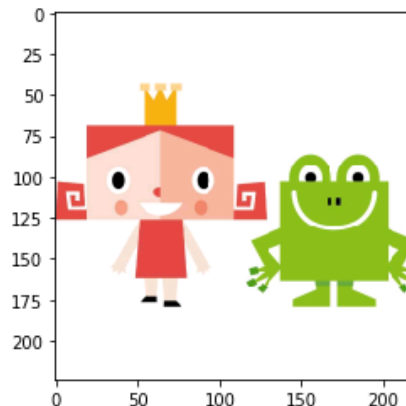
- Personified objects, combinations of objects representing a personage
- shower
- water
- drops
- soap
- brush
- ...

Label: 03.05.26 Animals of Series V standing

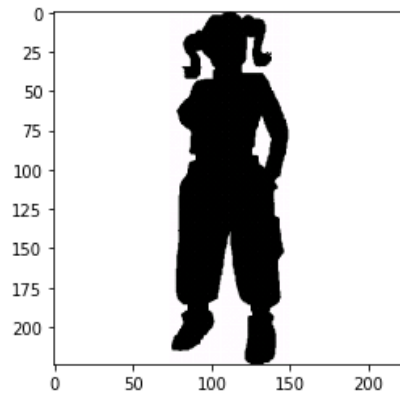


**• Intra-class variation, undefined hierarchy**

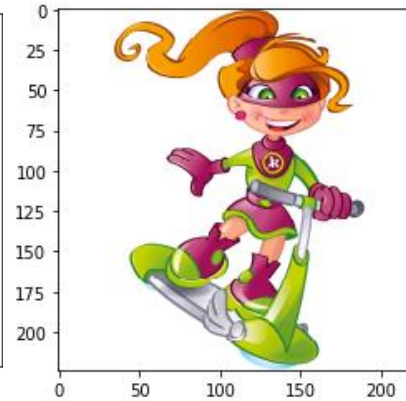
Label: 02.05.03 Girls



Label: 02.05.03 Girls



Label: 02.05.03 Girls



# Global Brand database

- Gateway to more than 45,700,000 records (incl. Madrid treaty)
- Allow for researching similar or identical trademark images
  - ➔ Strong need for image similarity tool
  - ➔ Machine learning based

# Global Design database

- Gateway to more than 13,470,000 records (incl. The Hague treaty)
  - ➔ similar image based similarity tools needed (even more challenging)


# WIPO Image Similarity Search



**FILTER BY**

Source **Image** Status Origin App. Year \* Expiration \*

Pick an image



edit delete

Pick a strategy

- Concept
- Shape
- Color
- Composite

Pick an image type

Verbal	0
<b>Nonverbal</b>	<b>65,232</b>
Combined	0
Unknown	0

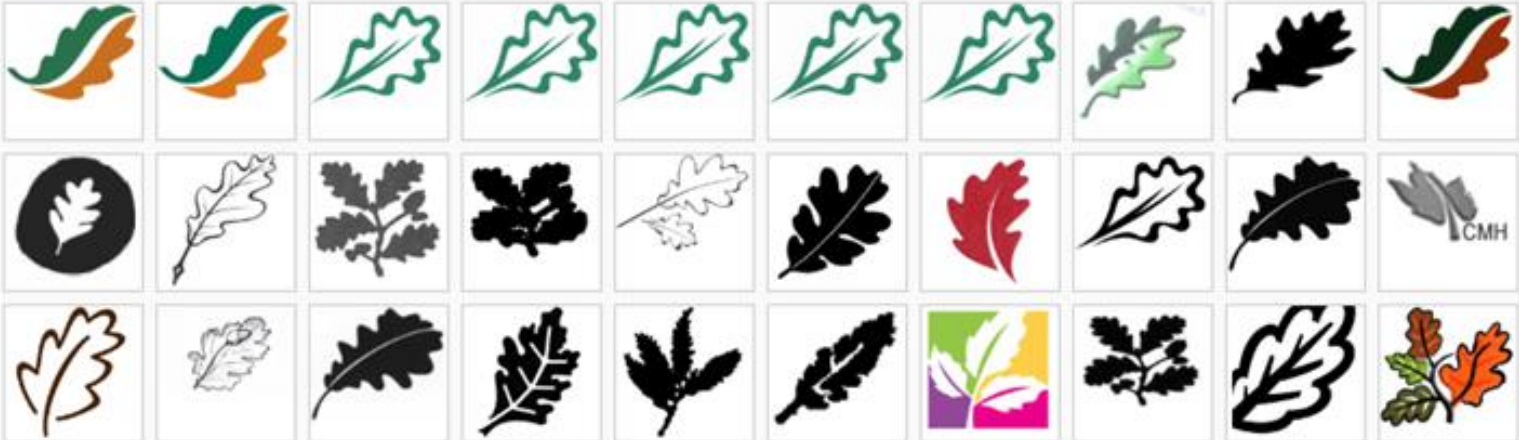
filter

**CURRENT FILTER**

ITY:Nonverbal \* IMAGE:Concept \*

Display: 100 per page options

Sort by Relevance - desc



# Global Brand database

- Gateway to more than 45,700,000 records (incl. Madrid treaty)
- Allow for researching similar or identical trademark images
  - ➔ Strong need for image similarity tool
  - ➔ Machine learning based

# Global Design database – Industrial Designs

- Gateway to more than 13,470,000 records (incl. The Hague treaty)
  - ➔ similar image based similarity tools needed (even more challenging)

# Image search: industrial designs



**FILTER BY** Source Status Designation Locarno Class Reg. Year \* Expiration \* Holder \*

AL Designs	69	BN Designs	50	BW Designs	0	CA Designs	43,245	CH Designs	31,052	CN Designs	6,120,653
CR Designs	707	DE Designs	194,709	EM Designs	843,450	ES Designs	250,529	FR Designs	780,945	GE Designs	1,249
ID Designs	27,557	IN Designs	59,718	IT Designs	49,115	JP Designs	598,850	JO Designs	1,145	KE Designs	754
KH Designs	621	KR Designs	866,183	LA Designs	240	MD Designs	2,301	MK Designs	377	MN Designs	663
MY Designs	5,152	MZ Designs	3	NZ Designs	11,863	TH ID	16,720	TN Designs	4,568	UA Designs	31,401
US Designs	0	WO Designs	40,314								


Display: List Sort: Value - asc filter

**CURRENT FILTER**  
STATUS:ACT \*


Display: 10 per page options 4 / 998,421

Sort by Reg. Date - desc


**NZ ID 428172**  
Sole for a Shoe  
2020-11-27  
FRANKIE4 IP 1 Pty Ltd



**NZ ID 428171**  
Sole for a Shoe  
2020-11-27  
FRANKIE4 IP 1 Pty Ltd



**NZ ID 428147**  
A Loader Grapple  
2020-11-27  
PAGE & MACRAE LIMITED



■ Trained a neural model to classify industrial design images (using the Locarno classification)

to come 2021: Image search for GDD

- Multiple images input
- Most similar designs retrieved



# Global Design Database Image Search

<https://www3.wipo.int/designdb/en/index.jsp>

- ❖ Image search for designs similar to trademark image search
- ❖ Challenges:
  - Multiple images per design
  - Very **different** images (visually) belonging to **same** design
  - **Similar** images can belong to **different** designs
  - How to get the most similar design from multiple images

Same design with different views:



Very similar views for very different designs (different Locarno class)



# Industrial Designs image search prototype

- ❖ Multiple-image query
- ❖ Retrieve the most similar designs from Global Design Db
- ❖ Extract the most similar design from the 1vs1 image similarity
- ❖ Under development

Example query with 3 images:



Top retrieved designs:



# Other AI projects in WIPO

- In close collaboration with business side
- Tailored to specific needs and expectations
- Can be exploratory, not always feasible:
  - Chatbot
  - HR workflow facilitation
  - Similarity / Filtering
  - ...

# Our strategy, what makes successful projects

- Careful choice of potentially successful projects
- Business driven - importance of integration
- Expectation management and proper evaluation
  
- Quick prototyping                    **“fail fast to innovate faster”**
  
- Technological survey, academic network
- Synergies through data sharing and exchange
- Open source technologies
  
- Use “recent/clean/big” data

Thank you



WIPO Translate & Patentscope

# WIPO Translate & Patentscope

## WIPO IP information roundtable

Bruno Pouliquen, Michal Ziemski

head

Advanced Technology Applications Center

World Intellectual Property Organization (WIPO)



## SIMPLE SEARCH

Using PATENTSCOPE you can search 98 million patent documents including 4.2 million published international patent applications (PCT).

[Detailed coverage information](#)

PCT publication 39/2021 (30.09.2021) is now available [here](#). The next PCT publication 40/2021 is scheduled for 07.10.2021. [More](#)

Check out the [new PATENTSCOPE features](#): CPC, NPL, Families ...

[Search Facility to Support COVID-19 Innovation Efforts](#)

Field

Front Page



Search terms...



[Query Examples](#)



# Assimilation: translation in PATENTSCOPE

<http://patentscope.wipo.int>

Sort by: Relevance View All List Length 10 Machine translation

Int.Class	Appl.No	Title	Applicant	Ctr	PubDate
B26B 21/10	201410380368.5	小尺寸剃刀系统	张鹭云	CN	26.11.2014
G01R 31/36	97193294.0	Electric shaver or apparatus cooperating therewith and data-evaluation arrangement	Braun Aktiengesellschaft	CN	21.04.1999

本发明的小尺寸剃刀系统被设计为剃刮小区域的毛发，该小尺寸剃刀系统包括：具有剃刀头的剃刀装置、至少一个剃刀片、以及手柄。剃刀头具有至少一个剃刀片。剃刀头与手柄连接。该剃刀装置选自由以下组成的组：手动的和电动的。该剃刀装置被设计为具有各种形状和尺寸。使用该小尺寸剃刀系统剃刮小区域例如腋窝和鼻孔中的毛发，对于剃刮者而言更方便、更容易。

## Result list translation

Sort by: Relevance View All List Length 10

Int.Class	Appl.No	Title	Applicant	PubDate
B26B 21/10	201410380368.5	小尺寸剃刀系统	张鹭云	26.11.2014
1214647		Electric shaver or apparatus cooperating therewith and data-evaluation arrangement		21.04.1999

[zhen]本发明的小尺寸剃刀系统被设计为剃刮小区域的毛发，该小尺寸剃刀系统包括：具有剃刀头的剃刀装置、至少一个剃刀片、以及手柄。剃刀头具有至少一个剃刀片。剃刀头与手柄连接。该剃刀装置选自由以下组成的组：手动的和电动的。该剃刀装置被设计为具有各种形状和尺寸。使用该小尺寸剃刀系统剃刮小区域例如腋窝和鼻孔中的毛发，对于剃刮者而言更方便、更容易。

## Translation of description and claims

Machine translation

1. (CN104162898) 小尺寸剃刀系统

National Biblio. Data Description Claims Drawings Documents

Note: Text based on automatic Optical Character Recognition processes. Please use the PDF version for legal matters

### 小尺寸剃刀系统

版权公告

本专利文档的公开内容的一部分包含受版权保护的材料。版权所有 者不反对任何人复制呈现在专利和商标局专利文件或记录中的专利文件 或专利公开内容，但在其它方面却保留 有版权权利。37 CFR 1.71 (d).

## Translation of title and abstract

IPC: B26B 21/10 B26B 21/52 B26B 21/40 B26B 21/56

Applicants: 张鹭云

Inventors: 张鹭云

Agents: 北京市提邦律师事务所 11389

Priority Data: (ZH) 小尺寸剃刀系统

Title: (ZH) Small size of the razor system

Abstract: (ZH) 本发明的小尺寸剃刀系统被设计为剃刮小区域的毛发，该小尺寸剃刀系统包括：具有剃刀头的剃刀装置、至少一个剃刀片、以及手柄。剃刀头具有至少一个剃刀片。剃刀头与手柄连接。该剃刀装置选自由以下组成的组：手动的和电动的。该剃刀装置被设计为具有各种形状和尺寸。使用该小尺寸剃刀系统剃刮小区域例如腋窝和鼻孔中的毛发，对于剃刮者而言更方便、更容易。

The small size of the shaver system is designed to shaving the area of the hair, the small size of the shaving system comprises: a razor head of the razor device, at least one of the razor blades, and a handle. The razor head has at least one of the razor blades connected to a razor cartridge and the handle. The razor device is selected from the group consisting of: the manual and electric the razor device is designed to have a variety of shapes and sizes. The use of the small size of the shaver system of a shaving small area, such as the armpits and nostril of the hair, for shaving the viewer to more conveniently and more easily.

COPYRIGHT NOTICE

Of the disclosure of this patent document contain material that is subject to copyright protection. The copyright owner has no facsimile reproduction by anyone presented in the patent and trademark office patent file or records of the patent document or the patent but otherwise reserves all copyright rights whatsoever.. 37 CFR 1.71 (d).

BACKGROUND OF THE INVENTION

The following information may aid in the understanding of the present invention. This is not an admission that any of the provided herein is prior art or with the present disclosure or claimed content-related material, or any explicitly or implicitly referenced documents of the prior art.

1. 技术领域

本发明总体上涉及剃刀和剃刮器的领域。

2. 相关技术的描述



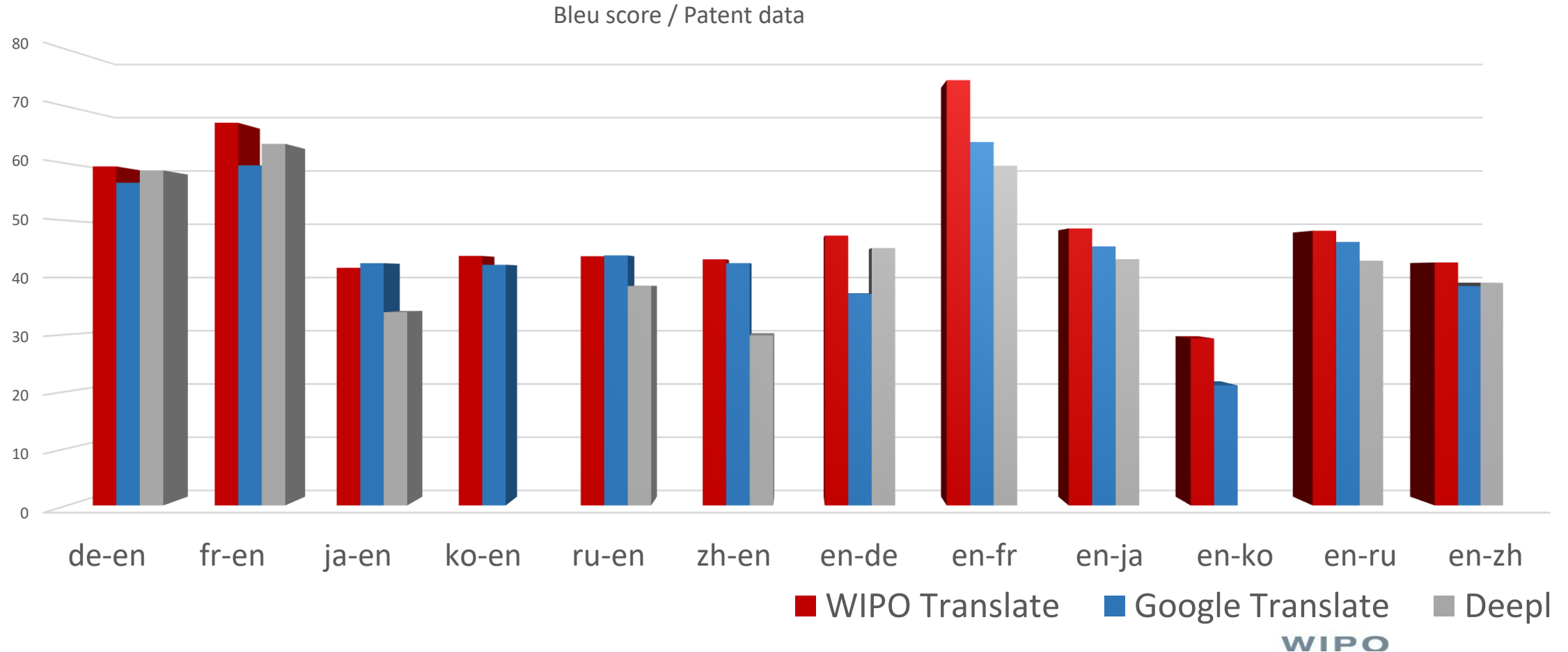
# WIPO Translate

- Updated models for:
  - Japanese, Korean, Chinese, Russian, German
- New model for Finnish
- Next Polish?
  
- Language covered for patents: Arabic, German, Spanish, Finnish, French, Italian, Japanese, Korean, Portuguese, Russian, Chinese, coming: Polish



# Competitive: WIPO Translate usually better compared to commercial tools using automatic metric

**BLEU: higher is better**



BLEU: measures similarity with a human translation reference.

We use recent in-domain translations, never used in training. Evaluation done on patent titles and abstracts only



# Current/future work

- New model for Finnish
- (Next Polish?)
  
- Working on translating subscripts/superscripts:
  - ... and italic/underline/bold
- Batch translation of fulltexts...
  - DPMA translated Chinese fulltexts and made it searchable
  - We translated selected patents for WIPO Green (any into English)