WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

用于知识产权行政管理的人工智能工具

由产权组织人工智能驱动

WIPO FOR OFFICIAL USE ONLY

内容

1. 产权组织人工智能简史
2. 产权组织人工智能项目内容提要（**已开展的**项目和正在研发的项目）
3. **重点**：
   - WIPO Translate（产权组织翻译工具），**用于**专**利**（**和其他**领域）
   - WIPO Image Similarity Search（产权组织图形近似性检索），**用于商**标
   - WIPO International Patent Classification auto-classifier（产权组织国际专利分类自动分类器），**用于**专利摘要（**和其他**）
   - WIPO Speech-to-Text（产权组织语音转文字），**用于会**议和大会
4. **向外拓展**产权组织的人工智能足迹

WIPO FOR OFFICIAL USE ONLY

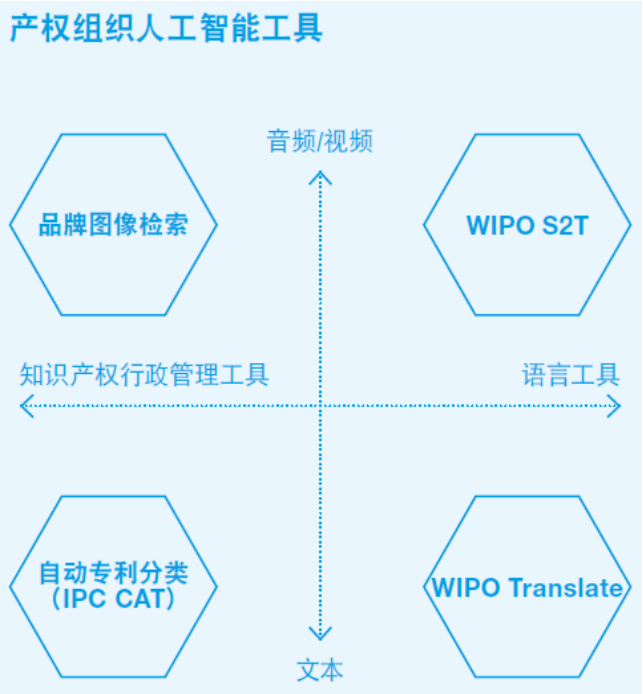
产权组织“深度监督”机器学习历史

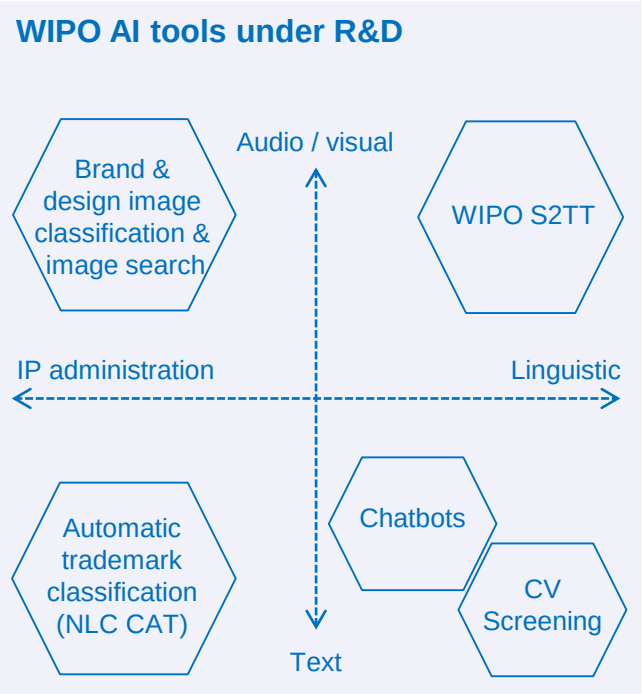
- 随着产权组织生成和收集的数据增加，我们从数据中获取价值的效率也必须提高——人工智能在其中发挥着重要作用！
- 由于各体系和部门的翻译需求不断增加，于2011年推出了统计机器翻译（SMT）工具——TAPTA
- 2016年推出了神经机器翻译（NMT）工具——WIPO Translate，以提高质量（仅比谷歌首次发布NMT晚5天）
- 产权组织自2016年以来，在图形近似性、语言处理等方面进一步使用人工智能，逐渐成为强大的人工智能市场参与者

WIPO FOR OFFICIAL USE ONLY

WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

产权组织积极发展人工智能内容提要





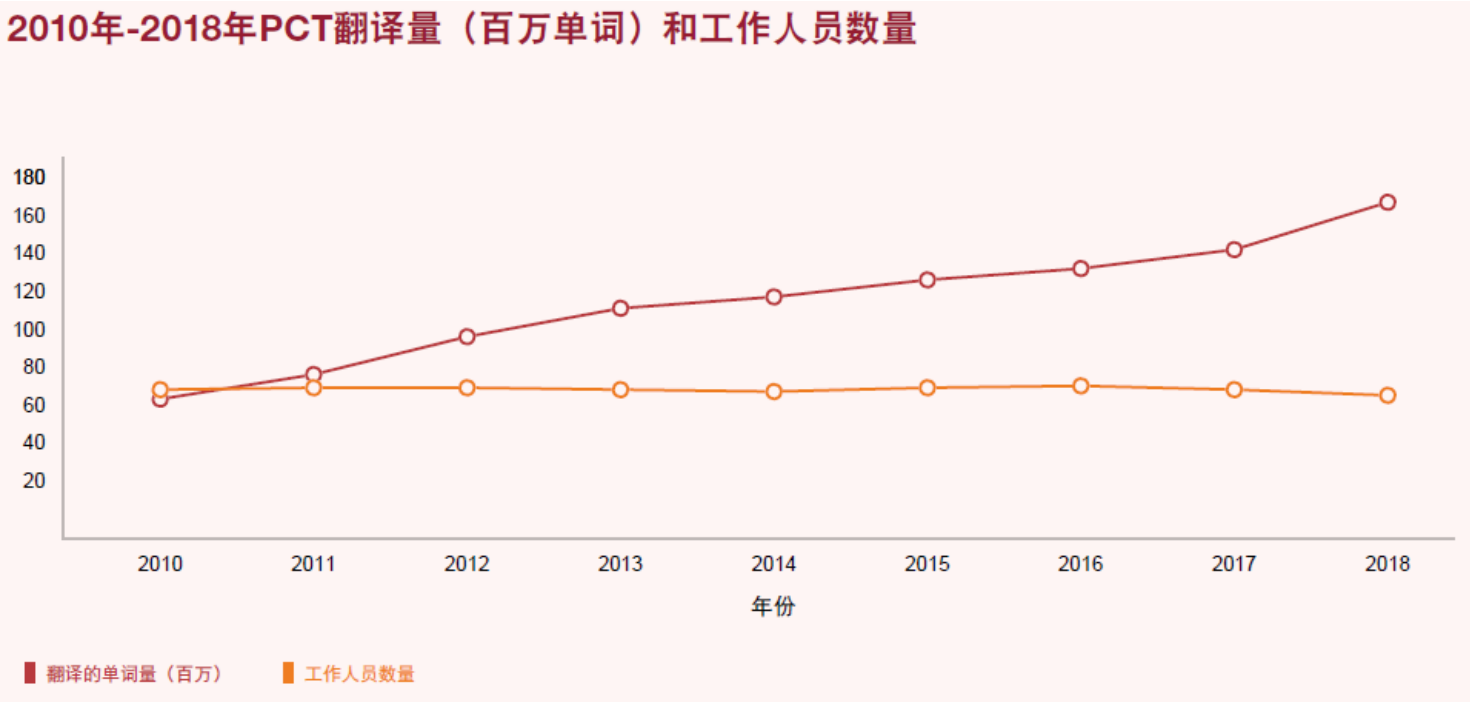
- 商标图形近似性检索
- NMT翻译和语音转文字(S2T)
- IPC文本分类（IPC CAT）

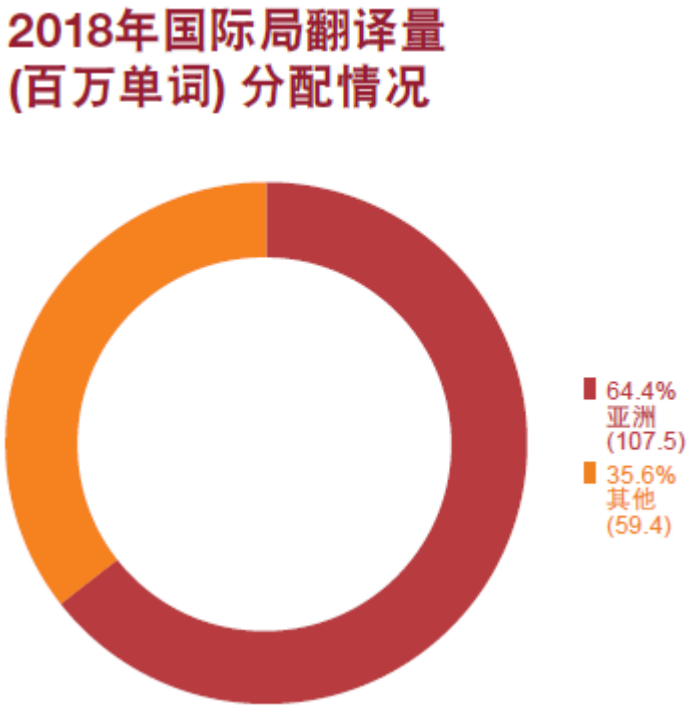
- 维也纳分类器与外观设计图形检索
- 语音转译文（S2TT）
- 尼斯文本分类（NLC CAT）
- 其他近似性

WIPO FOR OFFICIAL USE ONLY

WIPO Translate用于《专利合作条约》（PCT）

- PCT已用产权组织机器翻译工具*处理了超过10亿词的专利文本





*2011年推出了产权组织统计机器翻译工具（TAPTA），2016年推出了WIPO Translate（神经机器翻译工具）
来源：总干事提交2019年产权组织成员国大会的报告

WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

WIPO FOR OFFICIAL USE ONLY

WIPO Translate用于PATENTSCOPE

~7,000万 (2019年6月)

+159%

~2,700万 (2018年6月)

~100万 (2014年)

- 仅2018年内，翻译量估计为80亿字符
- 2020年初开始，翻译量达到300万词/天
- 65%是中文、日文和韩文

WIPO FOR OFFICIAL USE ONLY

WIPO WORLD INTELLECTUAL PROPERTY ORGANIZATION

WIPO Translate: 关键信息

- WIPO Translate建立在内部软件的基础上，但使用MarianNMT[1]开源技术
- 表现优于其他工具，包括市场领先的基于NMT的商业服务
- 关键优势：
  - 产权组织数据驱动的方法*：经过大量知识产权相关文本的高度训练
  - 注重引擎的量身定制，以适应特定领域
- 目前正在调整，以用于诸多非知识产权领域（例如供其他联合国机构使用）

https://www.wipo.int/wipo-translate/zh/index.html

[1] https://marian-nmt.github.io/

* 更多信息见附录。

WIPO FOR OFFICIAL USE ONLY

WIPO WORLD INTELLECTUAL PROPERTY ORGANIZATION

商标图形近似性检索——全球品牌数据库

选择一项检索策略:

1. **概念**
2. **形状**
3. 颜色
4. 组合

~ 4,000万
商标

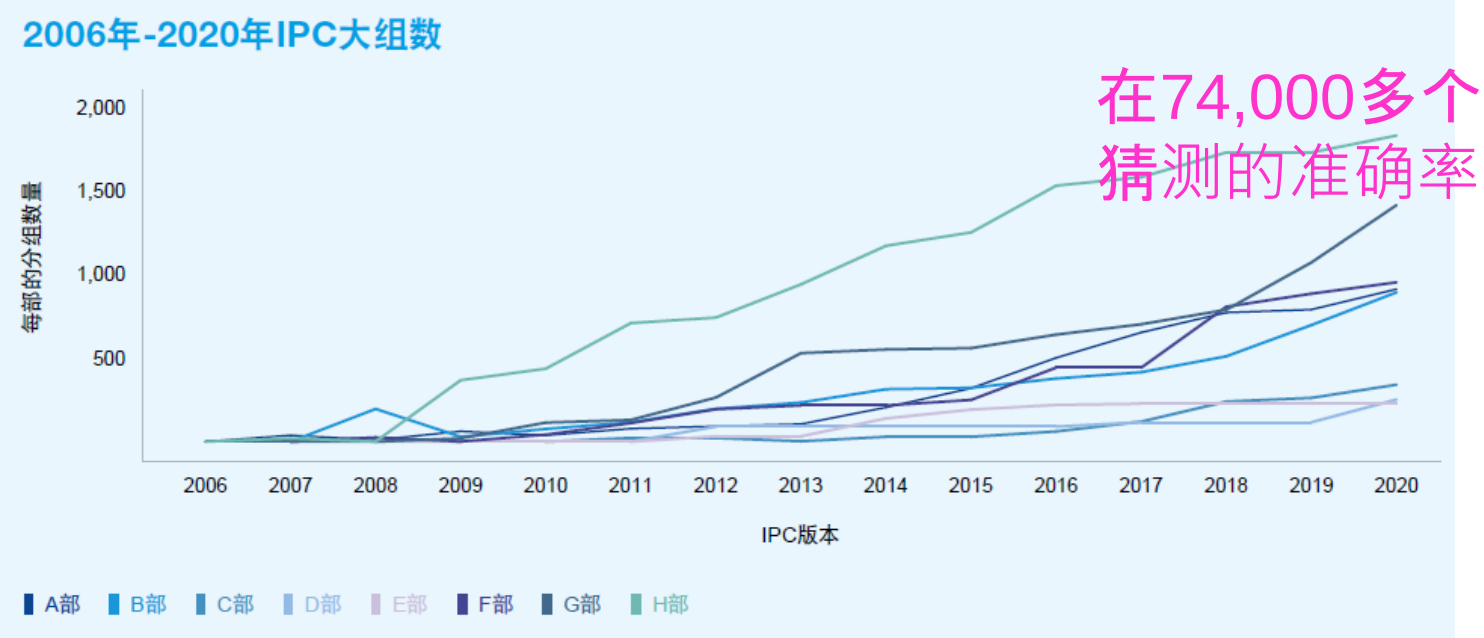
- 2019年4月开发出最佳此类人工智能算法并投入使用
- **每月超过6,000次基于**图形的检索
- 为许可目的提供应用程序接口

>> 亲自试用该检索工具<<

WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

WIPO FOR OFFICIAL USE ONLY

自动为专利分配IPC分类号（IPCCAT）



在74,000多个分类号中，前三个IPC猜测的准确率达到84%。

https://www.wipo.int/classifications/ipc/ipcpub/?notion=search&searchmode=ipccat

- IPCCAT-neural是产权组织基于人工智能的文本分类工具
- 自动在小组一级进行IPC分类（支持10种语言）
- **其他用途**：对科技相关文献进行分类

WIPO FOR OFFICIAL USE ONLY

WIPO Speech-to-Text: 音频/视频转录工具

- 内容与文本对齐
- 工具内文本编辑

Set playback speed: 2x 1.5x 1x

in the pre-grant and post-grant review, effective buffers and preven

patent-based incentives and necessity

With the drilling force for triggering drug research and development

developing countries.

The TRIPS Agreement provides these countries with policy space.

- 强大的多语言检索工具
- 可与WIPO Translate结合使用

WIPO Speech-to-Text search

trips | English | SCP31

Search WIPO meeting transcripts

Search results

Text snippet result	Conference	Date	Media timestamp
The TRIPS Agreement provides these countries with policy space.	SCP31	2019-12-02	56 minutes, 20 seconds
In spite of different national laws the flexibilities enshrined in the TRIPS Agreement.	SCP31	2019-12-02	2 hours, 16 minutes, 23 seconds
And especially LDCs on the understanding and use of flexibilities contained in the TRIPS Agreement.	SCP31	2019-12-02	2 hours, 31 minutes, 44 seconds
The effective use of compulsory licensing within TRIPS flexibilities has.	SCP31	2019-12-02	2 hours, 42 minutes, 57 seconds
TRIPS flexibilities.	SCP31	2019-12-02	2 hours, 31 minutes, 59 seconds
Of the TRIPS Agreement which is already in force in Brazil and this will be until Thirty first December	SCP31	2019-12-02	2 hours, 38 minutes, 7 seconds
And the mechanisms and the use of the flexible possibilities of the patent system are set out in TRIPS	SCP31	2019-12-02	23 minutes, 33 seconds
The major international patent instruments including the Paris Convention and the third TRIPS Agreement	SCP31	2019-12-02	2 hours, 30 minutes, 28 seconds
We find that this particular exception should be used within the rules provided in the TRIPS Agreement	SCP31	2019-12-02	2 hours, 37 minutes, 36 seconds
The significance of TRIPS flexibilities in the design of domestic patent laws is critical and provides	SCP31	2019-12-02	54 minutes, 47 seconds

WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

WIPO FOR OFFICIAL USE ONLY

WIPO Speech-to-Text：关键信息

- 基于产权组织会议，内部开发机器学习的例子
- 目前，产权组织在两个系列会议中使用该工具，但有望于2021年在产权组织所有会议中由此取代人工转录
- 为生成和翻译逐字记录报告节约可观预算
- 关键优势：
  - 产权组织数据驱动的方法*：经过大量知识产权相关文本的高度训练
  - 在准确转录口音浓重的英文发言方面表现出色
- 正在进行调整并许可诸多国际组织和机构使用

>> 观看使用中的WIPO S2T <<

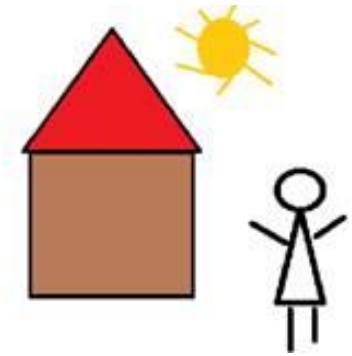
* 更多信息见附录

WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

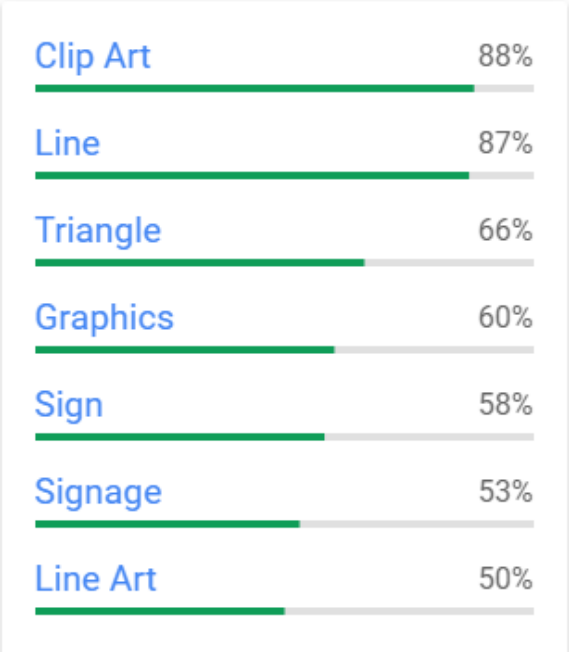
WIPO FOR OFFICIAL USE ONLY

产权组织人工智能与行业标准相比如何？

- 产权组织人工智能算法比**典型的商**业服务**更准确**（**例如**图形分类）



谷歌版本的人工智能“标签”



产权组织的图形分类工具

07.01.04	独立式住宅	12.891
07.01.06	其他形式的住宅	11.684
01.05.04	有阳光的太阳	11.08
04.07.03	代表一个人的几何图形或几何图形组合	10.558
01.05.25	太阳的其他表现形式	10.538
02.05.02	儿童的剪影或轮廓	9.63

WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

WIPO FOR OFFICIAL USE ONLY

向外拓展产权组织的人工智能足迹

- 产权组织注意到不断扩大的全球数字鸿沟，因此致力于提供对产权组织人工智能工具的包容性获取
- 为收回其人工智能工具的投资和运营成本，产权组织正在以合理的条款为其人工智能工具发放许可（例如向科技界和其他联合国组织）
- 迄今为止，产权组织已将其人工智能工具许可给17个不同的外部实体，另有31个正在积极进行谈判和讨论

WIPO FOR OFFICIAL USE ONLY

WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

谢谢！

联络方式：claudio.cocorocchia@wipo.int

用于知识产权行政管理的人工智能工具

由产权组织人工智能驱动

附录

1. WIPO Translate和语音转文字的单张幻灯片介绍
2. 产权组织数据驱动的方法
3. 产权组织语音转文字的主要使用案例

WIPO Translate：神经机器翻译工具

- 内部开发，使用MarianNMT[1] 开源软件
- 翻译文本和完整文档，并包含在线小工具和计算机辅助翻译（CAT）工具插件
- 提供不同的“现成”领域版本：例如联合国逐字记录和专利
- 每个版本都处理了数十亿字符的文本，因此高度精确
- 表现持续优于市场上领先的以神经机器翻译为基础的商业服务
- 关键优势：
  - 科技文本和文献的翻译
  - 产权组织数据驱动的方法*：经过大量知识产权相关文本的高度训练
  - 注重引擎的量身定制，以适应特定领域
- 为诸多国际组织和机构使用进行量身定制并发放许可

https://www.wipo.int/wipo-translate/zh/index.html

[1] https://marian-nmt.github.io/

WIPO FOR OFFICIAL USE ONLY

WIPO Speech-to-Text: 音频/视频转录工具

- 基于产权组织会议和大会数据，于内部开发
- 产权组织试点，但有望在2021年取代产权组织会议人工转录
- 音频/视频与文本对齐；工具内多语言全文检索与编辑
- 关键优势：
  - 速度：5分钟内处理1小时的音频/视频
  - 产权组织数据驱动的方法：经过大量知识产权相关文本的高度训练
  - 在准确转录口音浓重的英文发言方面表现出色
- 为诸多国际组织和机构使用进行量身定制并发放许可
- 研发中：直播字幕以及与WIPO Translate的全面融合（即刻生成原文的多语言版本）

>> 观看使用中的WIPO S2T <<

WIPO FOR OFFICIAL USE ONLY

WIPO WORLD INTELLECTUAL PROPERTY ORGANIZATION

产权组织数据驱动的方法

- 人工智能应涉及深度“监督” 机器学习，其中人类的专业知识和干预是关键——数据质量是产权组织方法的基础
- 数据量也很关键：没有大量数据，人工智能是无法学习和改进的
- 产权组织强大的GPU：大数据集需要强大的处理能力
- 在产权组织，定期训练（至少每年一次）或“维护”十分重要
- 领域重点：使用合作伙伴的数据进行训练有很多好处，并且产权组织团队是定制软件实例的专家
- 产权组织以快速取胜为目标，同时接受快速失败——迭代方法
- 追求100%准确不是可持续的策略，并且可能会浪费时间和精力——产权组织仍然坚持最高KPI

WIPO FOR OFFICIAL USE ONLY

WIPO Speech-to-Text：使用案例

- 快速交付会议记录
- 存档知识发现（通过检索功能）
- 改进内容可获取性（隐藏式字幕）

WIPO FOR OFFICIAL USE ONLY

WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION