NEW ERA OF CLASSIFICATION

ABDER MOUMEN, GARETH LORD, SALVATORE CHIARIZIA, UWE BOLTE | 20. JUNE 2024

NEW ERA OF CLASSIFICATION: WHY

- Massive increase of documentation with big developing economies as e.g. China and India contributing more and more patent and non-patent literature. 1 260 000 newly received documents were classified in 2023 alone.

- Time investment for intellectual classification of this rapidly growing annual number of new documents will exceed our capacities in view of time and money

- Consistent AI-based or AI-supported classification
  - will allow us to cope with full classification of the documentation also in future
  - will deliver a consistent classification picture

NEW ERA OF CLASSIFICATION: IMPLEMENTED PREVIOUSLY

- Promoting Clasma fields to US proficient
  - May 2023, 105 fields, 39.000 visits/year, 650 days/year saved
  - July 2023, 63 fields, 30.000 visits/year, 500 days/year saved
  - Dec 2023, 98 fields, 28.000 visits/year, 460 days/year saved

  Total 1610 days/year, app. 9 FTEs

- Auto circulation closing
  - Pilot project in Dir 1101, 1108, 1203 and 1211 [Aug 2023, 8.750 visits, 145 days one-off]
  - Clasma fields of Dir 1103, 1203, 1211 and field 760 in Dir 1014
    [April 2024, 79 fields, 7638 visits, 128 days, app. 30.000 visits/year, app. 500 days/year]
  - Clasma fields with high confidence level (>60%) for AI proposed symbols
    [May 2024, 54 fields, 4527 visits, 75 days, app. 18.000 visits/year, app. 300 days/year]

AI-BASED AUTO CIRCULATION CLOSING: PRINCIPLE

- Unreviewed CPC symbols given by other offices or non-authorized EPO examiners (incl. CPC symbols derived from IPC for non-EN language WO documents) **for a Clasma Field** are compared against AI suggestions
- If **all** the above intellectually allocated symbols match with AI symbols:
  - they are confirmed
  - all AI generated symbols for this Clasma Field with a confidence level above a defined threshold (currently set at 40%) are added to the document as confirmed symbols
  - the visit is closed
- These symbols appear as Auto-AI in the confirmed column of Canopée
- If **none or only some** of the intellectually allocated symbols match with AI symbols:
  - The visit is maintained and needs to be done by intellectual classifiers

AI-BASED AUTO CIRCULATION CLOSING: IMPLEMENTATION

From the 837 Clasma fields where visits can be created, 67 fields were excluded where either C-sets or special indexing (ALLOYS, SADIQ) are applied.

The AI-based auto circulation closing was implemented for 770 Clasma fields with effect of 14-06-2024.

During the pilot (Aug 2023) about 40% of visits were closed by the auto circulation closing approach. Accordingly, the classification workload for fields now affected by auto circulation closing is expected to be reduced by roughly 40%.

US PROFICIENCY: PRINCIPLE

- US proficiency: For US proficient Clasma fields, no circulations for documents classified by the USPTO are created.
- For the best 70% Clasma fields ranked by CPC-QA compliance, the respective fields were set to US proficient (indicator for the quality of the CPC symbols allocated by the USPTO).
- For these Clasma fields CPC symbols allocated to documents by the USPTO are endorsed by copying them to confirmed symbols by EPO.
- For these Clasma fields for each US document all AI generated symbols with a confidence level above a defined threshold (currently set at 40%) are allocated to the document as confirmed symbols and appear as AI-Classif. This process is called US booster.

US PROFICIENCY: IMPLEMENTATION

- In total 571 Clasma fields (~70%) were set US proficient with 266 fields (~30%) remaining open.
- The US proficiency setting was implemented with effect of 14-06-2024.

NEW ERA OF CLASSIFICATION: NEXT STEPS

- Review of the Class-OQC framework, aiming at
  - increasing the input from quality control
  - getting faster feedback on changes in quality
  - offering structured feedback for the AI engine

- This new concept for classification is a baseline approach that needs considerable fine-tuning of the AI-engine, aiming at increasing the compliance level (completeness and correctness) of AI-generated classification symbols.

- Develop and promote adapted search approach

NEW ERA OF CLASSIFICATION: CONSEQUENCES FOR OUR WORK

- Review structure of classification
  - less granular schemes
  - reduced deep-indexing
  - focussing on invention classes

- Adapt the search process
  - fully exploit ANSERA features
  - use the ranking mechanism
  - use search terms
  - get away from mere crossing of classes

NEW ERA OF CLASSIFICATION: CHANGE MANAGEMENT

- **Communication to DG1 operations directors** [COO + implementation team]
  (with this presentation)
  Give detailed description of the implemented changes and further course of action

- **Communication to TMs, Gérants and classifiers** [Line directors]
  (in dedicated meetings **within 1 week** from today)
  Create a thorough understanding and acceptance of the essential need to go this route in classification for sustainability purposes. Give details on the implemented measures and the consequences in everyday operations (classification and search)

- **Specific ANSERA training – by tailor-made courses** [Horizontal team Learning & Development]
  (remote and classroom training)
  In dedicated ANSERA trainings for examiners offer adapted search approaches for high quality search results based on new classification approach

NEW ERA OF CLASSIFICATION: SUPPORTING DOCUMENTS

- US proficient fields
- US non-proficient fields
- Auto circulation closing fields
- Non-auto circulation closing fields

- US proficiency manual
- Auto circulation closing manual