AI4media

Imagga Technologies

AI Technology in Image & Video Organisation

White paper - October 2022

Executive Summary

The automated organisation of large media collections has been fueled by the rapid development of Machine Learning (ML) algorithms that constantly improve accuracy and speed

Just ten years ago, image recognition technology could detect and classify only basic objects and shapes. **Thanks to deep learning advancements, image recognition algorithms for instant detection of all types of visual content are now a reality.**

This white paper is part of AI4Media's effort to align research with the industrial needs of media companies. Surveys are used to gather input from media companies and other organisations, aiming to provide a clearer picture of industry needs and understand current and future trends.

The goal of Imagga's online survey was to identify (i) the current level of adoption of automated image and video (re) organisation tools, (ii) the potential impact of the adoption of such novel tools, and who will benefit the most, and (iii) the most desired features such tools need to offer to be appealing and easily adopted by the stakeholders of media organisations.

Media companies have realised the importance of implementing AI-enhanced image and video (re)organisation technologies but have lagged in implementing such technologies due to the high cost, inflexibility of their legacy software, and lack of experienced AI staff internally.

* Visual Content – online content that's predominantly image-based, but can also include other forms of content such as diagrams, charts, infographics, online videos, screenshots, memes and slide decks.

Key messages

- AI-enhanced automated organisation of large media collections significantly aids media companies in reducing costs and, at the same time, providing new opportunities for visual content monetisation.
- Media companies have realised the importance of implementing AI-enhanced image and video (re)organisation technologies but have lagged in implementing such technologies as part of their workflows.
- Creating an easy-to-integrate demonstrator in the scope of AI4Media can showcase and help media companies adopt such technologies.
- Media companies lack an understanding of Trustworthy AI and its practical application in their ongoing visual content processes.

Table of contents

2

Introduction

Over the last few years, the image and video content organisation powered by Artificial Intelligence (AI) has grown exponentially. Due to the advancement of hardware and Machine Learning algorithms, visual content organisation has developed in previously unimaginable ways, breaking concepts, and widening the horizons of what's possible.

Images and videos have become essential mediums for all media organisations. The way we consume media has shifted dramatically during the last decade. Most media consumption has moved from traditional mediums like newspapers, magazines, and TV, to mobile-first mediums like social media platforms and, recently, the metaverse. This shift is putting significant stress on media companies to adapt and provide content that matches the new reality and demands.

Media companies in Europe and worldwide have accumulated vast digital archives and collections of images and videos over the years. Since these collections have been gradually and iteratively built over long periods of time, often by different departments and units of media companies, they are usually not well organised, having little to no metadata, such as tags, categories, and other types of annotations.

The automated organisation of large media collections has been fueled by the rapid development of machine learning algorithms that constantly improve accuracy and speed. In the last 10 years, the advancement of deep learning and image recognition technologies made it possible for almost real-time analysis of various forms of visual content. Particularly for video recognition, the incremental development in hardware able to process graphic content with lighting speeds opened new opportunities for advanced searchability and organization of media content.

Due to the lack of coherent media asset organisation tailored to the media company's business and services, utilising and monetising the ever-growing image and video assets is becoming a bothersome and expensive endeavour. In addition, both big traditional media companies and, more so, the new digital media platforms combine in their collections both media content created by these companies and their partners, but increasingly

also user-generated content (UGC). Such hybrid media archives need advanced content organisation solutions, require working in (near) real-time to safeguard viewers, meet laws and regulatory requirements, but also provide monetisation opportunities for new products based on already existing digital content.

AI4Media is an EU co-funded research initiative with 30 technology and media partners aiming to explore diverse aspects of AI in the media sector, advance AI technology research, and develop specific solutions for seven use cases. Led by Imagga, one of these use cases focuses on the development of an automated visual content (re)organisation demonstrator to address the above-described user needs and business problems. The demonstrator aims to improve novel methodologies based on advanced deep learning techniques, such as CNNs and RNNs, aimed at photo (re) organisation - tagging, categorisation, and facial recognition. The automatic tagging technology provides easy-to-use search functionality for photos and videos using keywords typed by the user. The categorisation technology provides a way to structure the photos in a top-level ontology of categories suitable for editorial and user generated photos. Finally, facial recognition automatically groups the people in the photos, enabling the search for specific images related to a particular person.

This white paper is part of AI4Media's effort to align AI research with the industrial needs of media companies, bring a better understanding of the problems and challenges in the sector and suggest novel technology solutions that might turn into growth engines. Surveys are used as a mechanism to gather input from media companies and other organisations aiming to paint a clearer picture of industry needs and understand trends.

3

Problems and Challenges

Automated image and video content organisation offers essential new capabilities for businesses of different venues that need effective digital content management. While holding great potential and already showing impressive results, there are challenges that AI-powered image and video content organisation faces.

Industry/Organisational Issues

The result of our survey shows that just 7% of the respondents state that they use automated solutions for image and video content organisation (see the section on survey results). Most media companies haven't implemented any form of AI-enabled automation regarding visual content and are purely dependent on human efforts to catalogue and organise content.

The slow adoption of new technologies is partly due to legacy software that's impossible or too expensive to adopt AI-based capabilities. Investment in such technological improvements might be challenging for many media companies due to the content monetisation shift enforced by search engines and the new social media platforms.

The lack of AI experts in the media industry prevents the penetration of AI-based image and video content organisations. For any machine learning model to be trained appropriately, a significant amount of well-organised and properly tagged visual content is required. The lack of a deeper understanding of the basic requirements for data collection and annotation for training AI models can be a significant stopper in even initiating the first steps in adopting automated solutions for image and video content organisation.

Technological Issues

Implementation of in-house AI-powered image and video processing solutions requires investment in powerful GPU (graphic processing unit - necessary for training deep-learning models in a time-efficient manner) hardware. Adoption of cloud services where the setup and scaling of the needed GPU are less economically burdensome is a great opportunity but often creates issues with safety and privacy of the visual information that might be shared on the cloud and possibly exposed to third parties.

Misunderstanding the capabilities of automated image and video organisation is another major obstacle to be addressed. Non-technical staff at media organisations often have unrealistically high expectations that exceed the current capabilities of AI-based automated image and video organisation. Image recognition technology has advanced significantly in recent years but one of the major issues with which automated content moderation is still struggling is recognizing context. Machine learning algorithms can find it difficult to differentiate between subtle cultural and social trends and phenomena. For example, if the algorithm is set to detect all nudity, this is what it would do — even if the nudity is related to art or important news pieces. A prominent example was the case from 2016 when Facebook removed the photo of the iconic Vietnamese 'napalm girl' who is naked.

Another important challenge AI technologies for image and video organisation platforms need to overcome is **multilingual support**. While they are getting better at it, there are still obstacles in the way. The process is not only about acknowledging the direct meaning of words and phrases, but also the social and cultural connotations that may make them offensive or inappropriate.

Video content, in particular, might be more challenging to index and organise due to the large size of the data that needs to be processed. Applying AI to each frame of the video generates high platform costs. A fast and accurate AI must be developed to overcome these hurdles to the **efficient and cost-effective organisation of live streaming and videos.**

Ethical Issues

Compliance with legal and regulatory requirements specific to a given country or industry might be a significant adoption barrier for AI-based automated image and video organisation solutions. The use of third-party providers, data of unclear origin, and an undocumented initial data collection process for the AI model might be of concern for some media companies in the public but also in the private sector.

Controversial AI models that tackle sensitive, unsafe or inappropriate content might be a challenge for media companies, exposing their staff to content that can create mental and psychological problems (images of war, violence, banned substances, etc.)

Last but not least, the lack of understanding, standardization, and assessment of the **trustworthiness of AI systems** is a significant challenge for the wider adoption of automated AI image and video organisation solutions. The problems of technology bias, transparency, and adaptation to legal and social norms are issues that need to be further addressed.

4

Industrial Needs

The main motivation for Imagga's AI4Media use case "Automated (re) organisation of large media collections of photos and video," is that almost all media companies, big and small, need to organise better the visual content they create, receive, and generally must deal with.

Such visual content includes both their photo or/ and video collections, or content they receive from or send to partners or suppliers/clients, as well as UGC (user-generated content) such as photos and videos provided by end users.

All media companies organise their visual content somehow, mostly using manual practices for tagging and categorization. However, this may not be optimal for search and exploration, so (re)organising media content archives is of significant importance. New methods and tools to do this automatically are needed. This requires understanding the current media content collection organisation, designing a new structure for reorganising it, and mapping the current content themes, categories, etc., to the new ones.

Media companies need to quickly and accurately find and retrieve specific content they need to serve their clients, i.e., by searching and finding photos and videos belonging to specific topics, themes, or categories or containing particular people, objects, or keywords/tags. It is also a common need for most media company content teams and operators to filter such visual content by source, author, license/rights, and whether it contains personal information.

Media content teams often search their media collections to repurpose or generate new content. Turns out a lot of the content media companies need on a daily basis to illustrate the news, events and facts have already been created and paid for, but they lay buried in a folder on a server or computer in the corporate network without proper metadata. After being used, few media companies have the tools and the time to describe, keyword and archive the materials properly. If image and video organisation is done properly, this will allow the easy discoverability of needed content and, as result, will save time, money, and workforce resources.

Finally, organising visual content to support better search and smarter exploration and discovery are critical for improving the effectiveness and monetisation opportunities for visual content that often can be used beyond the actual event for which this content was created.

The novel, creative methods for image and video content (re)organisation are critical for the adaptation and survival of media companies in the era of user-generated content. Due to the wide spread of "free", easily accessible citizen journalism and user generated content that's easy to appear on social media, media companies are struggling to keep their revenue streams from traditional print media and more recent business models as online subscriptions. Sending out big teams to create quality content is economically not viable, as freelancers with cheaper equipment and less need of assisting technical team can create relatively high-quality content. Everyone with a powerful, high-quality camera phone and good storytelling skills has become a treat for the professional guild. The way media companies create content needs to be adapted to the new reality. Because of the high cost of professional content creation, there's need for it to be re-purposed and re-used so it's economically viable. Investing in AI-enhanced tools for image and video re-organisation will make media companies better prepared for the future and the challenges it holds.

5

Survey Results

To further understand the need for novel approaches in visual content organisation and retrieval, we conducted a survey to collect firsthand opinions from those who work in media organisations or are providing services for such organisations.

The online questionnaire was distributed among Imagga's media clients and the AI4Media consortium and associate members.

With this online survey we aim to identify (i) the current level of adoption of automated image and video (re)organisation tools, (ii) the potential impact of the adoption of such novel tools, and who will benefit the most, and (iii) the most desired features such tools need to offer to in order to be appealing and easily adopted by stakeholders in the media organisations.

41 responses were gathered, the majority of which were from media companies but also organisations doing research and providing support for the media industry.

The pool of specialists we asked for collaboration was quite diverse - 25% of all respondents are photographers working in media companies, 25% are software developers, and the rest is a mix of digital asset managers, product managers, researchers, journalists, and technical support staff.

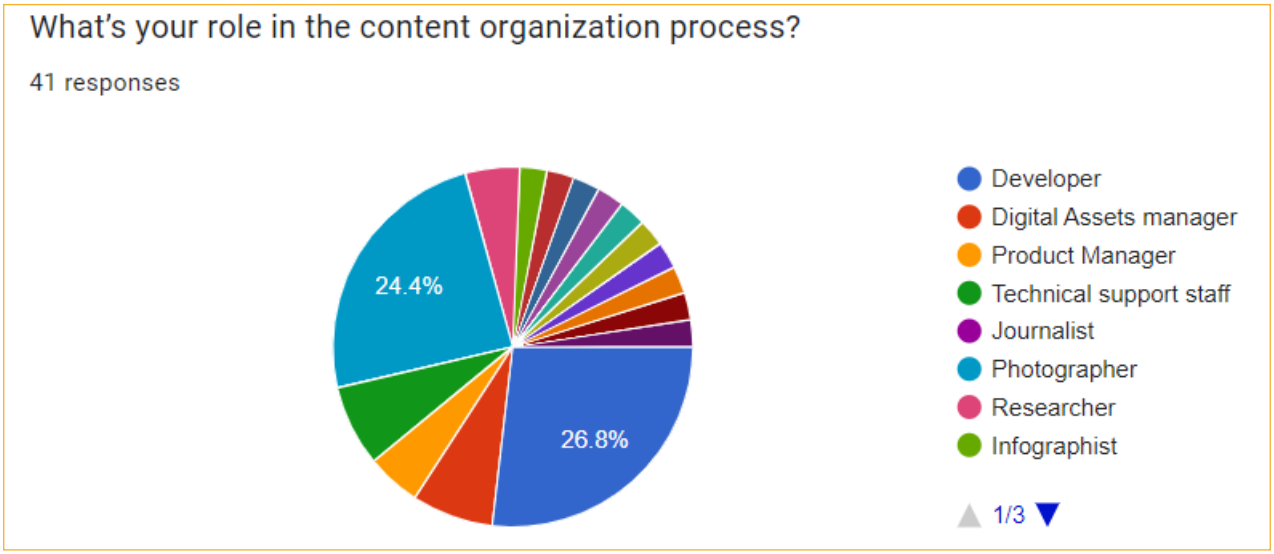

Figure 1: Role of survey participants with regard to the content organisation process

Use of Automated Tools for Visual Content Management

When asked how media companies currently manage their image and video collections, 41% of respondents say that they still rely on legacy organisational methods as personal computer programs for a visual organization (Mac Photos, etc.), mixed together with digital asset management systems of different sorts.

Only 7% of the respondents state they currently use automated image and video organisation solutions, and more than 51% have no automation whatsoever.

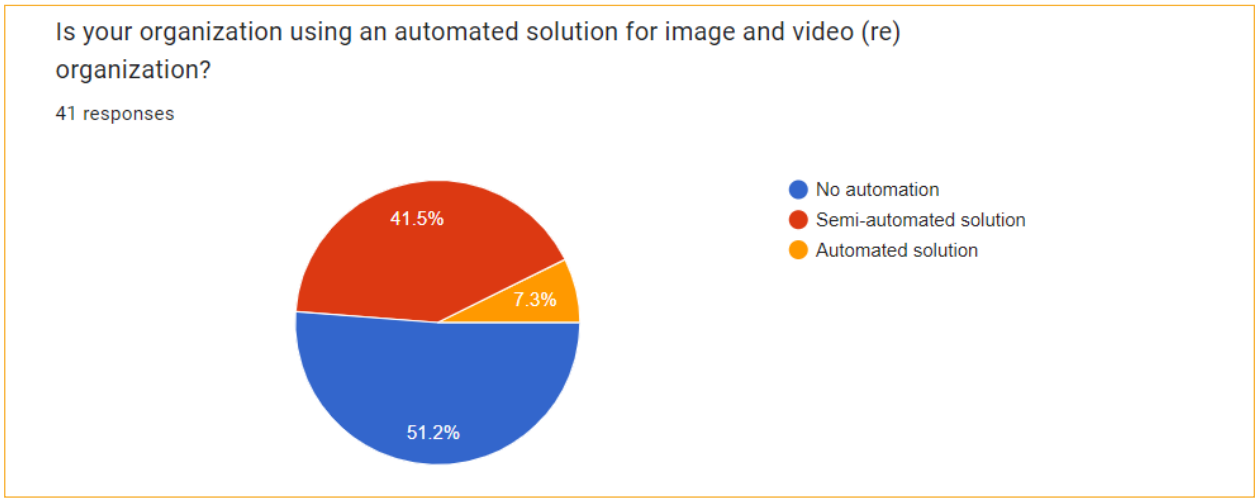

Figure 2: Current adoption of automated solutions for image and video organisation by media companies taking part in the survey

Biggest Challenges of Visual Content Management

The biggest challenge identified by the survey respondents is being able to search and discover image and video content that their organisation already has. Annotating images and videos seems to be another burdensome problem, as manual annotation requires specific linguistic skills and is time-consuming. Other significant issues seem to be checking for problematic and not safe for work content (NSFW), as well as managing duplicated content, permissions and rights management.

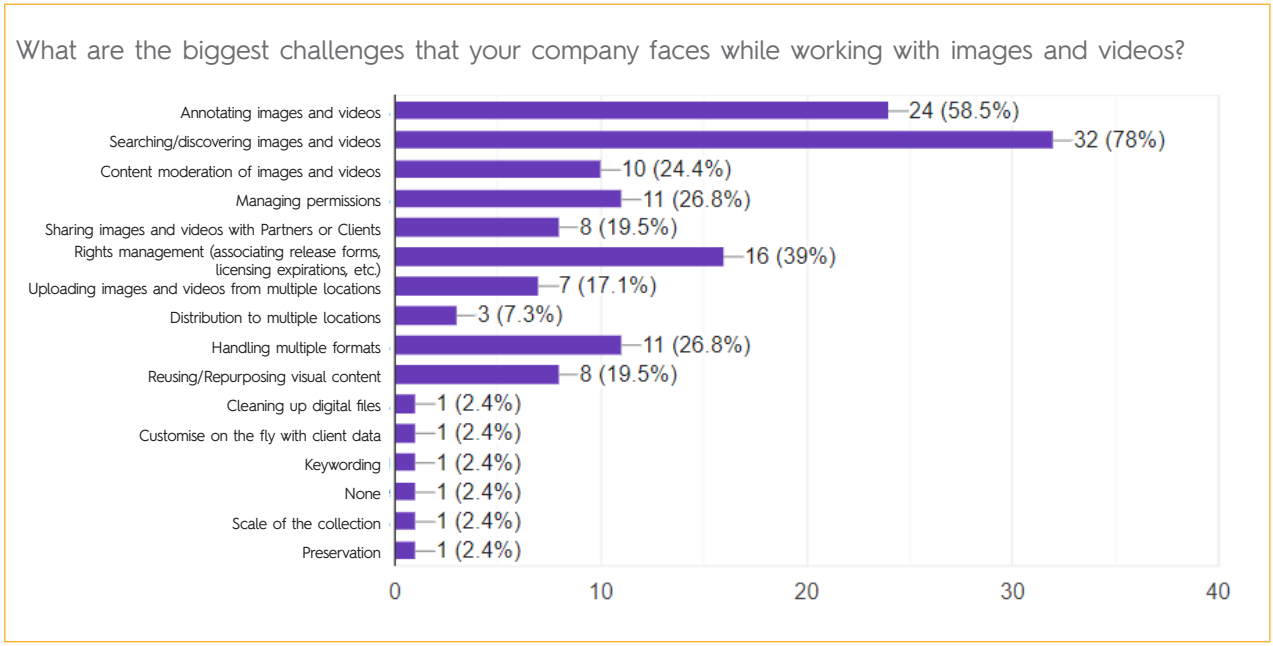

Figure 3: Challenges media companies face while working with images and videos

Potential Impact of AI Technologies on Visual Content Management Workflows

The majority of the participants firmly believe that implementing AI-enhanced image and video (re)-organisation will improve the discoverability and retrieval of images and videos (73% of respondents), will reduce the manpower needed for manual content annotation and organisation (61%), and will reduce the costs for storing images and videos within the organisation. Moreover, the introduction of automated solutions for image and video (re)-organisation will enhance the human efforts for visual asset organisation (66%), will make it possible to handle large volumes of visual content close to real-time (20%) and, on top of that, reduce the stress on people responsible for these tasks (23%).

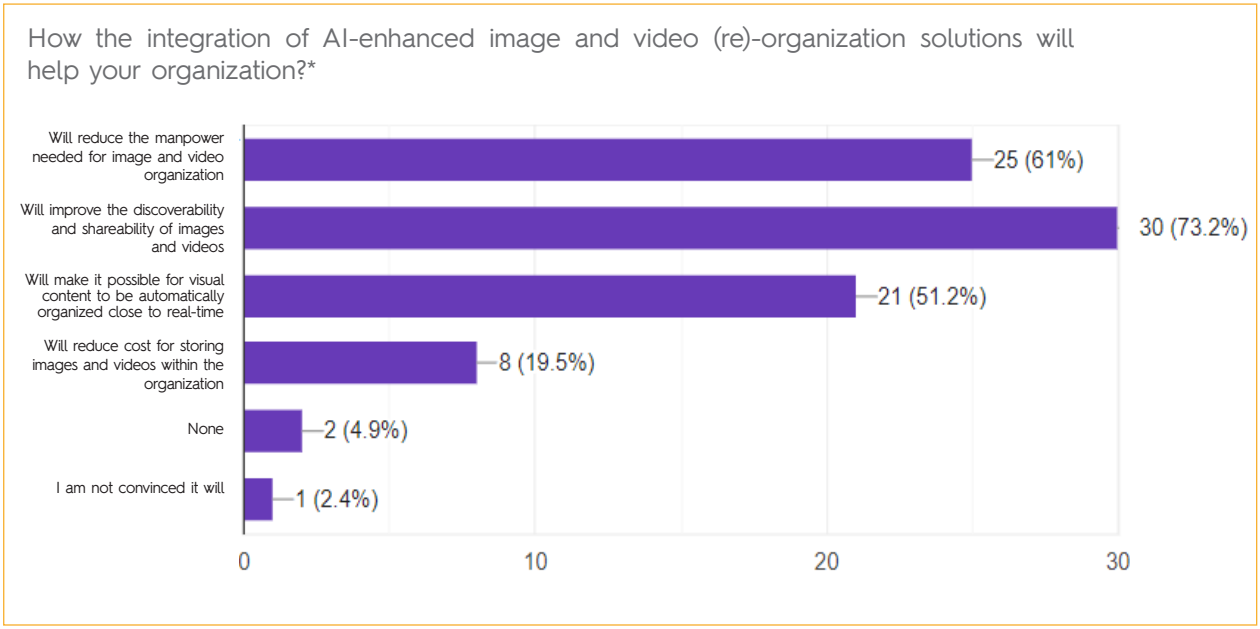

Figure 4: Impact of AI-enhanced image and video organisation solutions for media companies

The need for (re)-organisation of existing visual content is obvious as pointed out by 78% of the respondents. That is in correlation with the fact that the interviewed media companies currently do very little automation and struggle to properly annotate, classify and organise visual content. New content regularly created also requires attention, and automating these processes will greatly benefit the media organisations.

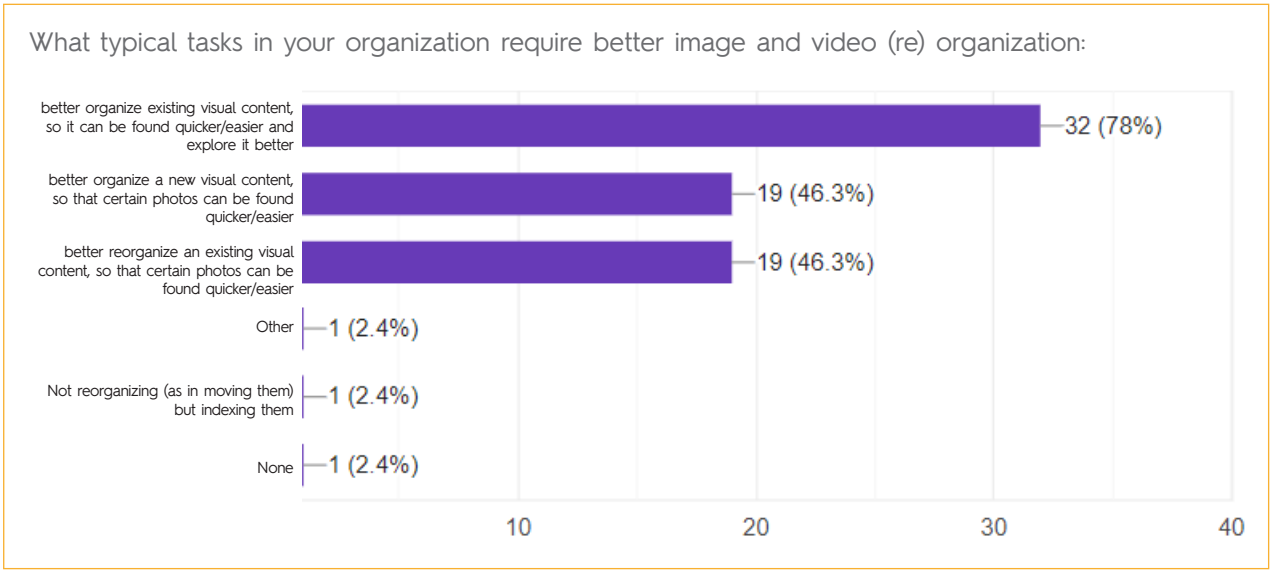

Figure 5: Typical tasks in media organisations that require better image and video organisation

Concerning the tasks within an organization that require better image and video (re) organization, according to the survey results:

A. 78% of respondents think it will improve the process of finding visual content with specific objects (object recognition is the AI technology that makes this possible),

B. 73% of respondents think it will help in finding content with a particular topic (using AI categorization that will be able to distinguish between public and political events, travel and leisure, etc.),

C. 54% of respondents think it will make the finding of people on editorial content easier (facial recognition and grouping),

D. Additionally, it will be beneficial if the implemented automated solution has features such as location-based tagging, emotion recognition, deduplication, extending support to PDF files, and others.

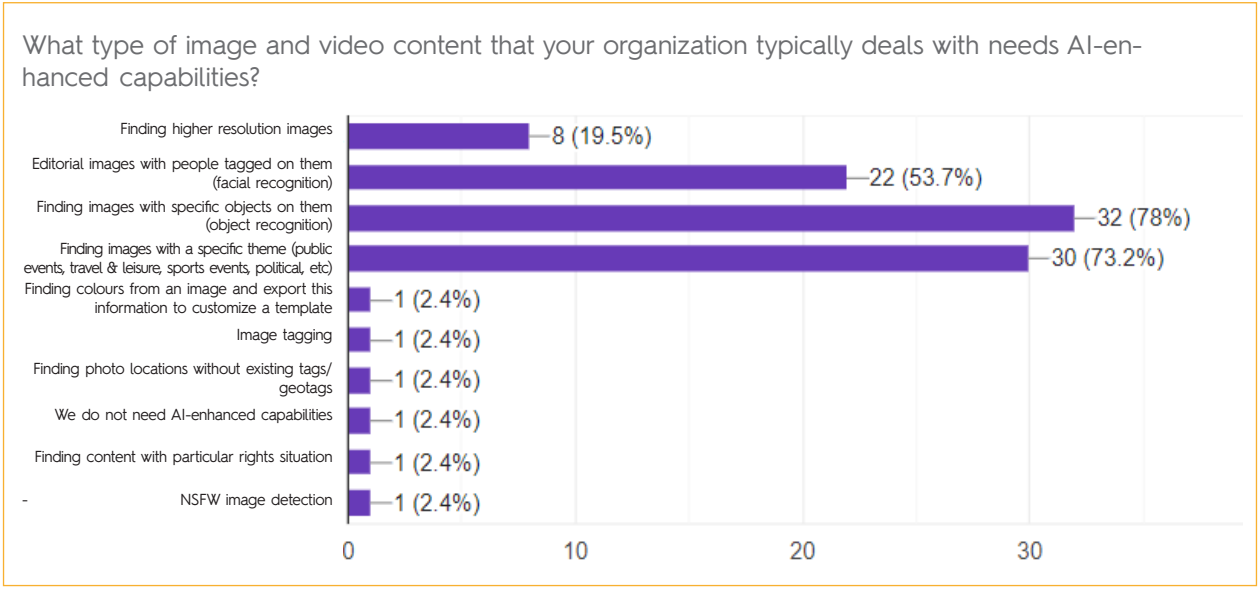

Figure 6: Types of visual content media companies deal with that need AI-enhanced capabilities

Parameters Affecting the Adoption of Automated AI Tools for Visual Content Organisation

When deciding on the practicality of using AI-enhanced tools for automatic image and video organisation, what matters most to the media companies is the precision and recall of the AI algorithms (73% of respondents), cost of implementation and ownership (70%), as well the ease of integration with any legacy software in use by the organisation. (63%).

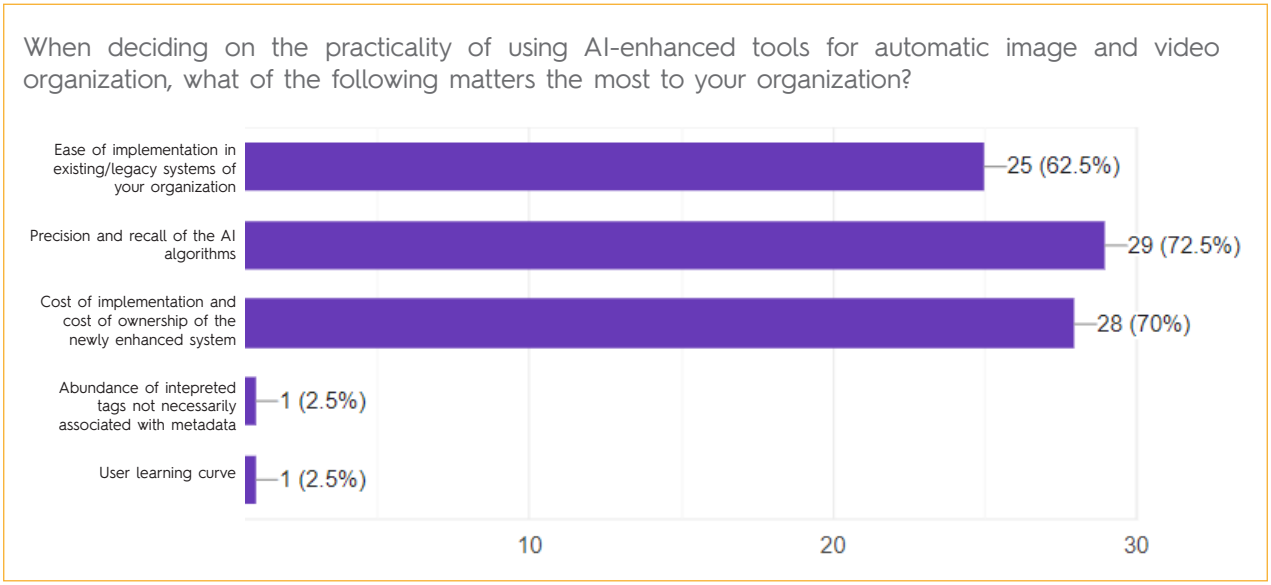

Figure 7: What matters the most when evaluating the implementation of AI-enhanced Image and video organisation

The reasons some media companies are reluctant to introduce automated organisation of images and videos include the inability of the current Digital Asset Management systems to integrate external AI tools, concerns about the cost of implementation and education requirements for the staff to handle the new functionality, concerns regarding privacy, and the need to use commercial cloud services.

Trustworthy AI for Media Content Organisation

On the topic of Trustworthy AI, only 17% of survey respondents admit they are very familiar with the concept while 32% have never heard about it.

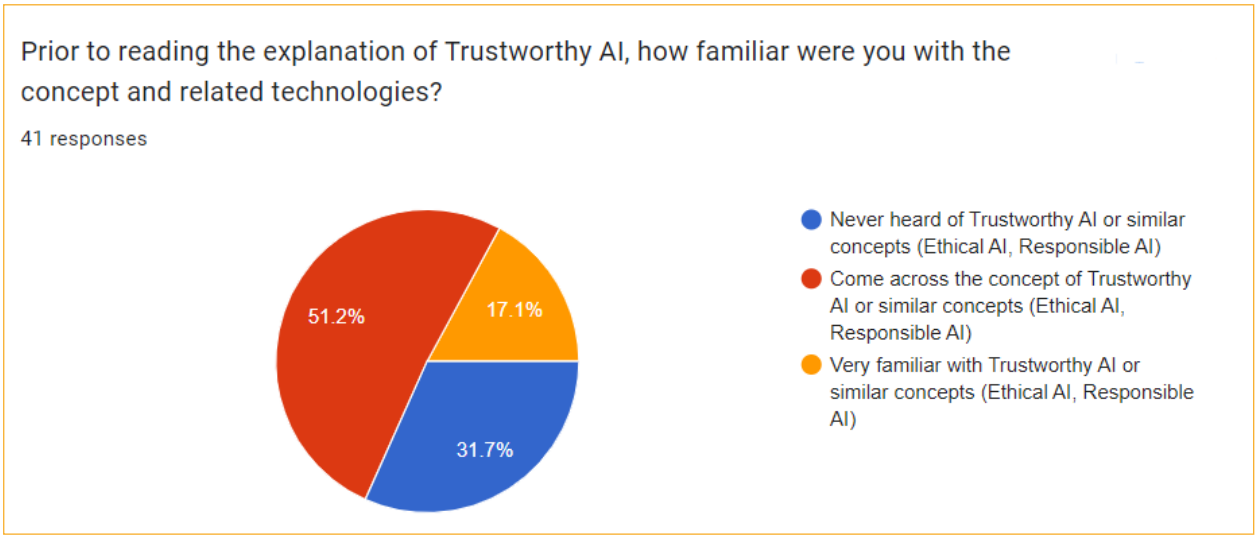

Figure 8: Familiarity with Trustworthy AI concept and related technologies

39% of respondents believe Trustworthy AI functions are essential, and 46% agree they are very useful. Unfortunately, only 7% of the respondents have come across Trustworthy AI functions integrated with actual tools or processes. Privacy protection, legal compliance, transparency, explainability, bias mitigation and robustness as features of Trustworthy AI are considered very important.

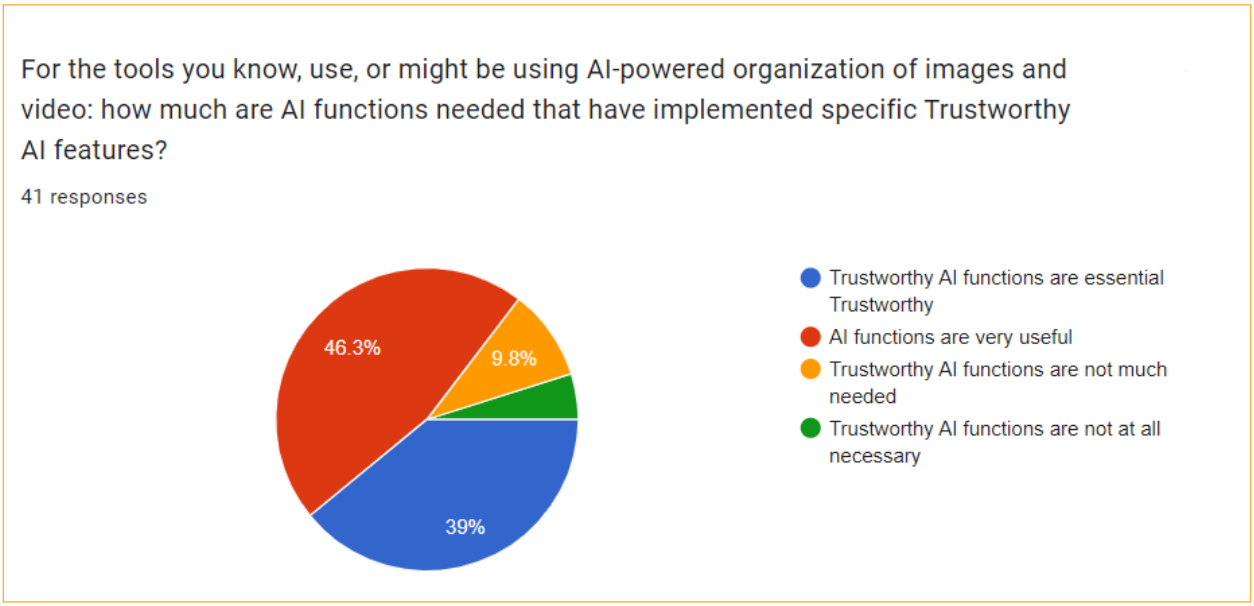

Figure 9: The need of Trustworthy AI

6

Conclusions

Media companies are well aware of the advantages that AI-enhanced image and video categorisation bring to the table. With the ever-growing volume of visual assets, the need for automation is very urgent. It will save money and human resources and create opportunities for monetisation of the visual content.

Media companies need to be able to quickly and accurately find and retrieve specific content, i.e. by searching and finding photos and videos belonging to certain topics, themes or categories, or containing specific people, objects, or keywords/tags. Being able to filter such visual content by source, author, license/rights, and if it contains personal information is also a common need for most media companies. Media content teams often search their media collections for repurposing existing content or generating new content out of it. Finally, organising visual content so that it supports not only better search, but also smarter exploration and discovery is critical. In time, a media company organises its visual content somehow, although this may not be optimal for search and exploration, so (re)organising media content archives is also important and new methods and tools to do this automatically are needed. This requires understanding the current media content collection organisation, designing a new structure for reorganising it and mapping the current content themes, categories, etc. to the new ones.

In order to address the above-described user needs and business problems, AI tools are being developed within AI4Media that aim to introduce novel methodologies based on advanced deep learning techniques such as CNNs and RNNs. An automatic tagging technology will provide an easy-to- use search functionality for photos and video using keywords typed by the user. A categorization technology will provide a way to structure the photos in a top-level ontology of categories suitable for personal photos. Last but not least, facial recognition will be able to automatically group together the people in your photos, which enables you to search specific photos related to a specific person.

Major stoppers in implementing such technologies are (i) the limitations of the legacy software used by companies, (ii) the high cost of implementing AI tools, (iii) the lack of experienced AI engineers within the media companies to ensure the proper implementation of AI-driven functionalities and integration with current processes, (iv) the complexity to comply with GDPR and various national and EU regulations on privacy and security protection.

To better serve the media sector, additional research efforts need to be done to address media company's needs: (I) better understanding of current visual content organization practices to secure smooth transition and adoption of AI automation; (ii) developing algorithms that take in consideration the political and social context under which the media companies operate; (iii) minimizing the efforts by single media company to develop, train and adopt an AI model by open sourcing or standardizing certain elements of the technology process (standardized data ontology for tagging and categorization, best practices and detailed guidelines on practical implementation of Trustworthy AI).

Media companies lack an understanding of Trustworthy AI and its practical application in their ongoing visual content processes. Significant effort will be needed to educate and help implement the concepts of Trustworthy AI into the current processes of media companies.

AI4media

Follow us @ai4mediaproject

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951911

info@ai4media.eu www.ai4media.eu