AI4media

Rai – Radiotelevisione Italiana
Centre for Research, Technological
Innovation and Experimentation (CRITS)

AI in Vision: High Quality Video Production & Content Automation

White paper - October 2022

A. Bruccoleri, R. Iacoviello, A. Messina, S. Metta, M. Montagnuolo, F. Negro

Executive Summary

Conceiving compelling scenarios for the usage of AI in the media domain is as simple as challenging is to assess to what extent those scenarios find application in the industrial practice.

This paper reports the path that RAI followed to approach this problem. In the context of the AI4Media project, RAI (the Italian public broadcaster and member of the AI4Media consortium) defines an articulated and quite dense business use case, in which the usage of AI technologies is spread to cover all the key functionalities of a typical media production workflow. The aim of this white paper is to validate the proposed vision through the collection of opinions gathered from industry experts and to drive the research community towards the most useful development directions. After a brief introduction to the use case, problems and challenges faced nowadays by industry players are tackled through the analysis of the results of a high-level survey conducted

by RAI among European stakeholders in the media domain. This is followed by an in-depth analysis of industrial needs obtained by means of a second more detailed survey addressed to RAI internal experts.

The results demonstrate that there is strong support from the industry experts to the introduction of AI in the media value chain and, at the same time, that clear indications about what to prioritise in terms of applied research and development can be derived from their feedback.

Key messages

- Media professionals strongly believe in AI technology and its application in the media domain, although it is still not clear how far we are from having fully operational and qualitatively high-level AI-driven functionalities;

- Several crucial tasks of the media value chain are not well covered by existing tools, new AI-driven tools are needed to fill this gap;

- AI technologies are seen as having the potential to offer fundamental support to the media value chain. They can give a great boost in terms of quality and creativity, but they shouldn't replace human work, which instead, according to experts, should always remain central;

- Trustworthy AI features are one of the key factors that affect the wide adoption of AI in the media chain, especially those related to Privacy Protection and Legal Compliance. The research community should push as much as possible to build trustworthy AI tools that respect user privacy and comply with relevant regulations.

To sum up

Introducing AI technologies into the production processes of a media company is not just about developing AI-driven smart tools that perform tasks, but, besides this, it involves a complex ecosystem with a delicate balance between artistic aspects, human rights to work, trustworthiness features and intellectual property of contents. Only by achieving this balance in a structural and lasting way, AI technologies will truly support human work, as expected, in the media company of the future.

Table of contents

High Quality Video Production & Content Automation

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 RAI, one of these use cases focuses on supporting newsrooms and media companies in reporting unexpected events (e.g., natural or man-made disasters, crime news).

Generally, whenever such events occur, it is necessary to quickly publish fresh news and updates on a constant basis. The related workflow may be split into three tasks, which can optionally run concurrently. First, timely coverage of the unexpected event must be ensured. Here, auxiliary sources (e.g., social media) and professional crews may act synergically to promptly report about the evolution of the event. Second, the broadcaster's archives and other advanced content (e.g., video synthesis from 3D environment or drone-based professional shootings) may be used to support the storytelling with e.g., (virtual) reconstructions and footage of existing landscapes. Finally, produced content may be enriched with graphics, maps, statistics, and reports about past events related to the present one. Here, open data repositories, professional newsgathering services and broadcast archives could be considered the most important data sources.

It is noticeable from the described scenario that the "time" variable plays a crucial role: providing and getting news on time can make all the difference and the entire production process is stressed to achieve this goal. Furthermore, content quality and reliability must always be properly and readily verified before flowing into the production chain. In this context, we envisaged several macro activities that can be considered as pertaining to the following main business areas:

· **Content access**: analysis and organisation of content for efficient search, retrieval and exploration;

· **Content production:** enhancement, creation and transformation of content;

· **Content monitoring:** content reliability assessment and social media analysis.

The **content access** area involves a fundamental family of tasks for AI-based metadata extraction, like speech recognition, natural language processing and object/scene or person classification. Once extracted, these metadata need to be carefully indexed and cross-referenced. AI tools for indexing help users discover such relations, thus enabling easier and more effective search functions. As a final step, AI-driven content summarisation allows users to browse among retrieved items in a user-friendly and quick way.

The **content production** area includes activities like enhancement, creation, and transformation of content. Indeed, this area comprises tasks that span from the enhancement of existing content (e.g., video super resolution, video denoising), going through the production of new content (e.g., video synthesis from 3D environment, exploitation of aerial shots for UAV-based shooting) up to cross-modal transformation of

content (e.g., text-to-video and data-to-graphics automation). All these tasks could be smartly performed through the introduction of AI tools and techniques in the workflows.

The **content monitoring** area presents a couple of strategic application domains in the media and creative workflows. These include tasks supporting the pillars of public media services, such as privacy preservation and deep fake and inappropriate content detection. Finally, yet importantly, this area is complemented by functionalities carefully monitoring data and emerging topics from social media through new AI-based solutions for e.g., sentiment analysis and polarisation detection.

Validating the vision

Through the definition and the analysis of use cases like the one introduced in Section 2, it becomes clear how technologies can support employees of a media company to efficiently fulfil their tasks and to be competitive at a time when the media industry is fighting for its economic sustainability.

This is particularly true when several workflows must deal with big amounts of data. AI already plays a crucial role in video production and content automation tasks, and it is expected to be even more impactful in the future.

However, to develop smart AI-driven solutions in the media environment, it is vital to collect information about everyday problems and challenges, main technological needs, and related functionalities in this domain. To this end, RAI conducted two surveys during 2021, one among external experts and one addressed to its own employees, with the aim to validate the proposed vision and to gather further insights about what to prioritise in the design and developments of future AI-based tools for news and media production. The next two Sections report the fundamental findings of these two surveys.

3

Problems and challenges faced by the industry players

Although the application of AI in the media business is not a novel trend anymore, it is far from being an established industrial practice.

Even though practitioners and integrators have developed a consistent number of projects including AI functionalities at various levels of maturity*, many criticalities – mostly concerned with performance and ethics – have not been fully solved yet.

The AI4Media project, bringing together academic excellence and top media players among the ones with the higher spirit for innovation, is a unique opportunity to highlight and solve some of these criticalities, and to introduce AI tools in media processes with a structured and well-founded approach. Following this philosophy and with the aim of driving the research community towards the most useful development directions, we started our analysis from a survey among international stakeholders focussed on the main aspects of our use case to bring out problems and challenges faced by industry players.

*EBU Metadata Developer Network (MDN) Workshop, https://tech.ebu.ch/events/mdn2022

Survey scope and details

The survey was organised around the three identified business areas (i.e., Content Access, Content Production and Content Monitoring). A section related to ethical, legal and safety principles applied to content and workflows (i.e., Trustworthy AI) was also included together with some general questions about the importance of AI within modern media companies. Every business area was split into several application domains (see Appendix A for the full list), each corresponding to a set of predefined questions. These were about the importance of AI to support that application domain and the availability and efficiency of related AI tools.

RAI exploited their international network of collaborators to solicit answers from experts working for media companies and broadcasters in different European countries such as France, UK, Switzerland, the Netherlands, and Denmark, obtaining a total of 35 valid responses. The roles accounting for most of the respondents were as follows: director, journalist, researcher, archivist, audio-video editor, and technician. Among the 35 participants, 22 declared that they have held their current role for more than 6 years, showing that an important part of the respondents had a senior position and significant experience.

Common challenges

Respondents were initially asked to highlight the major problems and challenges they encounter in their daily job. Among many generic answers, some clearly pointed towards a bunch of common pains like:

- Lack of metadata integration and media information along the value chain, lack of common terminology to define the problems, excessive detachment between information generation domains;

- Difficulty to design and implement software development processes to enable efficient transition to modern systems, especially due to shortage of human resources and suboptimal organisation;

- Lack of understanding of the potential of AI, especially in the strategic levels of management of the companies.

General impact of AI

We also inquired respondents about the potential impact of AI on the cost, quality, and time efficiency of media workflows. The results (see Figure 1) showed a substantial balance, with time and quality preferred over cost. This seems to indicate that for the most part the expectations for AI-driven tools introduction are about the increase in process efficiency and less about cost reduction/optimisation.
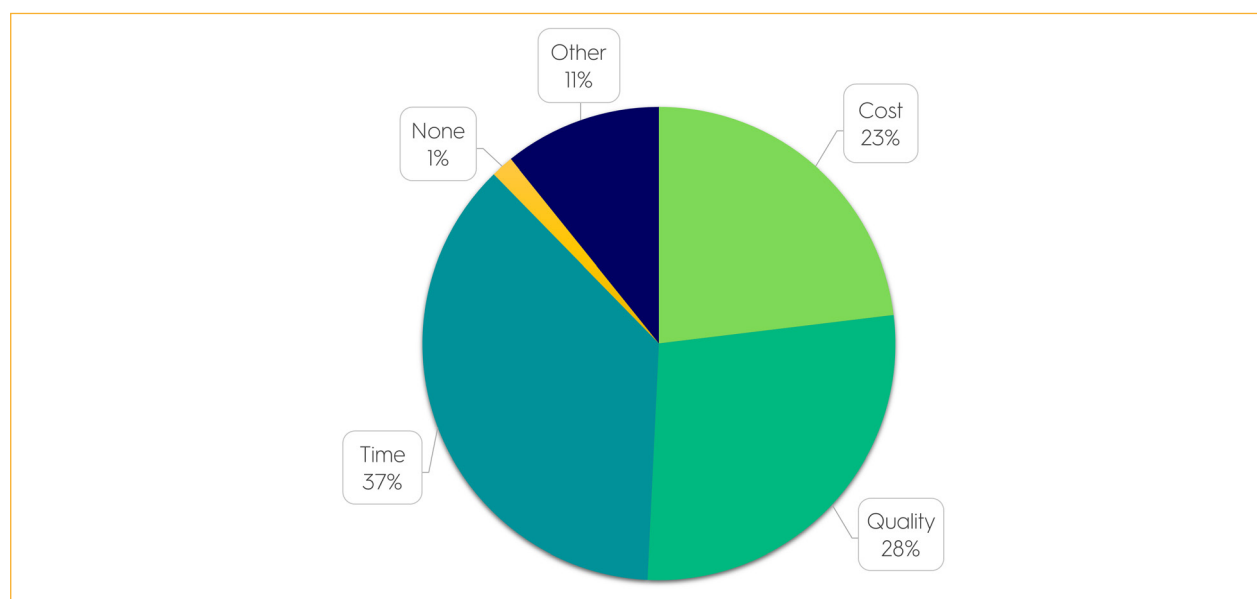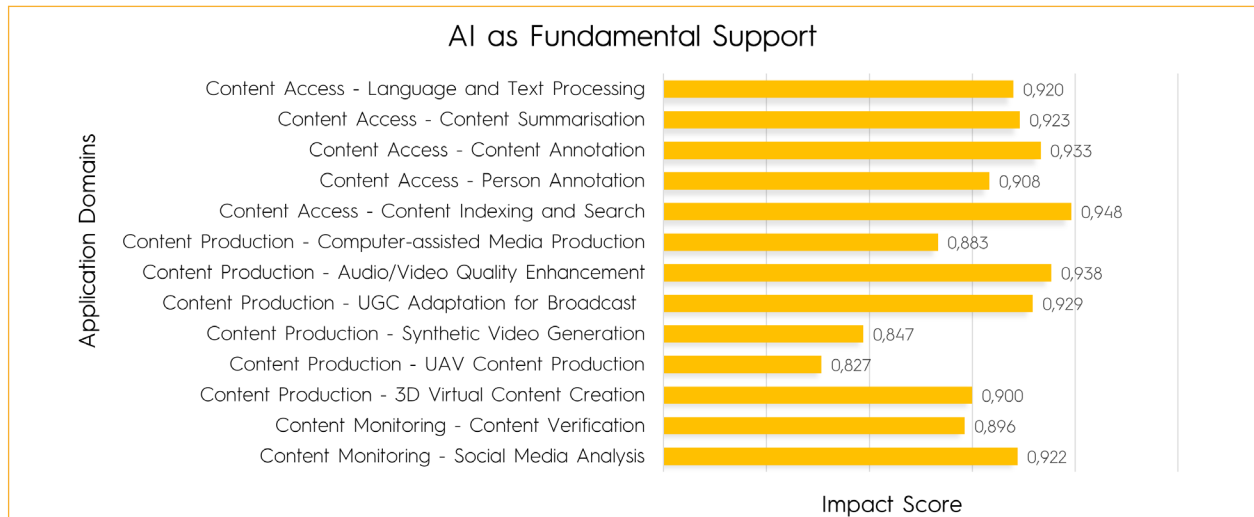

Figure 1. Aspects of AI impact

AI as Fundamental Support

Content Access - Language and Text Processing — 0,920
Content Access - Content Summarisation — 0,923
Content Access - Content Annotation — 0,933
Content Access - Person Annotation — 0,908
Content Access - Content Indexing and Search — 0,948
Content Production - Computer-assisted Media Production — 0,883
Content Production - Audio/Video Quality Enhancement — 0,938
Content Production - UGC Adaptation for Broadcast — 0,929
Content Production - Synthetic Video Generation — 0,847
Content Production - UAV Content Production — 0,827
Content Production - 3D Virtual Content Creation — 0,900
Content Monitoring - Content Verification — 0,896
Content Monitoring - Social Media Analysis — 0,922

Impact Score

Figure 2. Impact score of AI in application domains

AI impact on different application domains

We defined an "Impact score" representing the potential benefit that the exploitation and introduction of AI-based technologies could bring to a given application domain. This score, reported in Figure 2 for the different application domains, is calculated as the weighted sum of the four possible choices (i.e., "Strongly Agree", "Fairly Agree", "Disagree" and "Strongly Disagree") proposed to the statement *"AI will give a fundamental support to this application domain"*. The results show that "Content Indexing and Search", "A/V Quality Enhancement", "Content

Annotation", "UGC Adaptation for Broadcast", "Content Summarisation" and "Social Media Analysis" are the application domains where expectations of benefits from AI seem to be higher. Other important aspects explored in the survey include the usage of AI tools in the different application areas and the perceived usefulness/efficiency of these tools* .

Figure 3 shows that, except for a couple of cases where the percentage of respondents declaring the usage of tools is above 85% (i.e., "Language

Usage of Tools %

Content Access - Language and Text Processing
Content Access - Content Summarisation
Content Access - Content Annotation
Content Access - Person Annotation
Content Access - Content Indexing and Search
Content Production - Computer-assisted Media...
Content Production - Audio/Video Quality Enhancement
Content Production - UGC Adaptation for Broadcast
Content Production - Synthetic Video Generation
Content Production - UAV Content Production
Content Production - 3D Virtual Content Creation
Content Monitoring - Content Verification
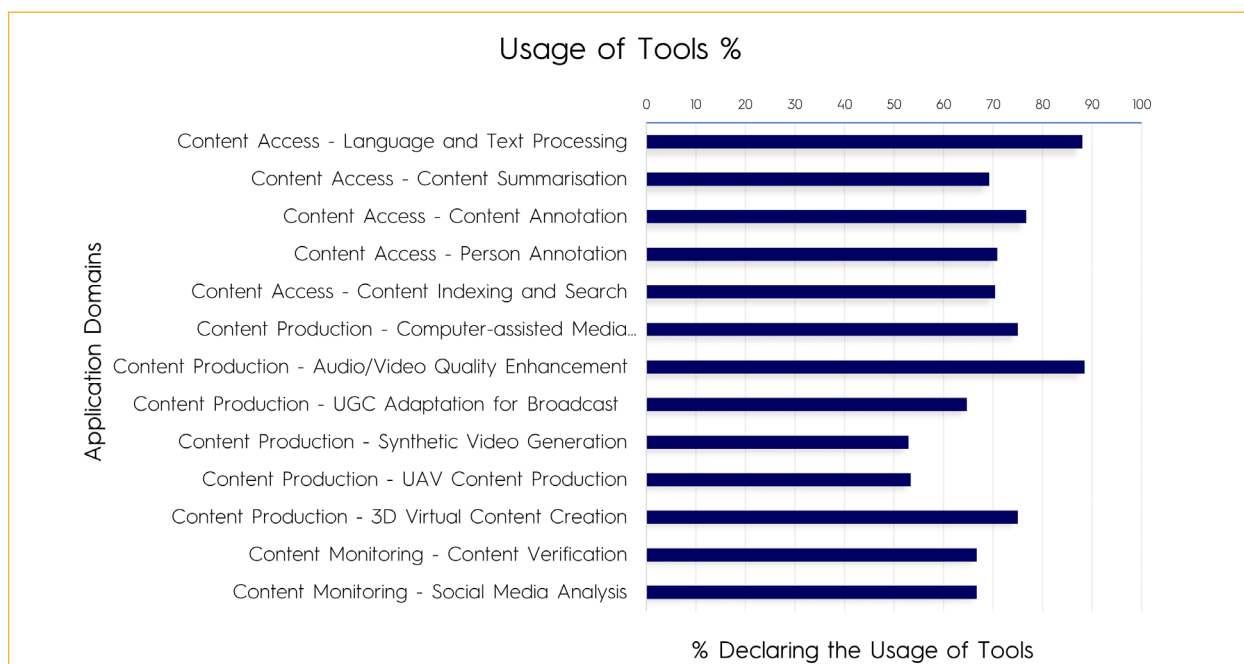Content Monitoring - Social Media Analysis

% Declaring the Usage of Tools

Figure 3. Usage of AI tools in application domains

* The two questions were respectively: "Do you already make use of AI tools or technologies that support one or more activities of this application domain?" and "Are the tools you are using so far suitable to carry out these activities efficiently?"
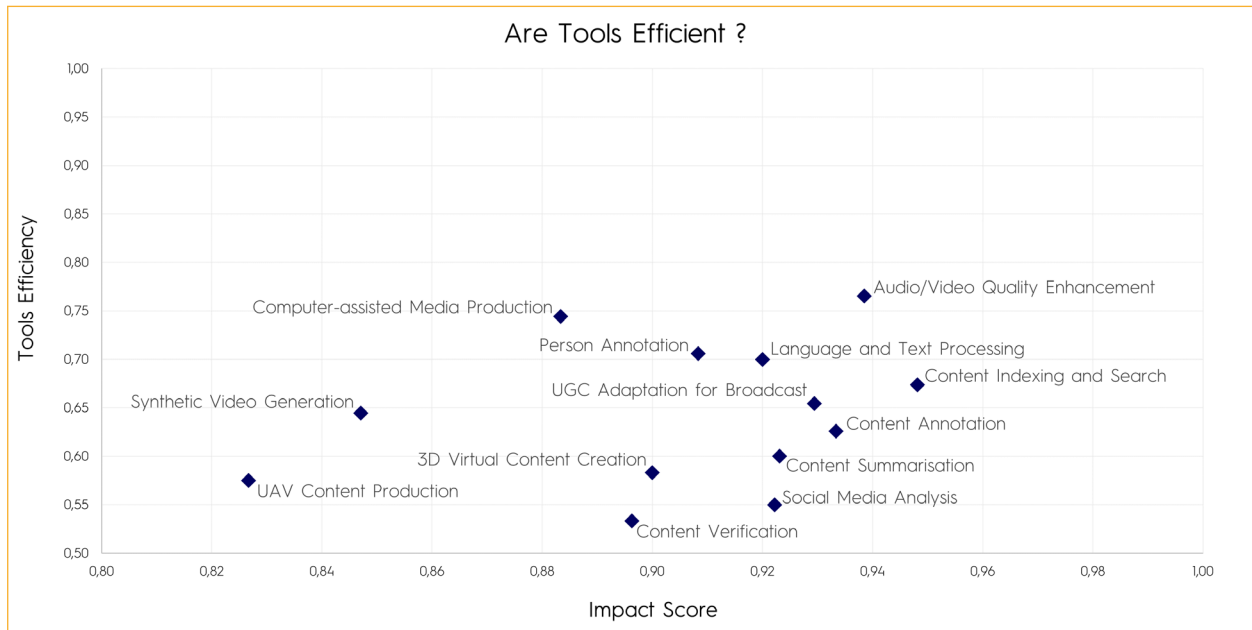
Figure 4. Tools efficiency versus impact score

and Text Analysis" and "A/V Quality Enhancement"), for most application domains the numbers are quite similar one to each other, showing a percentage in the range 50% to 70%. From Figure 4, it can also be noticed that some application domains with high expected impact of AI tools (e.g., with impact score higher than '0,9') do not have a very high tool efficiency e.g., "Social Media Analysis" ('0,92', '0,55'), "3D Virtual Content Creation" ('0,9', '0,58') and "Content Summarisation" ('0,92', '0,6'). This clearly indicates that there is much room for improvement of AI-driven tools in these domains.

Trustworthy AI aspects

Another aspect examined in the survey concerned the concept of trustworthy AI and the use of AI in modern media companies. Given the complexity of the subject and the consequent possible misunderstandings and biases that may follow in people's opinions, we kept questions very general, after giving a few examples to better contextualise the matter in the mind of the respondents (see Figure 5).

After reading the setup context, 30 out of 35 respondents stated that they had already come across the trustworthy AI concept or that they were

*In this section we cover **Trustworthy AI**, a concept that describes AI functions and systems that implement one or more of the features listed below:*

- *Preserving people's privacy (related to use of personal data)*
- *Complying with applicable regulation and legislation*
- *Being transparent/accountable by providing information on data sets, algorithms and models*
- *Explaining how predictions/outcomes were reached to the extent it is possible*
- *Listing measures taken to mitigate bias (increasing fairness)*
- *Detailing the level of technical robustness (safeguards against manipulation or attacks)*

Similar concepts are Ethical AI or Responsible AI. The aim is to adhere to ethical, legal and safety principles with technical or non-technical methods.

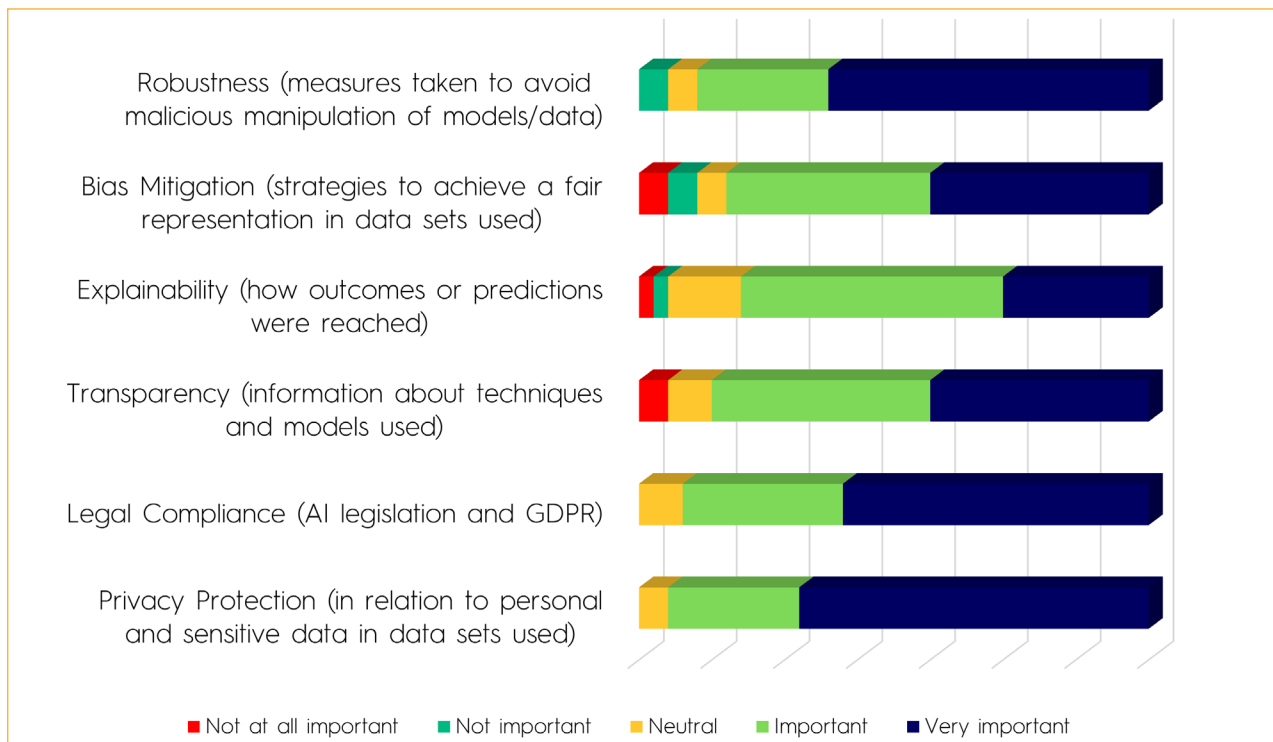Figure 5. Contextualisation of Trustworthy AI

Figure 6. Importance of Trustworthy AI features

already very familiar with it. Building upon this initial setup, we firstly explored how essential AI functions implementing specific trustworthy AI features are considered resulting in 94% of the respondents saying that Trustworthy AI functions are very useful or essential. Despite a good general acquaintance with the concept, only 11% of the respondents reported that some of the AI tools they have already used had some support for trustworthy AI.

Detailed insights were provided about the importance of each of the proposed trustworthy AI features in the production chain of a media company. Figure 6 clearly shows that respondents consider trustworthy AI features related to *Privacy Protection* and *Legal Compliance* quite important. Although in the same general direction, *Transparency, Explainability, Bias Mitigation* and *Robustness* collect a few opposite judgements.

Importance of AI adoption for media companies

We finally asked respondents to give their opinion about the importance of AI within modern media companies. The answers, on a 5-step scale from

"Strongly Disagree" to "Strongly Agree", reveal a generally strong trust in AI technology and its application in the media domain (see Figure 7).
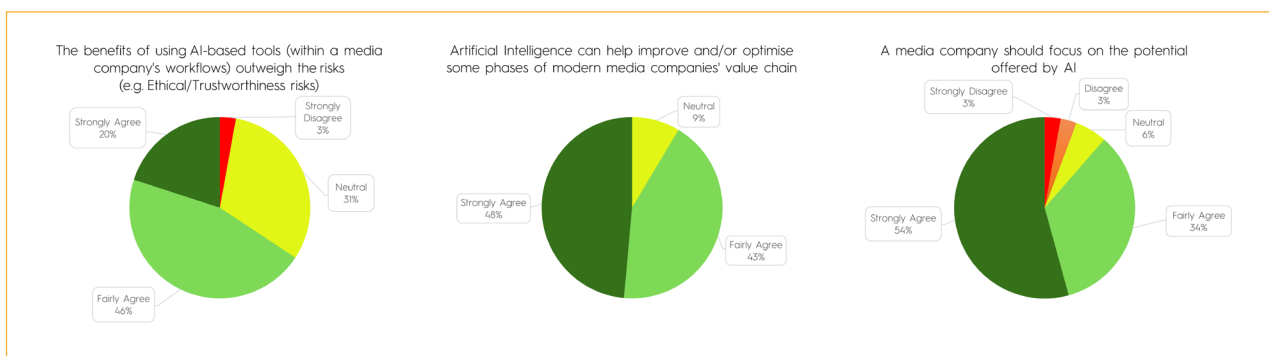

Figure 7. Importance of AI in the media domain

4

Analysis of Industrial Needs

The analysis of the external survey among the international stakeholders presented in Section 3, pointed out that, despite a certain reluctance to fully embrace AI due mostly to lack of resources and perceived shortage of skills, there is strong trust in the media business towards AI and what this technology can bring as a benefit to the media value chain.

It was also made clear that certain application domains will be impacted more by the introduction of AI than others and that not all existing AI tools and technologies are mature enough to meet industrial requirements. Driven by these indications, we moved a step further and gathered a deeper insight on these phenomena by conducting a new survey internally among RAI professionals.

More than 40 RAI employees contributed to the survey, bringing their perspectives about technological needs in the media company industry from different departments. They operate in different cities in Italy and their company roles span from management to technical and editorial ones. Even if company roles among respondents are not uniformly distributed, the sample being skewed towards "technical" profiles (e.g., ICT technician or production technician), the survey **provides qualitative insights about how different functionalities are perceived by people creating, using and leveraging them inside a media company.**

Industrial needs as "user stories"

To collect prioritised industrial needs and to drive the development of AI-based solutions, each application domain defined in the external survey was further split in "user stories", implying more specific functionalities, aligned with the breakdown analysis of requirements done in the AI4Media project. In total, we spotted 59 user stories, as reported in Appendix B. As an example, the user story "Spoken content transcription" is about the journalist wishing to be able to transform spoken language into text.

Respondents were asked to express their perception of AI in general and identify which user stories, among those 59, might be particularly useful in their daily activities. To allow us to better assess/compare respondents' feedback, the same subset of six questions/statements, some of them being open and some closed, was proposed for each functionality/user story, namely:

1. *Based on your professional or personal interests, would you like to express an opinion about this functionality? (Yes, No)*

2. *This functionality in the production chain or in processes you are involved in is/would be fundamental. (Strongly Agree, Fairly Agree, Disagree, Strongly Disagree, I don't know)*

3. *Do you already make use of tools or technologies that perform this functionality? (Often, Sometimes, Almost Never, Never)*

4. *The tools you are using so far are suitable to carry out these activities efficiently. (Strongly Agree, Fairly Agree, Disagree, Strongly Disagree)*

5. *Which tools implementing this functionality are you already using? (Open question)*

6. *Add, if you like, further considerations with respect to this functionality. (Open question)*

Importance of identified user needs

To assess, compare and rank the importance of the proposed user stories in the media company domain, i.e., to prioritise technological needs, we measured a global "Success score" for each user story (see Figure 8)* .

Even if suffering from the approximation due to the imperfect balance between involved experts' roles, this allows us to compare the importance of different user stories and to direct, in such a way, the development of AI-driven solutions.

The results clearly highlight that proposed user stories are generally perceived as important in the media company domain**. Here is the top-10 ranking among the proposed user stories:

1. UGC Video stabilization
2. Editorial scene segmentation
3. Video restoration
4. Audio denoising
5. Audio super resolution
6. Video denoising
7. Person identification in video
8. UGC video deblocking
9. Spoken content transcription
10. Person identification in audio

This partial ranking may suggest that UGC related activities together with audio & video improvement seem to be the main areas where the research community should focus on.

---

* For each user story, we first considered only the positive respondents to the question N°1, afterward evaluating answers to the statement N°2. We gave weights from '0' to '1' respectively from the "Strongly Disagree" selection (i.e., this user story is not fundamental at all) up to the "Strongly Agree" answer (i.e., this user story is fundamental). Finally, through the arithmetic mean, we measured the "Success score" for each user story.

**The Success score is in the range [0, 1]. Lower values (closer to zero) indicate that the specific user story is perceived as less important in video production and content automation workflows and should have low priority in terms of AI-based tools developments. Higher values (closer to one) indicate the opposite. In Figure 8, the lowest success scores are all above '0,5' and several user stories reach scores higher than '0,8'.

Importance of User Story

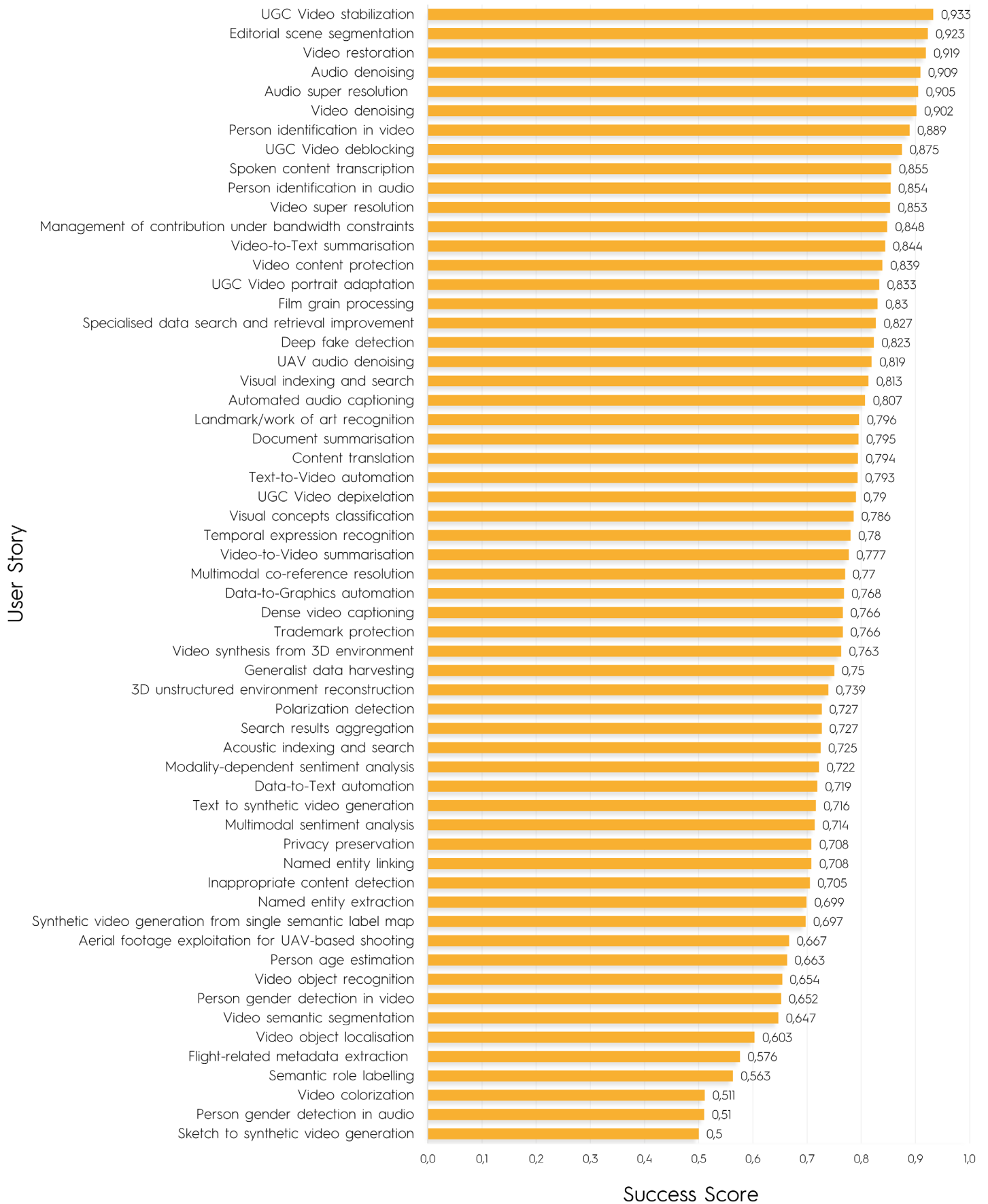| User Story	Success Score
UGC Video stabilization	0,933
Editorial scene segmentation	0,923
Video restoration	0,919
Audio denoising	0,909
Audio super resolution	0,905
Video denoising	0,902
Person identification in video	0,889
UGC Video deblocking	0,875
Spoken content transcription	0,855
Person identification in audio	0,854
Video super resolution	0,853
Management of contribution under bandwidth constraints	0,848
Video-to-Text summarisation	0,844
Video content protection	0,839
UGC Video portrait adaptation	0,833
Film grain processing	0,83
Specialised data search and retrieval improvement	0,827
Deep fake detection	0,823
UAV audio denoising	0,819
Visual indexing and search	0,813
Automated audio captioning	0,807
Landmark/work of art recognition	0,796
Document summarisation	0,795
Content translation	0,794
Text-to-Video automation	0,793
UGC Video depixelation	0,79
Visual concepts classification	0,786
Temporal expression recognition	0,78
Video-to-Video summarisation	0,777
Multimodal co-reference resolution	0,77
Data-to-Graphics automation	0,768
Dense video captioning	0,766
Trademark protection	0,766
Video synthesis from 3D environment	0,763
Generalist data harvesting	0,75
3D unstructured environment reconstruction	0,739
Polarization detection	0,727
Search results aggregation	0,727
Acoustic indexing and search	0,725
Modality-dependent sentiment analysis	0,722
Data-to-Text automation	0,719
Text to synthetic video generation	0,716
Multimodal sentiment analysis	0,714
Privacy preservation	0,708
Named entity linking	0,708
Inappropriate content detection	0,705
Named entity extraction	0,699
Synthetic video generation from single semantic label map	0,697
Aerial footage exploitation for UAV-based shooting	0,667
Person age estimation	0,663
Video object recognition	0,654
Person gender detection in video	0,652
Video semantic segmentation	0,647
Video object localisation	0,603
Flight-related metadata extraction	0,576
Semantic role labelling	0,563
Video colorization	0,511
Person gender detection in audio	0,51
Sketch to synthetic video generation	0,5

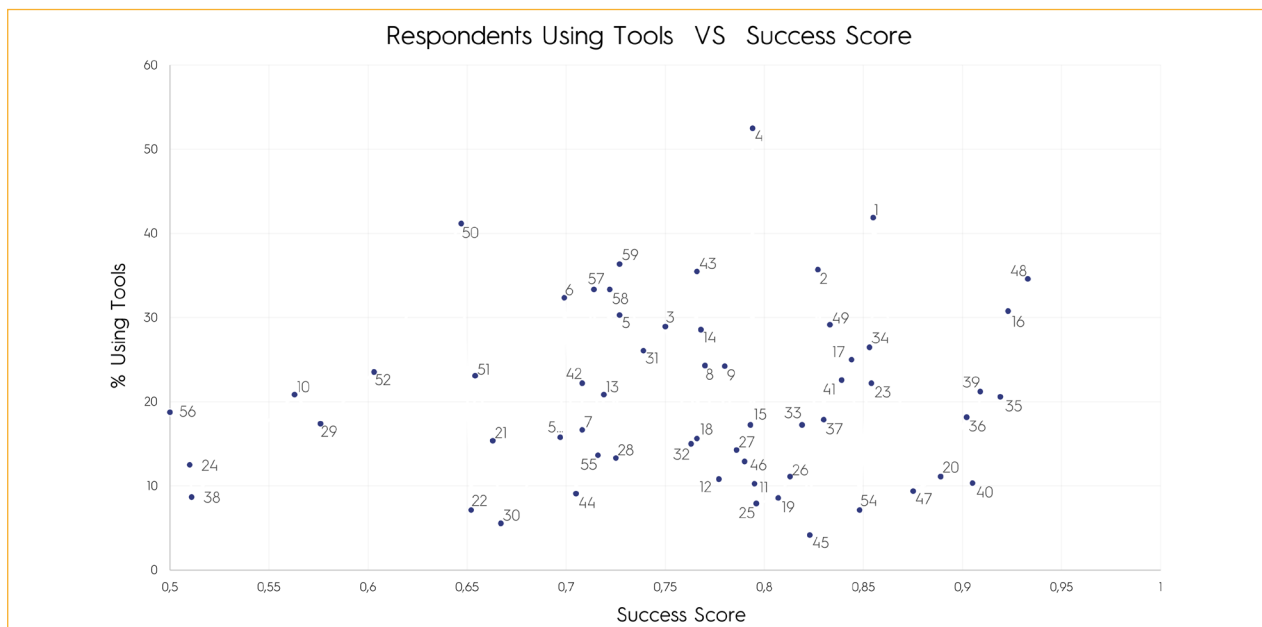Figure 8. Importance ranking of the assessed user stories

Figure 9. Respondents using tools for each user story

AI tools to support user needs

Another aspect examined was the availability and efficiency of AI tools addressing the identified user needs. For each user story, we asked respondents whether they use any AI-driven tools and, if yes, what they thought about their suitability and efficiency for the task at hand.

The results (see Figure 9 and Appendix B to get labels corresponding to numbers) show that there is only one functionality, i.e., 4-"Content translation", where more than 50% of respondents (52,5%) have reported that they use relevant AI support tools while in all other cases that percentage is significantly lower. The lowest use is reported for functionality 45-"Deep fake detection" where only 4,2% of respondents declare usage of some tool.

These findings imply either that for some of the user stories the existing tools are not reliable enough or that there are no tools specifically implemented for such tasks. Correlating the use of dedicated tools for a task with the perceived importance of this task in the media chain, **we see**

that there are many important tasks for which use of relevant tools is very low, including 40-"Audio super resolution", 36-"Video denoising", 20-"Person identification in video", 47-"UGC Video deblocking". These areas are where AI-driven solutions are even more necessary and would significantly improve and facilitate media workflows.

In addition, we examined the quality of existing tools for each user story (see Figure 10)*.

The analysis of survey responses shows that respondents in general do not have very positive feelings about existing tools implementing the proposed user stories**. In particular, several respondents are not happy with the quality of tools implementing high-impact user stories like e.g., "Video denoising", "Person identification in video", "Spoken content transcription". This outcome further highlights that **there is a strong need to improve existing tools to optimise strategic production workflows in the media company domain.**

---

* To assess the quality of tools, we considered questions N°3 and N°4, weighting both how often respondents use some tool to perform a task (question N°3, from "Often" with a higher weight to "Almost Never" with a lower weight) and how much these tools are perceived as suitable to carry out the activity efficiently (question N°4, from "Strongly Agree" with a higher weight to "Strongly Disagree" with a lower weight). Results are then added together to obtain a "Tools score" for each functionality (see orange bars in Figure 10).

Since for some user stories the number of respondents was low and their answers were quite different one to each other, the confidence interval has also been calculated (see blue bars in Figure 10).

** The tools score is included in the range [0, 1]. Lower values (closer to zero) indicate that tools used to perform a specific task are perceived as not well implemented/performing/usable and there is room to improve them (this is particularly important for user stories with high success score, shown on the right side of the x-axis in Figure 10). Higher values (closer to one) indicate the opposite.
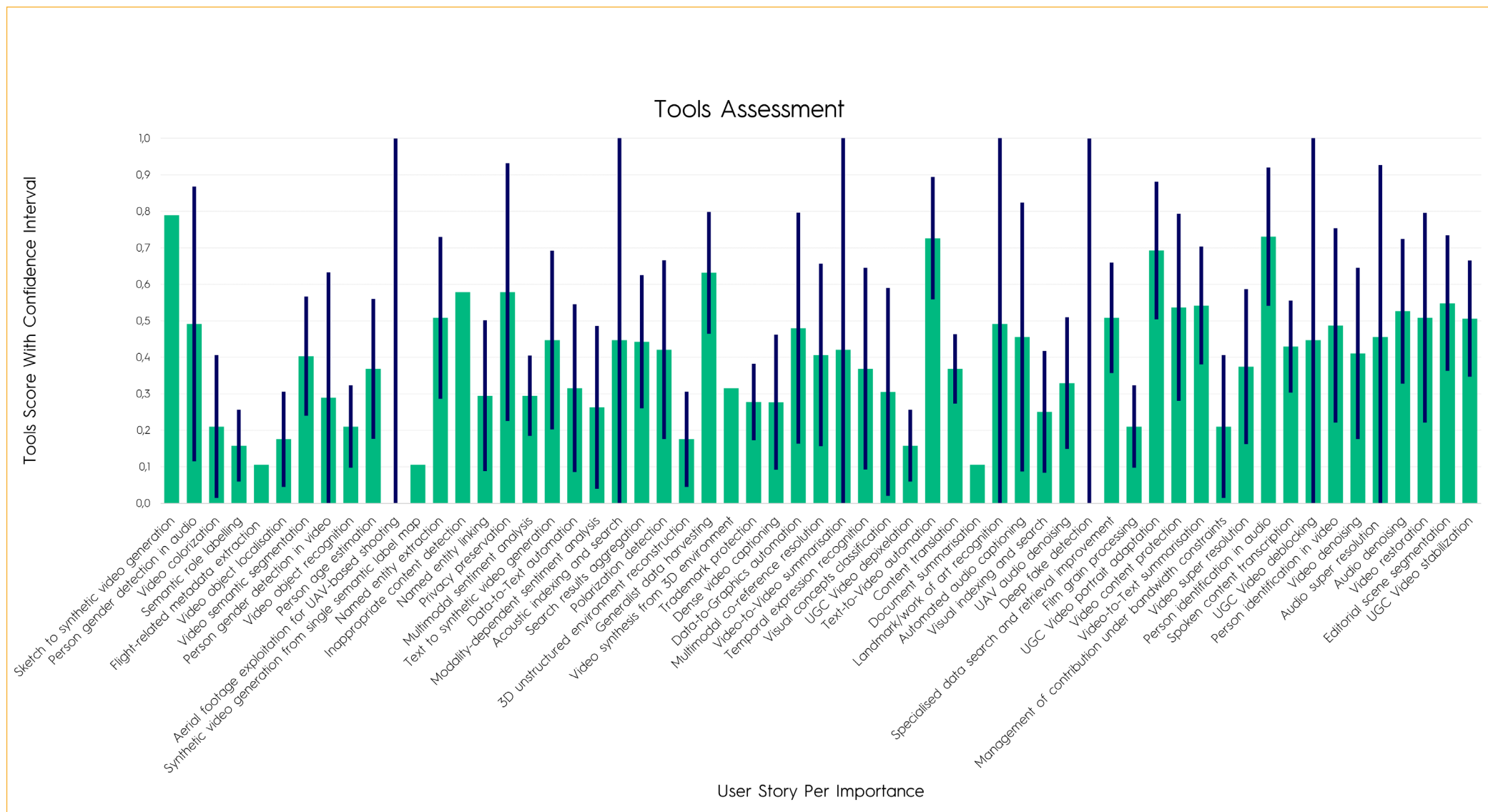
Figure 10. Assessment of existing tools

Some interesting insights

We also used open questions to get a more in-depth overview of the individual user stories and more personal opinions by the respondents. As an example, one of the user stories that aroused a lot of interest among the survey participants is the **Spoken content transcription**, which is already being used a lot. It is appreciated both by documentalists, for whom the correspondent functionality is of great support, and by journalists, especially in the case of content such as interviews (e.g., many interviews made for TV channels could easily become pieces for the web).

An interesting hint could be to apply the **Document summarisation** downstream of the Spoken content transcription, especially for archival documentation, but then it would be important to pay close attention to the result of this process because a bad synthesis could create more damage than offer benefits.

Quite some feedback was also given on the **Person identification in audio** user story. The automatic recognition of speakers in audio content would be very useful for archival documentation and would make it possible not to lose the historical memory of the voice-character connection, in particular for dialogues with frequent changes of the speaker, which happens quite often in audio content.

As for the **Video restoration** user story, assuming that vintage content still has its charm, it would be important not to use filters indiscriminately but to maintain an artistic sensibility in the renewal of the content. When dealing with the creation/modification of audio/video content, the artistic component, in general, is always considered the most important one by respondents, even if supported by AI-driven tools.

To end this brief overview of the feedback given through open questions, a mention is due to a user story particularly appreciated by the journalists participating in the survey, the **Management of contribution under bandwidth constraints**, which is considered a real turning point for mobile journalism.
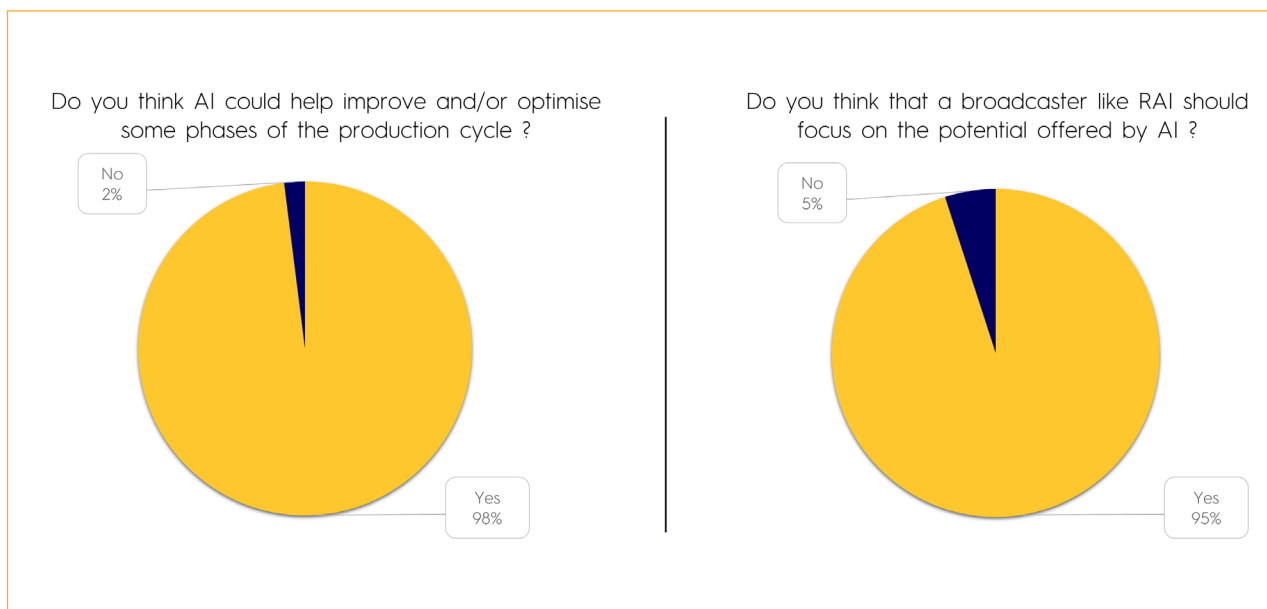
Do you think AI could help improve and/or optimise some phases of the production cycle ?

No
2%

Yes
98%

Do you think that a broadcaster like RAI should focus on the potential offered by AI ?

No
5%

Yes
95%

Figure 11. Importance of AI in the media company domain

AI's potential for the media sector: the big picture

Respondents agree that AI will play a crucial role in the future to improve and/or optimise media production workflows and that it is worth it for broadcasters like e.g., RAI to focus and invest on AI-driven technologies (see Figure 11). This confirms what was already observed in the external survey.

There is in general the certainty that **AI-based technologies are going to be very important in future workflows within media companies,** although it is not still clear to many respondents how far we are from having fully operational and qualitatively high-level functionalities. **AI technologies can give a great boost in**

terms of quality and creativity, but there is an explicit fear that these technological innovations could replace human work, which instead, according to the survey participants, should always remain central. The ethical issue related to the introduction of AI in media companies' workflows is also deeply felt, and it is necessary to consider very carefully the risks associated with the use of new AI-based functionalities in this domain.

Conclusions

This white paper gives a qualitative overview of how the possible incoming introduction of AI tools in media companies' workflows is perceived by different levels of experts in this area, such as managers, editors, and technicians.

In particular, the focus was on tasks related to high-quality video production and content automation, two core activities in the media value chain, on which the introduction of AI technologies could have a strong impact in the future.

Collected feedback can help define the most relevant topics in this domain to drive the research community towards the most useful research and development directions.

While on the one hand the trust that the vast majority of stakeholders place in the introduction of AI is very evident, on the other hand, the problems they expect to be able to solve are quite significant, and there is still not much clarity on how this transition to modern AI-driven systems could take place.

In a historical moment in which the way of making information and building multimedia products is evolving towards something new,

e.g., using UGC or content of the past that needs to be improved to go on air, many activities of the media value chain described in this white paper are considered fundamental. These industrial needs become even more pronounced considering the great lack of suitable tools to perform many of these activities, as highlighted by the survey participants.

Trustworthy AI features are also considered very important in this domain, especially those related to Privacy Protection and Legal Compliance, although we are still far from having AI tools fully encompassing such elements of trustworthy AI.

Appendix A

List of application domains proposed by the external survey:

- Content Access:
  - Language and Text Processing
  - Content Summarisation
  - Content Annotation
  - Person Annotation
  - Content Indexing and Search

- Content Production:
  - Computer-assisted Media Production
  - Audio/Video Quality Enhancement
  - UGC Adaptation for Broadcast
  - Synthetic Video Generation
  - UAV Content Production
  - 3D Virtual Content Creation

- Content Monitoring
  - Content Verification
  - Social Media Analysis

Appendix B

List of user stories proposed by the internal survey:

1. Spoken content transcription
2. Specialised data search and retrieval improvement
3. Generalist data harvesting
4. Content translation
5. Search results aggregation
6. Named entity extraction
7. Named entity linking
8. Multimodal co-reference resolution
9. Temporal expression recognition
10. Semantic role labelling
11. Document summarisation
12. Video-to-Video summarisation
13. Data-to-Text automation
14. Data-to-Graphics automation
15. Text-to-Video automation
16. Editorial scene segmentation
17. Video-to-Text summarisation
18. Dense video captioning
19. Automated audio captioning
20. Person identification in video
21. Person age estimation
22. Person gender detection in video
23. Person identification in audio
24. Person gender detection in audio
25. Landmark/work of art recognition
26. Visual indexing and search
27. Visual concepts classification
28. Acoustic indexing and search
29. Flight-related metadata extraction
30. Aerial footage exploitation for UAV-based shooting
31. 3D unstructured environment reconstruction
32. Video synthesis from 3D environment
33. UAV audio denoising
34. Video super resolution
35. Video restoration
36. Video denoising
37. Film grain processing
38. Video colorization
39. Audio denoising
40. Audio super resolution
41. Video content protection
42. Privacy preservation
43. Trademark protection
44. Inappropriate content detection
45. Deep fake detection
46. UGC Video de-pixelation
47. UGC Video deblocking
48. UGC Video stabilization
49. UGC Video portrait adaptation
50. Video semantic segmentation
51. Video object recognition
52. Video object localisation
53. Synthetic video generation from single semantic label map
54. Management of contribution under bandwidth constraints
55. Text to synthetic video generation
56. Sketch to synthetic video generation
57. Multimodal sentiment analysis
58. Modality-dependent sentiment analysis
59. Polarization detection

AI4media

Follow us 🐦 in f @ai4mediaproject