



Deliverable 7.2: Activities and outcome of the Pilots, second report

Gerhard Rudinger (APA-IT)

30/3/2018

Work Package 7: Pilots, Testing and User Feedback

InVID - In Video Veritas: Verification of Social Media Video Content for the News Industry

Innovation Action

Horizon 2020, Research and Innovation Programme

Grant Agreement Number 687786

Dissemination level	PU
Contractual date of delivery	31/3/2018
Actual date of delivery	30/3/2018
Deliverable number	D7.2
Deliverable name	Activities and outcome of the Pilots, second report
File	invid_d7.2_v1.0.docx
Nature	Report
Status & version	Final, V1.0
Number of pages	54
WP contributing to the deliverable	WP7
Task responsible	APA-IT
Other contributors	AFP DW
Author(s)	Gerhard Rudinger / APA-IT
Quality Assessors	Evlampios Apostolidis / CERTH Jan Thomsen / Condat
EC Project Officer	Alberto Rabbachin
Keywords	Testing, evaluation, user feedback

Abstract

In WP7, the technologies integrated into the InVID platform and applications have been tested and evaluated in various editorial cases and trials. This document reports the results of test cycles 4 to 6, which were held between M19 and M27 of the project. The first part of the document gives a general description of the test cycle concept in the InVID project. This is followed by an overview of the tested components and applications within the project's test cycles 4 to 6 and also the evaluation methods that were applied in the test cycles for these components and applications. Then the main part of the document reports the major results of test cycles 4 to 6. This part is structured according to the different components and applications. For each component and application, the focus of the tests, the number of feedback comments and the major outcomes of the test cycles are reported.

Table of contents

1	Introduction	8
1.1	History of the document	8
1.2	Glossary of acronyms.....	8
2	Overview of test cycles four to six.....	9
2.1	Objectives of WP7.....	9
2.2	General description of the test cycles.....	9
2.3	Overview of test cycles 4 to 6.....	10
3	Results of test cycles four to six.....	16
3.1	General remarks	16
3.2	Video Fragmentation & Annotation Service	16
3.3	Web Application for Video Fragmentation & Reverse Image Search	18
3.4	Near Duplicate Detection Service.....	20
3.5	Logo Detection Service	21
3.6	Forensic Analysis Service	23
3.7	Context Aggregation & Analysis Service	24
3.8	Rights Management Service	26
3.9	InVID Multimodal Analytics Dashboard and Tool for Social Media Retrieval and Topic Detection	27
3.10	InVID Verification Plugin.....	29
3.11	InVID Verification Application	40
3.12	InVID Mobile Application	52
3.13	InVID Core Platform API	53
4	Conclusions and outlook for the next test cycles.....	54

List of Figures

Figure 1: Tested components in test cycles 4 to 6	11
Figure 2: InVID Verification Plugin: Overall impression.....	30
Figure 3: InVID Verification Plugin: Results for the question “Was the plugin easy to use?” .	30
Figure 4: InVID Verification Plugin: Results for the multiple response question “What was missing in your experience?”	31
Figure 5: InVID Verification Plugin: Results for the question “Would you recommend the InVID plug-in to your newsroom or a colleague?”	31
Figure 6: InVID Verification Plugin: Results for the question “Did you like the InVID plugin?”	32
Figure 7: InVID Verification Plugin: Results for the question “Are you satisfied with the results provided by the analysis feature?”	32
Figure 8: InVID Verification Plugin - Analysis: Results for the multiple response question “What is the most useful for you?”	33
Figure 9: InVID Verification Plugin - Keyframe segmentation: Results for the question “Are you satisfied with the results provided by the keyframe segmentation feature?”	33
Figure 10: InVID Verification Plugin - Keyframe segmentation: Results for the multiple response question “In your opinion, what should we improve?”	34
Figure 11: InVID Verification Plugin - YouTube thumbnail: Results for the question “Are you satisfied with the results provided by the YouTube thumbnail feature?”	34
Figure 12: InVID Verification Plugin - YouTube thumbnail: Results for the multiple response question “In your opinion, what should we improve?”.....	35
Figure 13: InVID Verification Plugin - Twitter advanced search: Results for the question “Are you satisfied with the results provided by the Twitter advanced search?”	35
Figure 14: InVID Verification Plugin - Twitter advanced search: Results for the multiple response question “In your opinion, what should we improve?”	36
Figure 15: InVID Verification Plugin - Image magnifier: Results for the question “Are you satisfied with the results provided by the image magnifier?”	36
Figure 16: InVID Verification Plugin - Image magnifier: Results for the multiple response question “In your opinion, what should we improve?”.....	37
Figure 17: InVID Verification Plugin - Metadata: Results for the question “Are you satisfied with the results provided by the Metadata feature?”.....	37
Figure 18: InVID Verification Plugin - Metadata: Results for the multiple response question “In your opinion, what should we improve?”	38
Figure 19: InVID Verification Plugin - Picture forensic: Results for the question “Are you satisfied with the results provided by the forensic feature?”	38
Figure 20: InVID Verification Plugin - Picture forensic: Results for the multiple response question “In your opinion, what should we improve?”.....	39
Figure 21: InVID Verification Application – external testers: Overall impression	41
Figure 22: InVID Verification Application – internal testers: Overall impression	41
Figure 23: InVID Verification Application – external testers: Results for the question “Did you like the InVID plugin?”	42
Figure 24: InVID Verification Application – internal testers: Results for the question “Did you like the InVID plugin?”	42

Figure 25: InVID Verification Application – external testers: Results for the question “Was the Verification Application easy to use?”	43
Figure 26: InVID Verification Application – internal testers: Results for the question “Was the Verification Application easy to use?”	43
Figure 27: InVID Verification Application – external testers: Results for the question “Is the division of the screen in cards appropriate?”	44
Figure 28: InVID Verification Application – internal testers: Results for the question “Is the division of the screen in cards appropriate?”	44
Figure 29: InVID Verification Application – external testers: Results for the question “Is the use of the verification cards on the right side of the screen clear?”	45
Figure 30: InVID Verification Application – internal testers: Results for the question “Is the use of the verification cards on the right side of the screen clear?”	45
Figure 31: InVID Verification Application – external testers: Results for the question “Is the naming of the different cards self-explanatory about their functionality?”	46
Figure 32: InVID Verification Application – internal testers: Results for the question “Is the naming of the different cards self-explanatory about their functionality?”	46
Figure 33: InVID Verification Application – external testers: Results for the question “Is the placement of feature results in separate windows or the verification cards appropriate?”	47
Figure 34: InVID Verification Application – internal testers: Results for the question “Is the placement of feature results in separate windows or the verification cards appropriate?”	47
Figure 35: InVID Verification Application – external testers: Results for the question “Video player: Is the behaviour of cursor, click in keyframes, jumping of screen sections when zooming understandable?”	48
Figure 36: InVID Verification Application – internal testers: Results for the question “Video player: Is the behaviour of cursor, click in keyframes, jumping of screen sections when zooming understandable?”	48
Figure 37: InVID Verification Application – external testers: Near duplicates: Results for the question “Are the offered filters and their toggling clear?”	49
Figure 38: InVID Verification Application – internal testers: Near duplicates: Results for the question “Are the offered filters and their toggling clear?”	49
Figure 39: InVID Verification Application – external testers: Near duplicates: Results for the question “Is the presentation of results in relation to the current video clear?”	50
Figure 40: InVID Verification Application – internal testers: Near duplicates: Results for the question “Is the presentation of results in relation to the current video clear?”	50
Figure 41: InVID Verification Application – external testers: Results for the question “Are you satisfied with the results provided by the near duplicate feature?”	51
Figure 42: InVID Verification Application – internal testers: Results for the question “Are you satisfied with the results provided by the near duplicate feature?”	51

List of Tables

Table 1: History of the document.....	8
Table 2: Glossary of acronyms	8
Table 3: Time plan of the project's test and validation cycles.....	9
Table 4: Tested components in test cycles 4 to 6	11
Table 5: Use of the different testing methods for the different applications/components (where TC refers to Testing Cycle)	13
Table 6: Testing times of test cycles 4 to 6.....	15
Table 7: Number of received feedback comments for the Video Fragmentation & Annotation Service	17
Table 8: Number of received feedback comments for the Web Application for Video Fragmentation & Reverse Image Search	18
Table 9: Number of received feedback comments for the Near Duplicate Detection Service	20
Table 10: Number of received feedback comments for the Logo Detection Service	22
Table 11: Number of received feedback comments for the Forensic Analysis Service.....	23
Table 12: Number of received feedback comments for the Context Aggregation & Analysis Service	24
Table 13: Number of received feedback comments for the Rights Management Service.....	26
Table 14: Number of received feedback comments for the InVID Multimodal Analytics Dashboard and Tool for Social Media Retrieval and Topic Detection.....	28
Table 15: Number of received feedback comments for the Verification Plugin.....	29
Table 16: Number of received feedback comments for the InVID Verification Application	40
Table 17: Number of received feedback comments for the InVID Mobile Application	52
Table 18: Number of received feedback comments for the InVID Core Platform API.....	53

1 Introduction

In WP7, the technologies integrated into the InVID platform and applications are tested and evaluated in various editorial cases and trials. The primary aim is to collect user feedback both on the tools and components themselves, their usability and appropriateness for various tasks, and the results of the system as a whole (e.g. in terms of reliability and accuracy).

Using an iterative approach over the entire project duration (with overall nine validation cycles), the feedback and results obtained from each cycle of the trials are used to improve the subsequent versions of the applications, platform and components. Tests and evaluations are focused on video news emerging from social networks and media websites.

In full accordance with the InVID DoA (Description of Action), this deliverable extends the reportings of D7.1, by covering the pilot testing carried out as part of the project's development and validation cycles 4 to 6, which run from M19 to M27.

1.1 History of the document

Table 1: History of the document

Date	Version	Name	Comment
16/03/2018	V0.90	Gerhard Rudinger	Initial version
20/03/2018	V0.91	Jan Thomsen	QA
23/03/2018	V0.92	Evlampios Apostolidis	QA
30/03/2018	V1.00	Gerhard Rudinger	Final version

1.2 Glossary of acronyms

Table 2: Glossary of acronyms

Acronym	Explanation
API	Application Programming Interface
JSON	JavaScript Object Notation
UGC	User Generated Content
UGV	User Generated Video
(G)UI	(Graphical) User Interface
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
REST	Representational State Transfer
XML	Extensible Markup Language

2 Overview of test cycles four to six

2.1 Objectives of WP7

As discussed in D7.1 (see Section 3), through the work carried out in WP7, the technologies integrated into the InVID platform and applications are tested and evaluated in various editorial cases and trials. This is done with the help of different user groups. The primary aim of the conducted evaluations is to collect the users' feedback about the usability and appropriateness of the exposed tools and components for performing various tasks, and the overall efficiency of the InVID system in terms of reliability and accuracy.

Driven by the objectives of the InVID project, the performed tests and evaluations focus on the analysis of newsworthy videos distributed via social networks and media websites, as well as on User Generated Content provided by users' communities that have been developed by regional newspapers.

2.2 General description of the test cycles

2.2.1 Time plan of the test cycles

In InVID we follow an agile development methodology which includes nine test and validation cycles. The findings of each evaluation cycle and the feedback collected from the participants (which could be both internal users from the InVID consortium and external users from outside the consortium) are exploited to improve the following versions of the exposed analysis components, integrated applications and the overall InVID platform.

The test cycles 4 to 6 took place from month 19 to month 27 (see Table 3).

Table 3: Time plan of the project's test and validation cycles

Month	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
Test cycle							Pilots prep.			Test cycle 1	Test cycle 2	Test cycle 3	Test cycle 4	Test cycle 5	Test cycle 6	Test cycle 7	Test cycle 8	Test cycle 9																		

2.2.2 Phases of a test cycle

As described in D7.1, each test cycle consists of the following phases:

- Preparations for the testing
 - Registration of the applications/components that will be included in the test cycle
 - Exposition of a stable version of each involved component and application for testing

- Gathering all the needed data for getting access to the exposed technologies, such as URLs and access credentials
- For each individual application/component, documentation of the applied changes since its last tested version
- Development of the established testing guidelines
- Setup of templates for recording the outcomes of the evaluations
- Definition of the exact time-plan for the testings
- Testing
 - Re-assessment of the issues found and fixed based on the prior test cycles
 - Testing of the different applications/components based on the guidelines and evaluation scenarios that were determined by the developer(s) of each technology
 - Reporting of the evaluation outcomes with the help of the provided testing templates
- Follow-ups
 - Merging of the results collected from the different testers
 - Provision of the aggregated evaluation outcomes (on the technology-level), to the technology partners of InVID
 - Organization of a conference call among the technology providers and the technology testers (involving only members of the InVID consortium and not external testers), to clarify potential questions and comments regarding the test results

2.3 Overview of test cycles 4 to 6

2.3.1 Applications tested in test cycles 4 to 6

In general we included as many applications and components as possible in each test cycle. Applications or components that were not yet ready for testing were skipped in the respective test cycle. This could occur for components which had a longer development phase and were not yet stable enough for testing.

Figure 1 shows the overall InVID architecture diagram. All components that were evaluated in the test cycles 4 to 6 are marked with a red bounding box. Some analysis components that are not directly accessible through individual user interfaces, were tested through the InVID applications that integrate these components.

Table 4 lists the components evaluated in each test cycle.

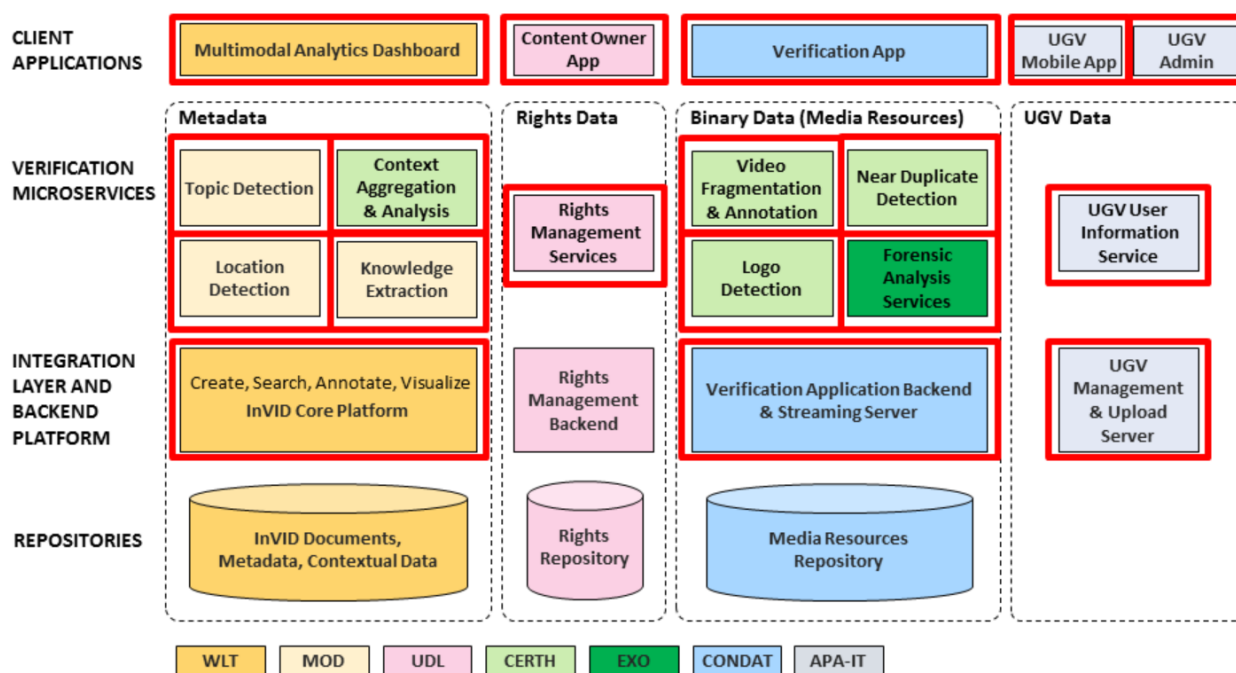


Figure 1: Tested components in test cycles 4 to 6

Table 4: Tested components in test cycles 4 to 6

Applications	Test cycles		
	4	5	6
Video Fragmentation & Annotation Service	Yes	Indirect	Yes
Web Application for Video Fragmentation & Reverse Image Search	Yes	Indirect	Yes
Near Duplicate Detection Service	Yes	Indirect	Yes
Logo Detection Service	Yes	Indirect	No
Forensic Analysis Service	No	Indirect	Yes
Context Aggregation & Analysis Service	Yes	Indirect	Yes
Rights Management Service	Yes	Indirect	Yes
Tool for Social Media Retrieval and Topic Detection	Yes	No	No
InVID Multimodal Analytics Dashboard	Yes	No	No
InVID Verification Plugin	Yes	Yes	Yes
InVID Verification Application	Partial	Yes	Yes
InVID Mobile Application	Yes	Yes	Yes
InVID Core Platform-API	Yes	No	No

In the table above, by "indirect" testing we mean that the corresponding component was tested through its integration in the Verification Plugin and the Verification Application. Moreover, "partial" testing indicates that, the InVID Verification Application was in a longer development phase and a version with reduced functionality was evaluated in the relevant test cycle.

Test cycle 5 was focused on the evaluation of the main integrated InVID tools and applications by users external to the consortium. The participants to this test cycle, assessed the functionality of the InVID Verification Plugin, the InVID Verification Application and the InVID Mobile Application. As a side activity, other services were tested indirectly in this test cycle through their integration in the evaluated technologies.

2.3.2 Testing methods used in test cycles 4 to 6

A detailed overview of existing testing and evaluation methodologies was given in Section 2 of D7.1. From the reported methods, the following ones were used in test cycles 4 to 6:

Targets of testing

- **Integration testing:** Interface testing was applied on all individual analysis components, accessible either via their user interfaces or their APIs.
- **System testing:** Testing of the integration of systems was used for the integrated tools and applications that are accessible mostly via user interfaces, i.e. the Verification Plugin, the Verification Application, the Mobile Application and the Multimodal Analytics Dashboard.

Remark: Unit testing is part of the development process and it is not addressed in WP7.

Objectives of testing

- **Alpha testing:** All exposed applications and components were tested by users within the consortium.
- **Beta testing:** Part of the integrated tools and applications, namely the Verification Plugin, the Verification Application and the Mobile Application, were tested by potential and/or existing users/customers.
- **Functional testing:** In all test cycles the functionality of the components/applications was tested.
- **Non-functional testing:** In all test cycles the non-functional aspects of the components/applications, such as their reliability, were tested.
- **Usability testing:** In all test cycles the testers evaluated the degree to which the system can be used by specified users with effectiveness, efficiency and satisfaction, in a specified context of use.
- **Performance testing:** The performance of each individual component/application was tested in the conducted test cycles. The exposed technologies were evaluated both in a qualitative manner (e.g. is the efficiency of this technology sufficient enough for a journalist?) and a quantitative way (e.g. are the produced results fully accurate and complete?).

- **Regression testing:** In the test cycles the functionality of the components/ applications was re-tested, to ensure that no defects have been introduced as a result of the changes made based on the feedback of previous test cycles.

Testing techniques

- **Exploratory testing:** Besides testing with guidelines, every tester also evaluated each technology by following an exploratory manner.
- **Fault-based techniques:** Fault-based techniques were mainly used in the integration (API) tests. The tests on API level enable a stable interaction between the different components of the system and thus also a stable basis for all user interfaces.
- **Scenario testing:** Scenario testing in the test cycles was based on a typical journalistic workflow for video verification and provided information on how the exposed and tested technologies help with these tasks.
- **Walkthrough testing:** Test cycle 4 included a walkthrough testing via a dedicated session that involved both the testers and the developers of the technologies. In test cycle 6 a walkthrough testing was performed for the InVID Verification Application

These testing methods and techniques were chosen because they are suitable for providing adequately detailed and well-justified feedback on the usability and efficiency of the evaluated applications and components. The different testing methods were often used in combination, for example, functional, non-functional and performance testing were performed in one test sequence.

Table 5 gives an overview of the testing method(s) that was(were) used for assessing each component and/or integrated system, in the test cycles of the reporting period. The tests were performed manually, and in the case of the API tests also automatically with automation tools (JMeter). The automatic API tests ensure the functionality of the tested APIs and reduce time and effort for re-testing the APIs after a change.

Table 5: Use of the different testing methods for the different applications/components (where TC refers to Testing Cycle)

Application	Targets of testing		Objectives of testing						Testing techniques			
	Integration testing	System testing	Alpha testing	Beta testing	Regression testing	Functional testing	Non-functional testing	Performance testing	Exploratory testing	Fault-based techniques	Scenario testing	Walkthrough testing
Video Fragm. & Annotation Service	TC4 TC6	TC4 TC5 TC6	TC4 TC6	TC5	TC4 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC6	TC4 TC6	

Web App. for Video Fragm. & Reverse Image Search	TC4 TC6	TC4 TC5 TC6	TC4 TC6	TC4 TC5 TC6	TC4 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC6	TC4 TC6	TC4
Near Duplicate Detection Service	TC4 TC6	TC4 TC5 TC6	TC4 TC6	TC5	TC4 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC6	TC4 TC6	
Logo Detection Service	TC4	TC4 TC5	TC4	TC5	TC4	TC4 TC5	TC4 TC5	TC4 TC5	TC4 TC5	TC4	TC4	TC4
Forensic Analysis Service	TC6	TC5 TC6	TC5 TC6	TC5	TC5 TC6	TC5 TC6	TC5 TC6	TC5 TC6	TC5 TC6	TC6		
Context Aggregation & Analysis Service	TC4 TC6	TC4 TC5 TC6	TC4 TC6	TC4 TC5 TC6	TC4 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC6	TC4 TC6	TC4
Rights Management Service	TC4	TC4 TC5 TC6	TC4 TC6	TC5	TC4 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC6	TC4 TC6	TC4
Tool for Social Media Retrieval & Topic Detect.		TC4	TC4		TC4	TC4	TC4	TC4	TC4	TC4	TC4	TC4
InVID Multimodal Analytics Dashboard	TC4	TC4	TC4		TC4	TC4	TC4	TC4	TC4	TC4	TC4	TC4
InVID Verification Plugin		TC4 TC5 TC6	TC4 TC6	TC4 TC5 TC6	TC4 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC6	TC4
InVID Verification Application	TC4 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC5	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC6
InVID Mobile Application		TC5 TC6	TC4 TC6	TC5	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC4 TC5 TC6	TC5 TC6	
InVID Core Platform-API	TC4	TC4			TC4	TC4	TC4	TC4	TC4	TC4		

2.3.3 Testers

The components and applications were tested by different groups of testers

- Members of the consortium: Partners with journalistic background (AFP, DW)
- Members of the companies of the consortium, but external to the project (AFP, DW)
- External testers: e.g. testers from Storyful, BBC, Tiroler Tageszeitung, Vorarlberger Nachrichten (regional newspapers in Austria), France 24, Almasy Alyoum (Daily independent newspaper in Egypt).
- IT specialists from the companies of the consortium for the testing of the technical interfaces (APIs) (AFP, APA-IT)

2.3.4 Testing dates

Table 6 shows the time period for the implementation of test cycles 4 to 6. These periods do not include the preparation and the follow-up phase of each test cycle.

Table 6: Testing times of test cycles 4 to 6

Test cycle	Time period
Test cycle 4	2017-09-01 to 2017-09-30
Test cycle 5	2017-12-04 to 2017-12-20
Test cycle 6	2018-02-01 to 2018-02-21

3 Results of test cycles four to six

3.1 General remarks

Following a reporting approach similar to the one in D7.1, the results of the three test cycles are structured according to the different components and applications. Each of the following subsections is dedicated to a different component/application. It starts with a short description of the relevant technology and then it provides a description of the performed tests. The number of individual responses/comments about the tested technology is reported at the test-cycle-level. Finally, the main outcomes about the performance of the evaluated technology are listed.

3.2 Video Fragmentation & Annotation Service

3.2.1 Description of the service

The Video Fragmentation and Annotation Service is a web service (API) that performs a temporal decomposition of a video into three different levels of granularity; scenes (i.e. semantically and temporally coherent segments that correspond to the story-telling parts of the video), shots (i.e. sequences of frames captured uninterruptedly by a single camera) and sub-shots (i.e. visually coherent parts of a video shot; useful when analysing single-shot videos). Consecutively, it identifies the semantics of the video at the most fine-grained level (either shots or sub-shots, depending on the type of analysed video) by detecting a number of high-level visual concepts after analysing one representative key frame per video fragment.

3.2.2 Tests

A dedicated test of this service was executed in test cycles 4 and 6. The testing of the service's API was performed by IT technicians from AFP and APA-IT.

Test cycles 4 and 6 focused on the evaluation of:

- The service's responses on a variety of incoming requests; both the format of the response and the response time were taken under consideration.
- The functionality of the supported types of analysis, namely the one that performs video segmentation and the one that performs video segmentation and concept-based annotation.
- The service's communication interface for retrieving the analysis results.
- The service's compatibility with the supported video platforms.

Table 7: Number of received feedback comments for the Video Fragmentation & Annotation Service

Test cycle	Feedback comments
Test cycle 4	6
Test cycle 5	--
Test cycle 6	9

3.2.3 Major outcomes of the test cycles

As the Video Fragmentation and Annotation Service is accessible through its API, all tests were performed on the communication interface level. The results of the conducted evaluations within test cycles 4 and 6 indicated a significant improvement of the service's functionality. This improvement was based on the collected feedback after test cycle 4, and the efforts made to address the users' recommendations and requirements.

Feedback on the service's documentation:

The documentation was improved significantly compared to the one evaluated in test cycle 4, and it was considered as "Good" in test cycle 6.

Feedback on reliability:

The reliability of the API was rated as "Not so good" in test cycle 4, due to problems observed with various Facebook URLs. This problem was fixed after test cycle 4, and the service's reliability during test cycle 6 was rated as "Neutral" and "Good". The "Neutral" effect was caused by the delay in the processing of incoming analysis requests. However, this delay was a result of the heavy use of the service by the InVID Multimodal Analytics Dashboard over the testing period. Appropriate actions to address this issue have already been taken. In particular, the service has been upgraded in terms of processing resources by installing a clone of it in another machine and distributing the incoming requests in the two machines with the help of a load balancing mechanism. In addition, the video collection strategy of the InVID Multimodal Analytics Dashboard has been updated in order to exclude long videos and screencasts of video games, which are of limited interest in terms of newsworthiness. Additional stress tests will be made to assess the processing capacity of the new infrastructure and quantify the improvement after the applied changes.

Feedback on response time:

The evaluation of the response time of the service ranged from "Neutral" to "Very good".

Feedback on robustness and error handling:

The robustness and error handling capability of the service was marked as "Neutral" to "Very good".

Last but not least, additional feedback was given, in the form of recommendations, for fine-tuning the service's response in case that a previously submitted and analysed video is no longer available.

3.3 Web Application for Video Fragmentation & Reverse Image Search

3.3.1 Description of the service

The Web Application for Video Fragmentation & Reverse Image Search, accessible through its user interface, allows the user to extract a set of representative key frames from a video, and to search for occurrences of these key frames on the Web, through the reverse image search functionality of the Google search engine.

3.3.2 Tests

This web tool was tested in test cycles 4 and 6. This analysis component is integrated into the InVID Verification Plugin, and was therefore assessed also through the testing of this technology (see Section 3.10).

Test cycle 4 focused on:

- The user interface of the tool
- The key frame selection efficiency
- The exploitation of the analysis results
- The compatibility of the tool with the supported video platforms

Test cycle 6 focused on the:

- The response time of the tool
- The reliability of the tool in terms of access and use
- The compatibility of the tool with the newly supported platforms

Table 8: Number of received feedback comments for the Web Application for Video Fragmentation & Reverse Image Search

Test cycle	Feedback comments
Test cycle 4	4
Test cycle 5	--
Test cycle 6	7

3.3.3 Major outcomes of the test cycles

The testers stated that the tool helps a journalist a lot to get more insights about the re-use of a video, and the added value of this web application for a journalist is ranging from "High" to "Very high".

Feedback on the user interface:

The user interface, which was formed according to the users' feedback after the previous test cycles and was further fine-tuned on specific aspects, was labeled as "Good".

Feedback on the key frame selection efficiency:

Based on the collected feedback, also, from users of the InVID plugin (which integrates this tool), the key frame selection mechanism of the service was updated to provide on demand (i.e. after a user is clicking on a newly added button that appears right below the initial collection of extracted key frames) an additional set of keyframes, thus allowing a more detailed investigation for previous occurrences of the video on the Web. The efficiency of the updated mechanism was marked as "Good".

Feedback on the exploitation of the analysis results:

The exploitation of the analysis results (i.e. the extracted representative key frames) for assessing the originality of a video, was considered as "Very easy".

Feedback on the compatibility with the supported video platforms:

The compatibility of the web application with the supported video platforms was evaluated as "Good". However, it was noticed that the processing of videos from Vimeo and DailyMotion takes longer than the processing of videos from YouTube and Twitter. Also some Vimeo clips were not processed correctly. Finally, the analysis of videos shared via Dropbox was problematic during test cycle 4, but this issue was fixed and the corresponding analysis pipeline worked smoothly in test cycle 6.

Feedback on the response time:

The evaluation outcomes about the response time of the tool, ranged from "Neutral" to "Very good", depending on the platform that hosts the video (see relevant comment above). Moreover, the heavy use of the underlying service by the InVID Multimodal Analytics Dashboard (see also Section 3.2 about the Video Fragmentation & Annotation Service) resulted, from time to time, in longer delays of the analysis. Nevertheless, this issue was solved during the test cycle by extending the underlying service with a prioritization mechanism that allows a more time-efficient and balanced handling of incoming analysis requests from different components and users of the InVID ecosystem.

Feedback on the reliability:

The reliability of the tool was considered as "Good".

Feedback from test cycle 4 focusing on expanding the compatibility of the tool to additional widely used video sharing platforms, such as Instagram and LiveLeaks, was taken under consideration. The web application was updated in order to support the suggested platforms, and its updated functionality was evaluated in test cycle 6.

3.4 Near Duplicate Detection Service

3.4.1 Description of the service

The Near Duplicate Detection Service, which is accessible via its API, aims to identify near-duplicate content. It gives a similarity value based on indexed videos and therefore allows the detection of near-duplicates of an image or video. This a strong indication of prior use of this media item in the past, and thus an evidence that it is not original.

The functionality of the Near Duplicate Detection Service is available to the InVID users through the InVID Verification Application.

3.4.2 Tests

A dedicated test of this service was executed in test cycles 4 and 6. The testing of the API was performed by a group of IT technicians from AFP and APA-IT.

Both test cycles (4 and 6) focused on the service's:

- Speed
- Robustness and error handling
- Performance in terms of detecting near-duplicate videos

Evaluations of the accuracy and appropriateness of the analysis results were done by journalists through the testing of the InVID Verification Application (see Section 3.11 about the Verification Application).

Table 9: Number of received feedback comments for the Near Duplicate Detection Service

Test cycle	Feedback comments
Test cycle 4	8
Test cycle 5	--
Test cycle 6	4

3.4.3 Major outcomes of the test cycles

The service is very well documented, reliable and no major bugs were found during test cycles 4 and 6.

Feedback on documentation:

The documentation of the service was assessed as "Good".

Feedback on reliability:

The reliability of the service was considered as "Very good".

Feedback on response time:

The response time of the service was rated as "Good" to "Very good".

Feedback on robustness and error handling:

The robustness and error handling capability of the service was marked as "Good".

Additional feedback for the service was related mainly to the improvement of the error handling mechanism of the service (e.g. handling the case of missing or invalid parameters in the incoming request) and improvements in the error messages of the service.

3.5 Logo Detection Service

3.5.1 Description

The Logo Detection Service (accessible both via its API and its UI) evaluates the existence of a given logo within a video or image. The service integrates an extendable pool of logos of particular interest for journalists, and after processing a video or image, it provides information about the occurrence of these logos in the video frames or within the image, aiming to assist investigators in identifying the origin of the media item under investigation.

The user interface of the Logo Detection service was developed only for demonstrations purposes and internal testing within the InVID project. The functionality of the Logo Detection Service is accessible to the InVID users, via the InVID Verification Application.

3.5.2 Tests

The user interface and the API of this service were already in a very good shape. So, a dedicated test of this service was only performed in test cycle 4.

As the user interface of the service is not intended for external use, but only for internal tests and demonstrations within the InVID consortium, all tests were done by testers of the InVID project with journalistic background. The testing of the service's API was done by IT technicians.

The conducted tests focused on:

- The robustness and error handling capability of the service
- The time efficiency of the analysis

Further tests regarding the accuracy and appropriateness of the analysis results were performed by journalists through the testing of the InVID Verification Application.

Table 10: Number of received feedback comments for the Logo Detection Service

Test cycle	Feedback comments from testing the user interface	Feedback comments from testing the API
Test cycle 4	4	7
Test cycle 5	--	--
Test cycle 6	--	--

3.5.3 Major outcomes of the test cycles

Major results from the user interface tests:

Feedback on robustness:

The robustness of the service in general was assessed as "Good"; however, sporadically submitted videos for analysis weren't shown in the user interface of the service.

Feedback on the time efficiency of the analysis:

The testers had mixed subjective expectations about the needed time for logo detection within a video. Their feedback ranged from "Good" to "Not good at all".

Additional feedback suggested further improvements in usability aspects. For example, the testers were not sure if they have to play the video in order to initiate logo detection. Another recommendation was to inform a user when a new logo has been added in the existing pool of searchable logos.

Major results from the API tests:

The results of the conducted tests were very good. This is a reasonable outcome due to the improvements made based on the feedback of the previous test cycles.

Feedback on documentation:

The documentation of the service was rated as "Good"; however, the explanation of some new functionality was missing in the documentation.

Feedback on reliability:

The reliability of the service was considered as "Very good" in general, but Vimeo videos failed to be processed.

Feedback on response time:

The response time of the service was labeled from "Good" to "Very good". Only the response time needed for initiating the analysis of a video from Dailymotion was a bit slower.

Feedback on robustness and error handling:

The robustness and error handling capability of the service was rated as "Good".

3.6 Forensic Analysis Service

3.6.1 Description of the service

The Forensic Analysis Service provides a set of developed forensic filters that assist users in detecting manipulation of videos.

3.6.2 Tests

A dedicated test of this service was executed in test cycle 6. The testing of the service's API was done by IT technicians from AFP and APA-IT. The focus of the tests was the functionality and the error handling capability of the API.

Further tests of the Forensic Analysis Service, concerning the accuracy and appropriateness of the analysis results were done by journalists, through the testing of the Verification Application which integrates this analysis component (see Section 3.11 about the Verification Application).

Table 11: Number of received feedback comments for the Forensic Analysis Service

Test cycle	Feedback comments
Test cycle 4	--
Test cycle 5	--
Test cycle 6	9

3.6.3 Major outcomes of the test cycles

The service is very well documented and reliable. No major bugs were found in test cycle 6. The main feedback was directed to the improvement of error handling capability of the service (see below).

Feedback on documentation:

The documentation of the service was rated as "Very good".

Feedback on reliability:

The reliability of the service was considered as "Very good".

Feedback on response time:

The response time of the service was evaluated from "Good" to "Very good".

Feedback on robustness and error handling:

The robustness and error handling capability of the service was marked as "Neutral" and "Not so good". The testers suggested further improvement of the error handling mechanism of the service by performing a more precise and stricter checking of the

input parameters (e.g. invalid filter parameter values do not cause an error) and improving the error messages of the service.

3.7 Context Aggregation & Analysis Service

3.7.1 Description of the service

The Context Aggregation & Analysis Service (accessible via an API and a user interface) aims to facilitate the verification of the content derived exclusively from the YouTube and Facebook platforms.

3.7.2 Tests

As mentioned above, this service provides a user interface and an API. Both were tested in all three test cycles. Dedicated tests of the user interface were done in test cycles 4 and 6 by testers of the consortium with journalistic background. The testing of the service's API was done by IT technicians from AFP and APA-IT.

This analysis component is integrated into the InVID Verification Plugin, and was therefore assessed also through the testing of the Verification Plugin (see Section 3.10).

In test cycle 4, the focus of the user interface and API tests included:

- The robustness of the component and error handling
- The new method for extracting locations
- The updated verification words list

In test cycle 6 the conducted tests focused on:

- The evaluation of execution time improvement
- The update to support native Twitter videos
- The detailed report per hour in the weather report

Table 12: Number of received feedback comments for the Context Aggregation & Analysis Service

Test cycle	Feedback comments from testing the user interface	Feedback comments from testing the API
Test cycle 4	22	7
Test cycle 5	--	--
Test cycle 6	4	0

3.7.3 Major outcomes of the test cycles

Major results from the user interface tests:

The testers mentioned that this tool helps a journalist to get more insights about a video based on the contextual data. The added value of this tool for a journalist was rated as "High" to "Very high".

Feedback on robustness and error handling of the application:

The robustness of this tool was rated as "Not so good" to "Good" in test cycle 4, because of the need to refresh the user interface from time to time. The latter was significantly improved based on this feedback, and these problems were not observed anymore in test cycle 6.

Feedback on processing time:

The processing time of the integrated analysis component was considered as "Good" and "Very good".

Reported problems with the detection of locations in test cycle 4 were solved with the integration of the Recognize module from webLyzard.

Major results from the API tests:

In test cycle 4 some bugs were found, and a number of suggestions for improvements were made, whereas in test cycle 6 no further bugs and suggestions occurred. The testers stated that this is a very good and complete API, which is easily integratable into a GUI.

Feedback on documentation

The documentation of the service was rated from "Good" to "Very good".

Feedback on reliability:

The reliability of the API was considered as "Very good".

Feedback on response time:

The response time of the service was marked from "Good" to "Very good".

Feedback on robustness and error handling:

The robustness and error handling capability of the service was labeled as "Very good".

3.8 Rights Management Service

3.8.1 Description of the service

The InVID Rights Management Service (accessible both via an API and a user interface) deals with the copyright aspects related to the reuse of UGC. It helps to discover the owner of an interesting piece of UGC, to contact the owner, set a copyright negotiation framework and establish the required rights agreements to reuse the asset.

3.8.2 Tests

The user interface of the Rights Management Service was tested in test cycles 4 and 6. As the API for the Rights Management Service is in excellent status, it was tested only in test cycle 4. The testing of the user interface was done by testers of the consortium with journalistic background. The testing of the service's API was done by IT technicians from AFP and APA-IT.

In test cycles 4 the user interface tests focused on:

- The registration of YouTube, Twitter and Facebook videos
- The evaluation of the reimplementations of the user interface

In test cycles 4 the API-tests focused on:

- The registration of YouTube, Twitter and Facebook videos
- The content authentication with the new identity provider service

In test cycles 6 the user interface tests focused on:

- The creation of reuse requests
- The acceptance or rejection of a reuse request as a content uploader

Further tests on this technology were done by journalists through the testing of the Verification Application which integrates this service.

Table 13: Number of received feedback comments for the Rights Management Service

Test cycle	Feedback comments from testing the user interface	Feedback comments from testing the API
Test cycle 4	5	6
Test cycle 5	--	--
Test cycle 6	6	--

3.8.3 Major outcomes of the test cycles

Major results from the user interface tests:

The testers stated that this tool helps a journalist a lot to clear the rights situation about a newsworthy video. The user interface was in a very stable and usable status. The main negative feedback from test cycle 4 was related to some observed issue with Twitter and Facebook videos. These problems were fixed and successfully re-tested in test cycle 6. The major feedback from test cycle 6 was a recommendation to add an unlimited or automatic renewable license period in the request of reuse form. Additional minor suggestions were made, e.g. to include a link to a webpage explaining what this organization is about in the request of reuse or some wording and suggestions for the wording.

Major results from the API tests:

The results from the API tests were very good. The API is very well documented, reliable and no major bugs were found during the tests.

Feedback on documentation:

The documentation of the service was rated as "Good".

Feedback on reliability:

The reliability of the service was marked from "Good" to "Very good".

Feedback on response time:

The response time of the service was considered as being "Good" to "Very good".

Feedback on robustness and error handling:

The robustness and error handling capability of the service was labeled as "Good".

3.9 InVID Multimodal Analytics Dashboard and Tool for Social Media Retrieval and Topic Detection

3.9.1 Description of the service

The InVID Multimodal Analytics Dashboard is a tool for media analytics and monitoring. In the InVID project the existing dashboard of webLyzard is enhanced with new functionalities, such as the display of collected and selected videos, the ability to track evolving stories on the Web, and many more.

The Tool for Social Media Retrieval and Topic Detection collects tweets from various news accounts on Twitter, video documents from various social networks, and metadata for each social media document. Through the InVID Multimodal Analytics Dashboard, the social media can be explored and the metadata can be used, both as a means to filter/query the documents, as well as to provide additional context for the user when a document is selected.

The story detection clusters the collected documents and identifies them as belonging to distinct news stories. Also, filtering by news topic is provided as functionality. Label stories are accurately based on keywords. The most relevant documents are provided for each story.

3.9.2 Tests

The InVID Multimodal Analytics Dashboard and the integrated Tool for Social Media Retrieval and Topic Detection was tested in test cycle 4 by testers of the consortium.

The tests focused on the evaluation of:

- The new supported Streamgraph visualisation
- The improved colour scheme in visualisations for better readability
- The relevance of different sorting and ranking settings for a journalistic workflow.

Table 14: Number of received feedback comments for the InVID Multimodal Analytics Dashboard and Tool for Social Media Retrieval and Topic Detection

Test cycle	Feedback comments
Test cycle 4	11
Test cycle 5	--
Test cycle 6	--

3.9.3 Major outcomes of the test cycles

The newly introduced stream graph visualisation was rated as being more intuitive than the storyflow visualisation. Moreover, the preferred sorting of stories (by relevance or by date) depends on the use case of a journalist. A sorting by relevance is important when the journalist is looking for trending topics and the sorting by date is more important in case of a breaking story. Furthermore, the “related keyword” section was considered as good, and the testers suggested an improvement of this technology in order to enable the selection of a combination of keywords that helps to narrow down the results. Finally, for the story views, the testers requested a way to explore all documents in a cluster, rather than only the first ten of them.

3.10 InVID Verification Plugin

3.10.1 Description of the service

The InVID Verification Plugin wraps up a number of tools and services, developed to help journalists in the verification/fact-checking process. It provides a simple way to use those media verification technologies directly in the browser. Moreover, it enables users to provide their opinion about the usability of the integrated tools and suggest improvements of these technologies, via the integrated instant feedback mechanism.

3.10.2 Tests

The InVID Verification Plugin was tested in test cycles 4 and 5 by 36 testers outside of the consortium. The testers included people from companies, such as France 24, BBC and Storyful. The results of these tests were collected with the help of an online survey containing 40 questions regarding the functionality of both the InVID Verification Plugin and each individual component of it.

Additional tests of this integrated technology were done by internal testers of the consortium with journalistic background.

Table 15: Number of received feedback comments for the Verification Plugin

Test cycle	Survey responses	Feedback comments
Test cycle 4	7 survey responses	15 items
Test cycle 5	29 survey responses	14 items
Test cycle 6	--	10 items

3.10.3 Major outcomes of the test cycles

The findings of the survey indicate that both the InVID Verification Plugin and the integrated components in it were very well appreciated by its users. The data shown in the following diagrams represent the opinion of 36 testers that are external to the project consortium, regarding several aspects about the usability and functionality of this technology.

Results of the survey for general questions

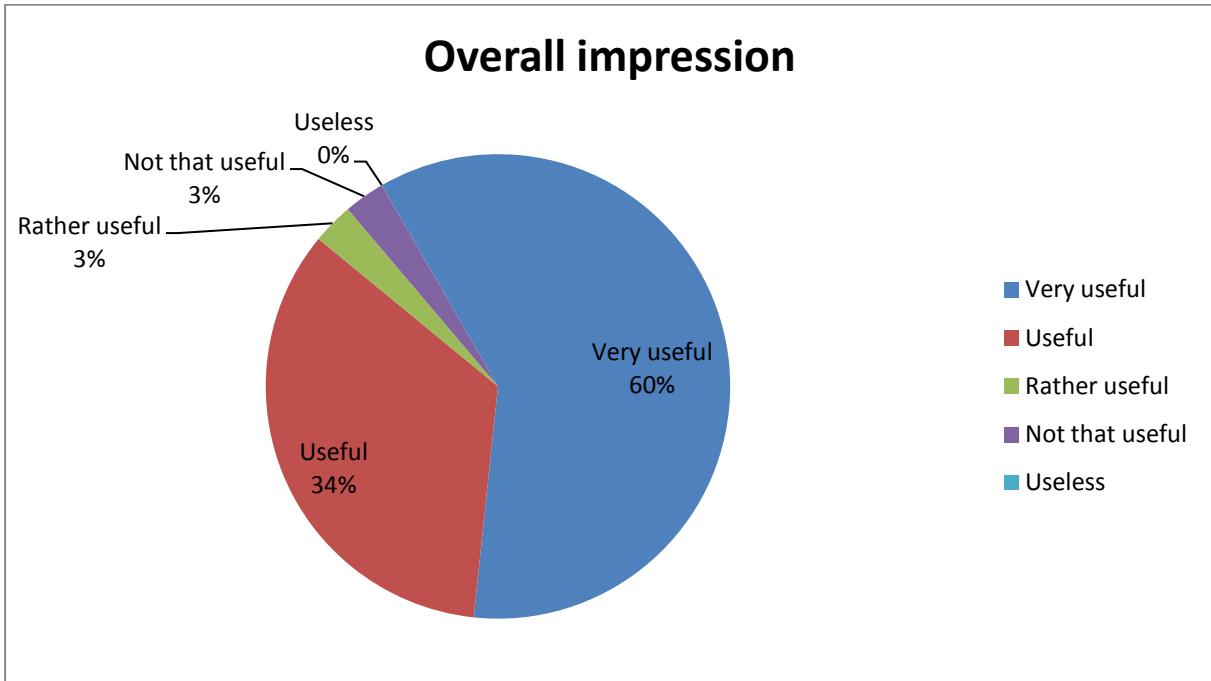


Figure 2: InVID Verification Plugin: Overall impression

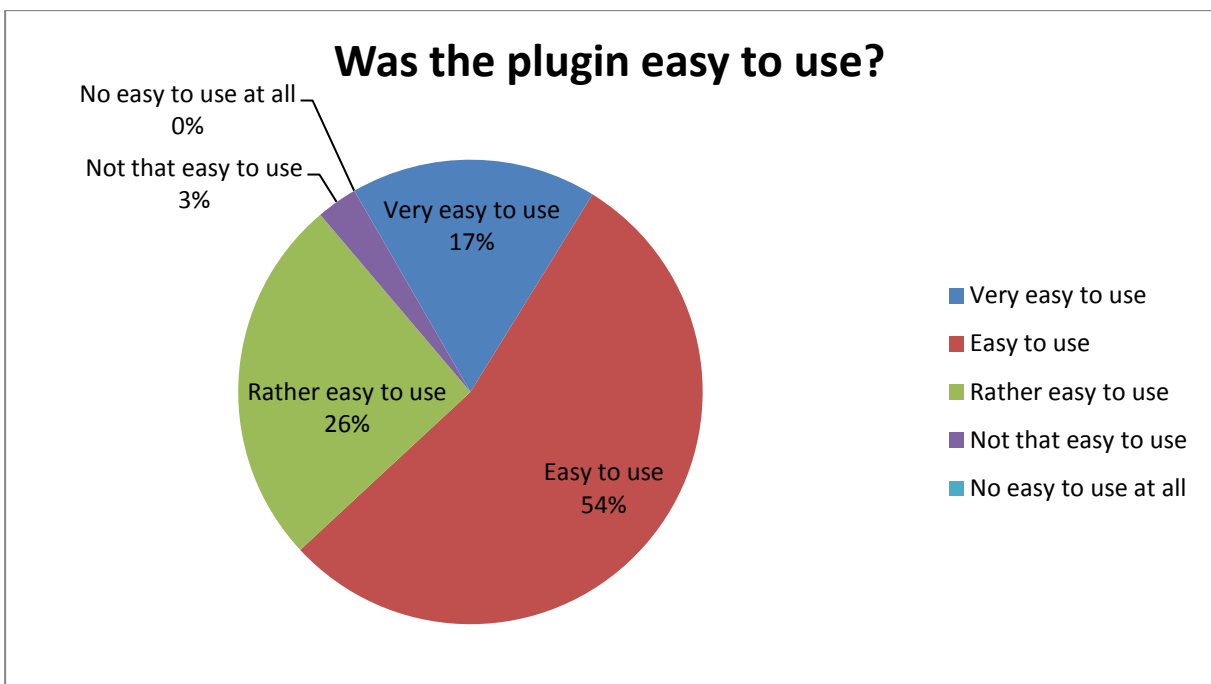


Figure 3: InVID Verification Plugin: Results for the question “Was the plugin easy to use?”

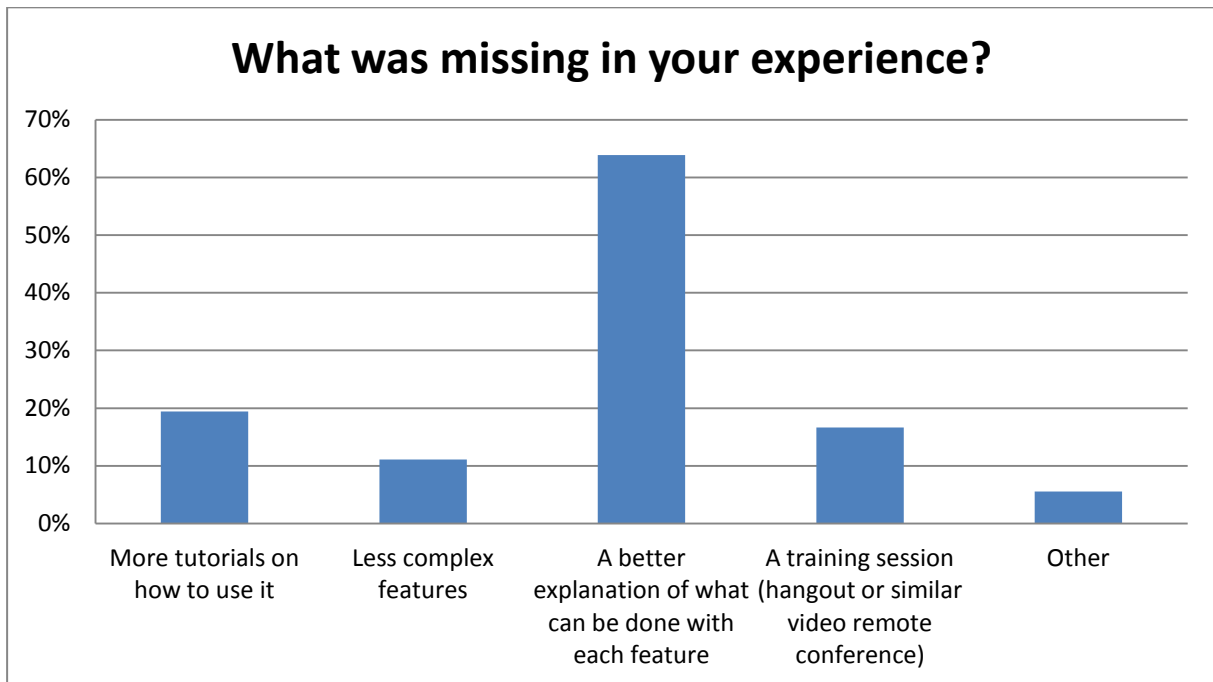


Figure 4: InVID Verification Plugin: Results for the multiple response question “What was missing in your experience?”

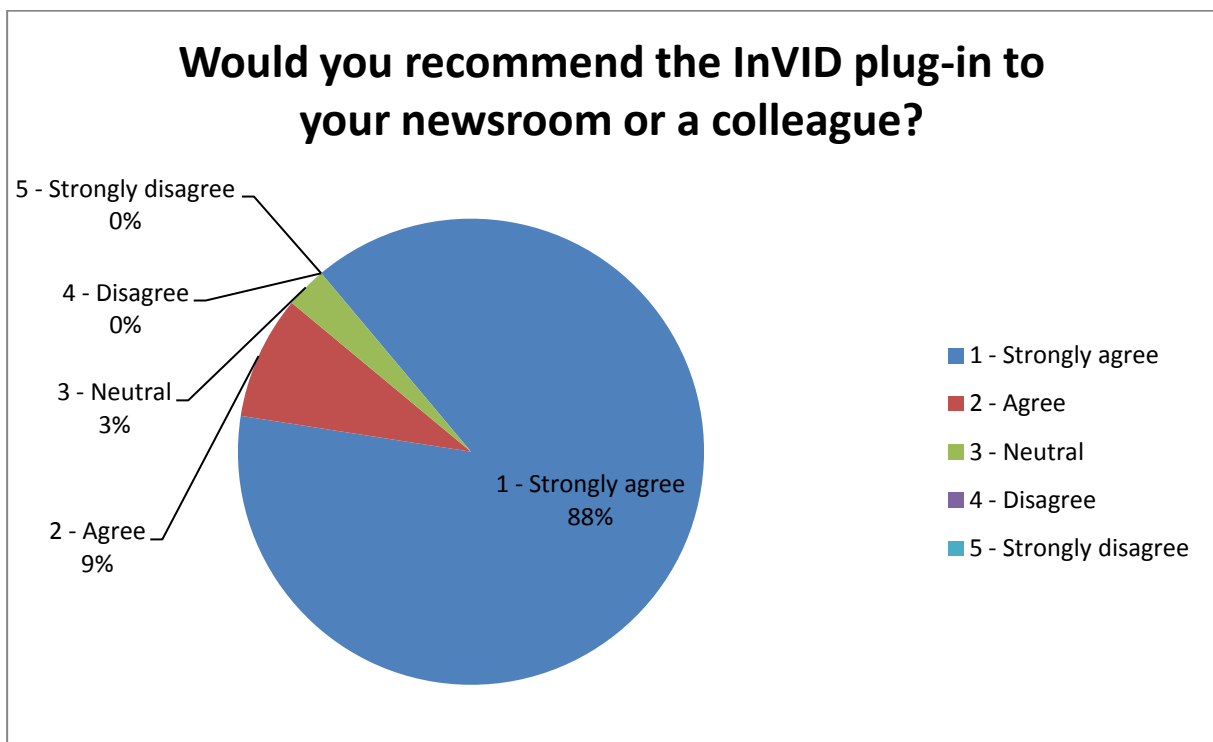


Figure 5: InVID Verification Plugin: Results for the question “Would you recommend the InVID plug-in to your newsroom or a colleague?”

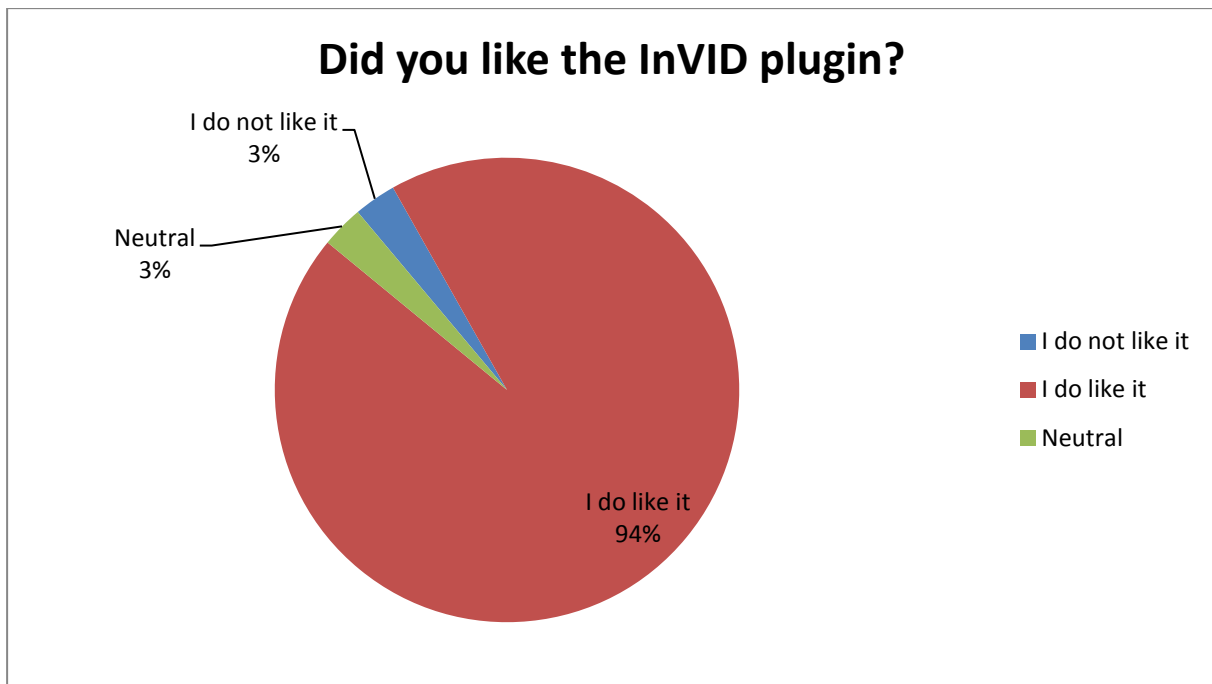


Figure 6: InVID Verification Plugin: Results for the question “Did you like the InVID plugin?”

Results of the survey for questions regarding the analysis feature (Context Aggregation & Analysis Service)

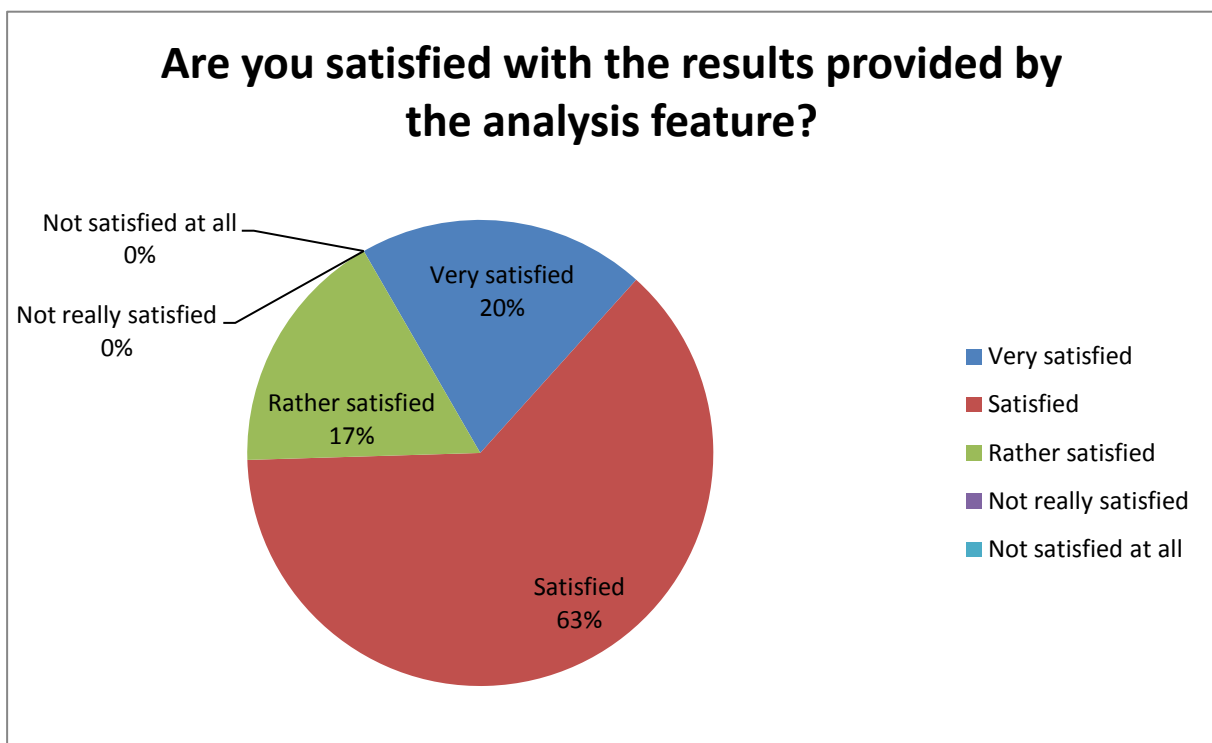


Figure 7: InVID Verification Plugin: Results for the question “Are you satisfied with the results provided by the analysis feature?”

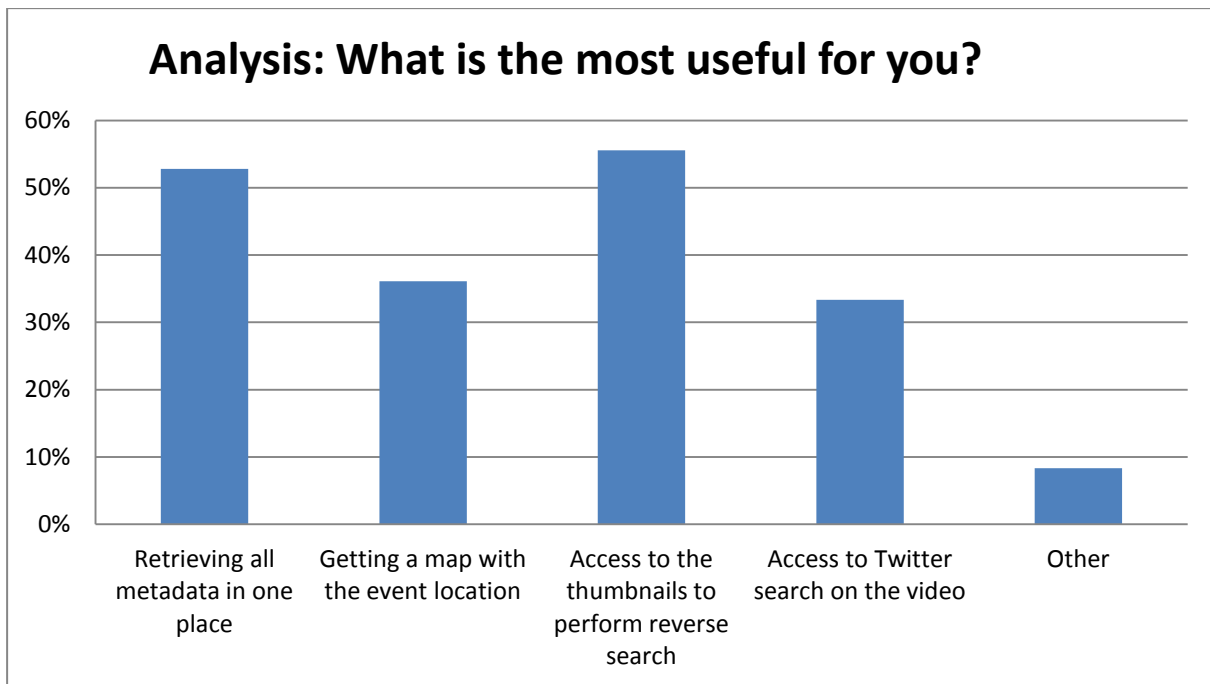


Figure 8: InVID Verification Plugin - Analysis: Results for the multiple response question “What is the most useful for you?”

Results of the survey for questions regarding the keyframe segmentation (Video Fragmentation and Annotation Service)

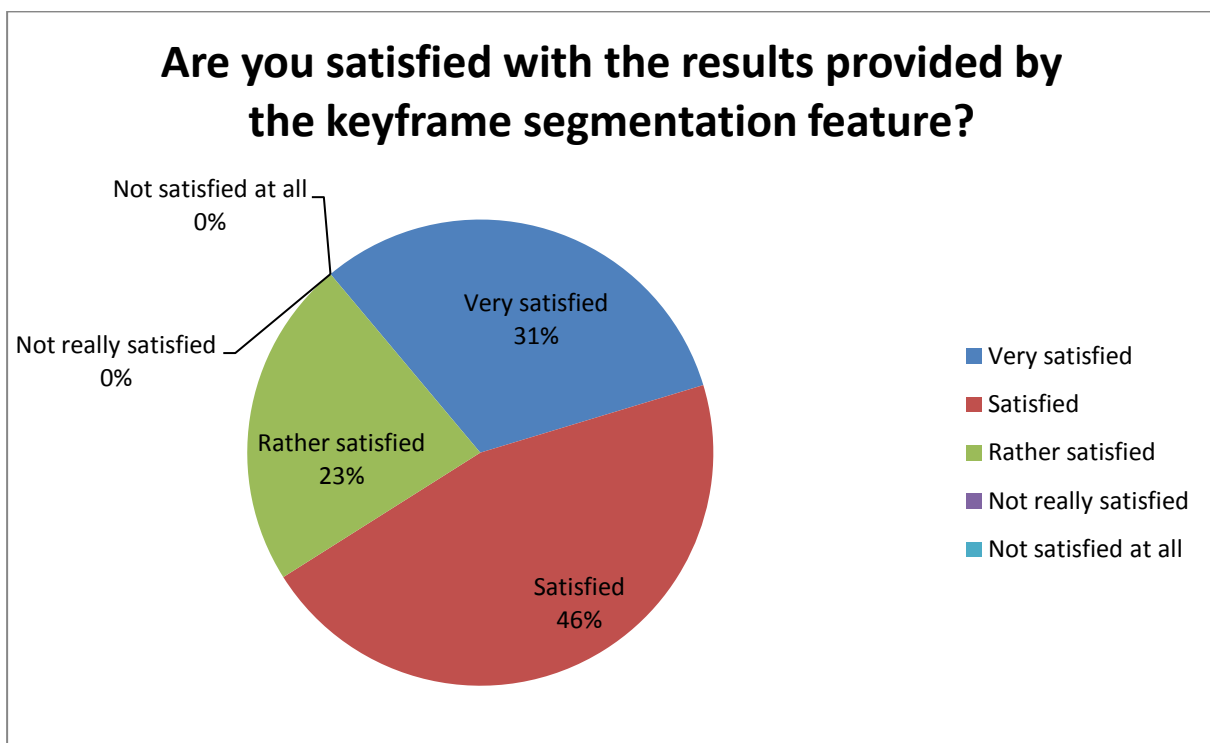


Figure 9: InVID Verification Plugin - Keyframe segmentation: Results for the question “Are you satisfied with the results provided by the keyframe segmentation feature?”

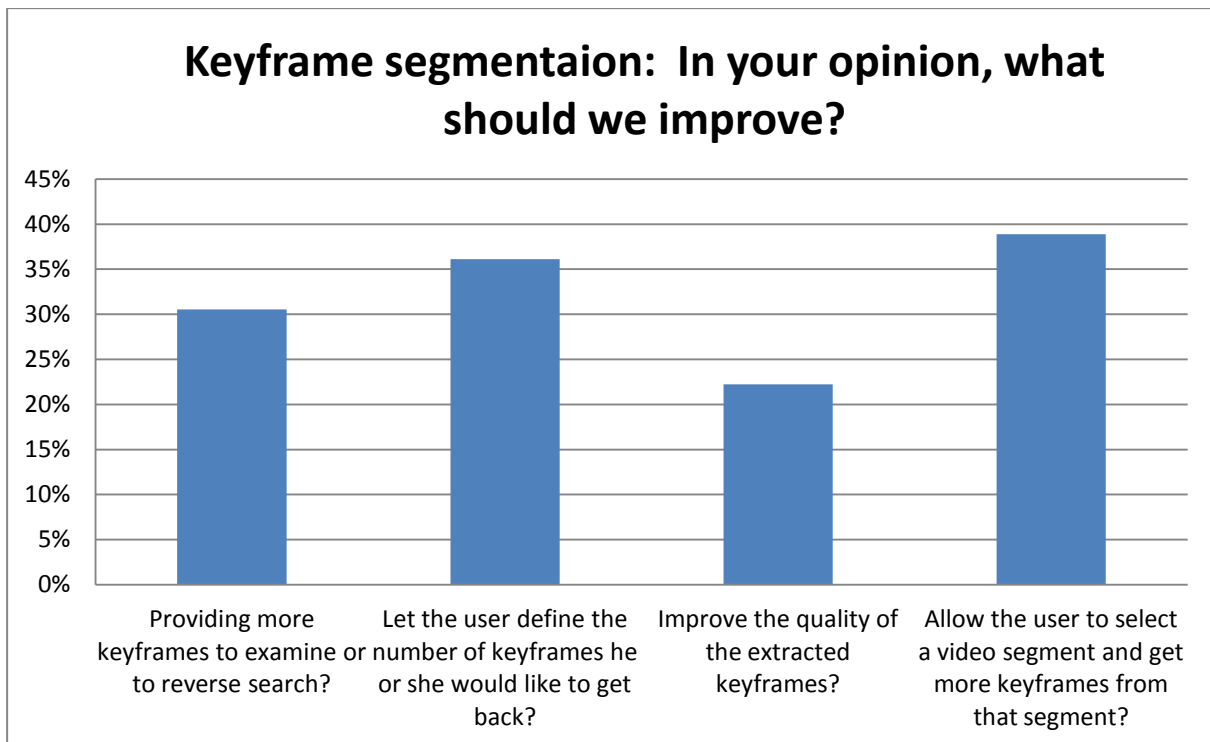


Figure 10: InVID Verification Plugin - Keyframe segmentation: Results for the multiple response question “In your opinion, what should we improve?”

Results of the survey for questions regarding the YouTube thumbnail feature

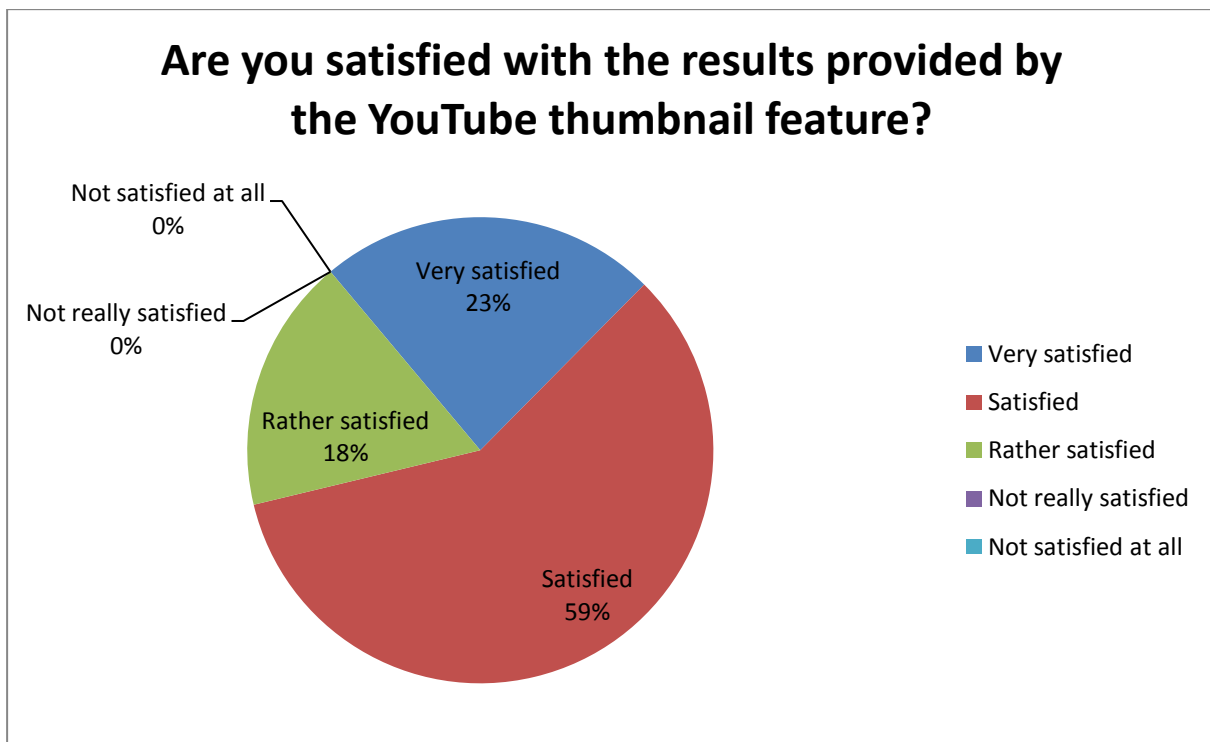


Figure 11: InVID Verification Plugin - YouTube thumbnail: Results for the question “Are you satisfied with the results provided by the YouTube thumbnail feature?”

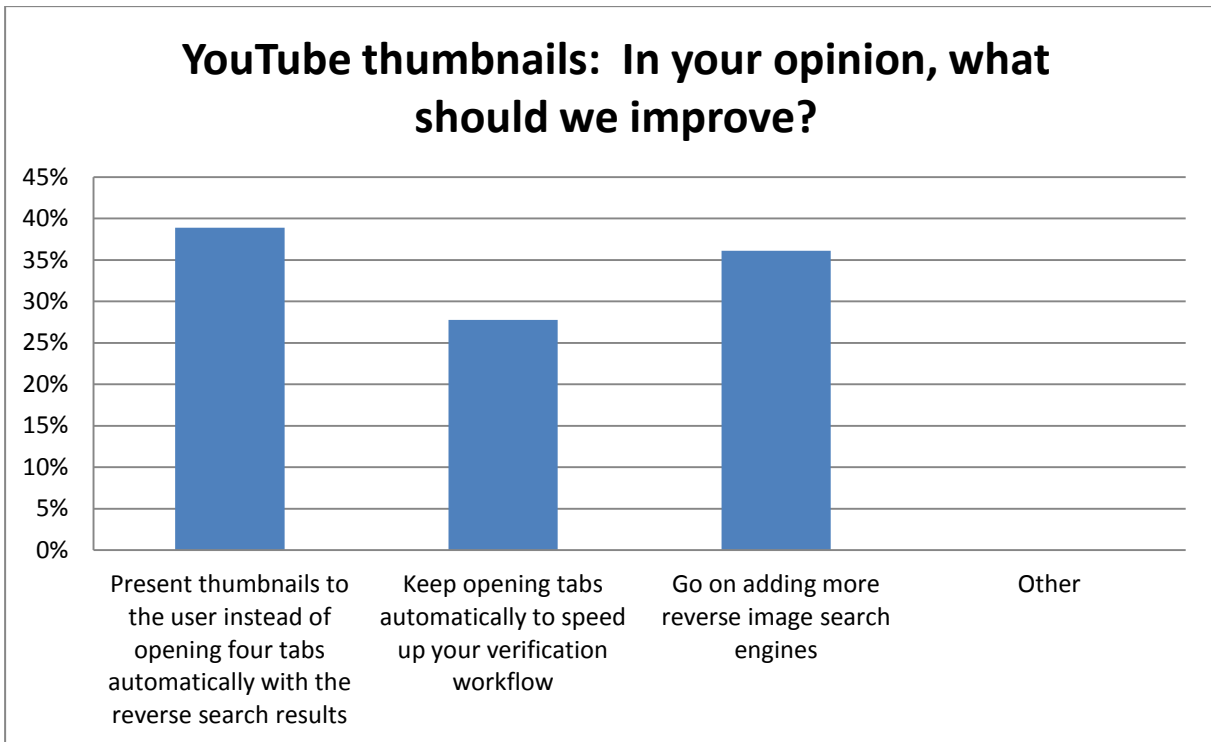


Figure 12: InVID Verification Plugin - YouTube thumbnail: Results for the multiple response question “In your opinion, what should we improve?”

Results of the survey for questions regarding the Twitter advanced search

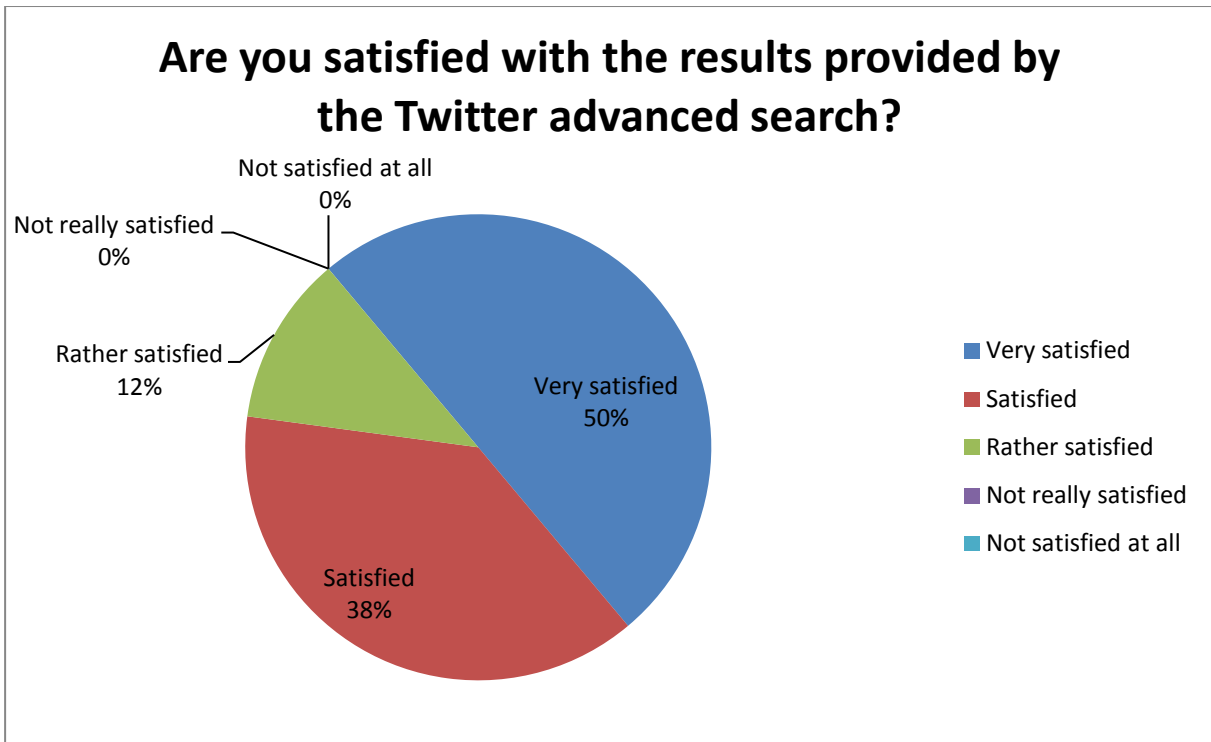


Figure 13: InVID Verification Plugin - Twitter advanced search: Results for the question “Are you satisfied with the results provided by the Twitter advanced search?”

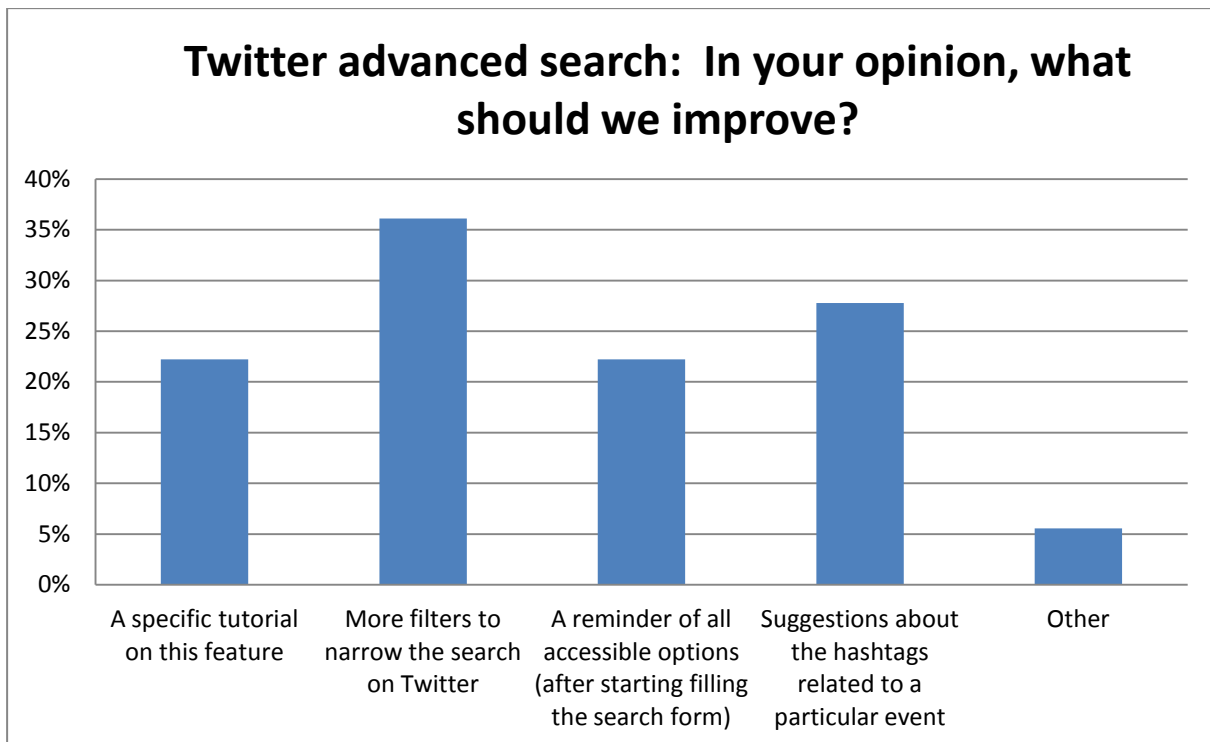


Figure 14: InVID Verification Plugin - Twitter advanced search: Results for the multiple response question “In your opinion, what should we improve?”

Results of the survey for questions regarding the image magnifier

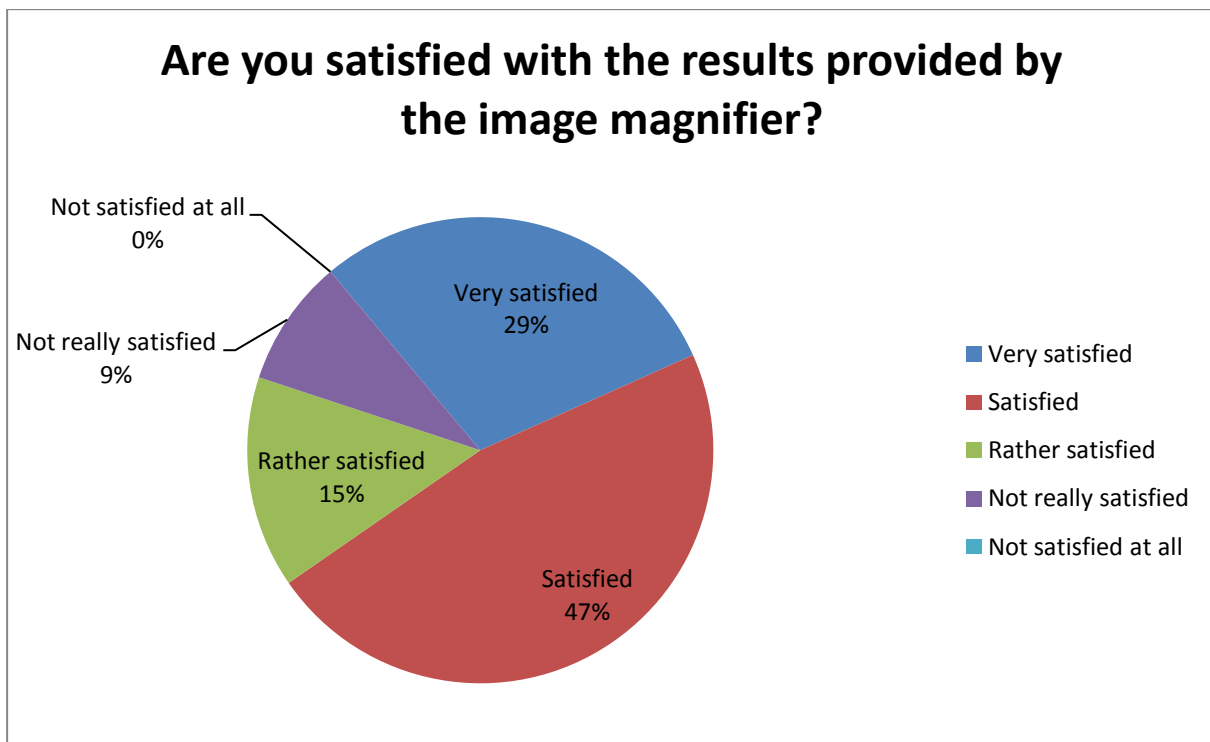


Figure 15: InVID Verification Plugin - Image magnifier: Results for the question “Are you satisfied with the results provided by the image magnifier?”

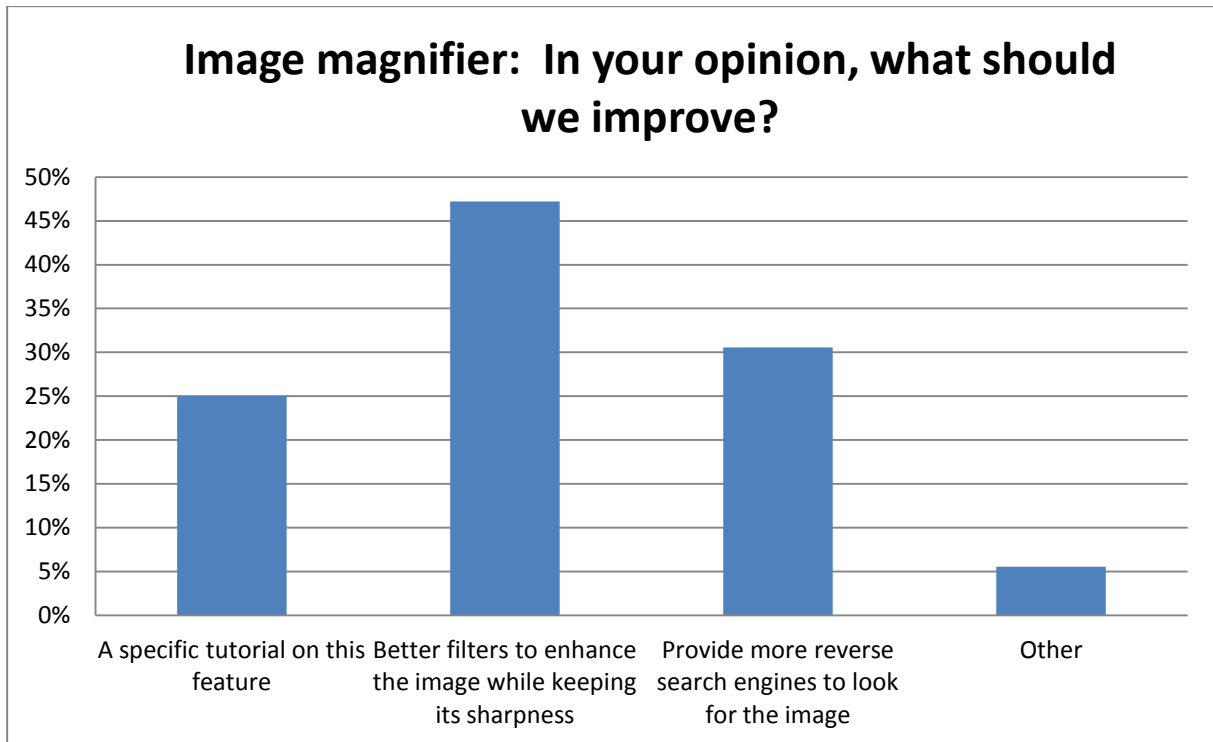


Figure 16: InVID Verification Plugin - Image magnifier: Results for the multiple response question “In your opinion, what should we improve?”

Results of the survey for questions regarding the metadata feature

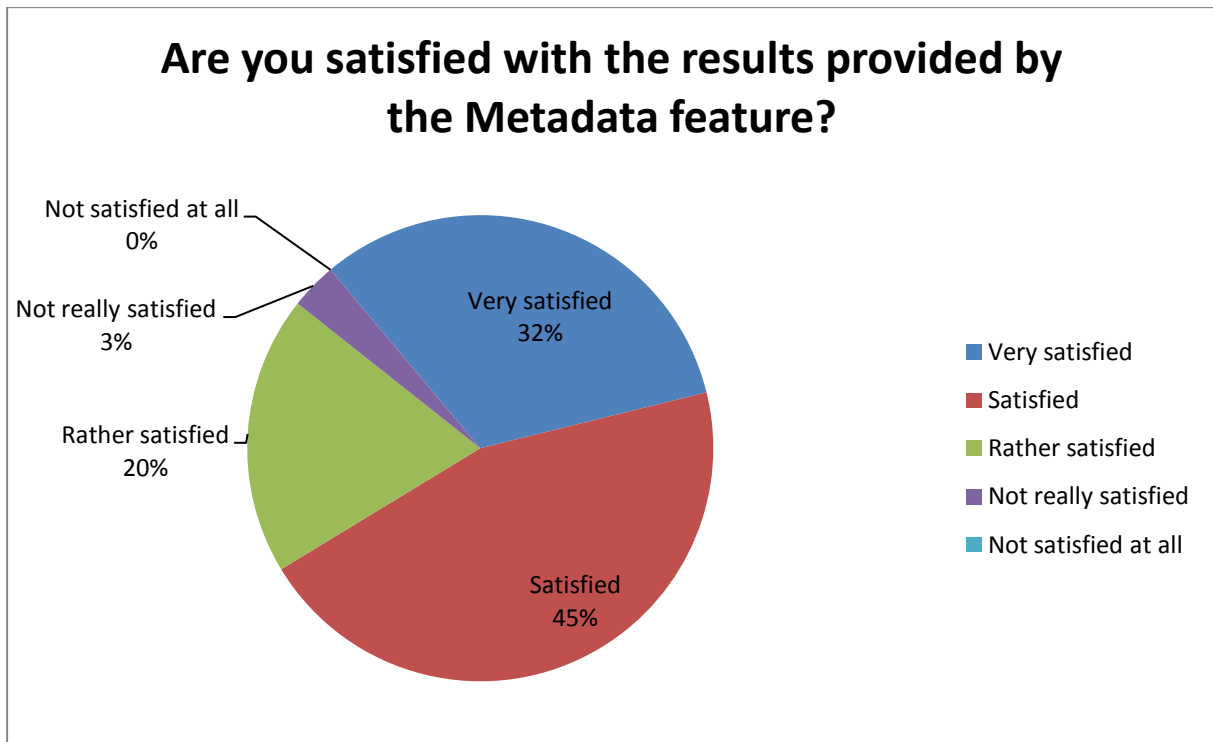


Figure 17: InVID Verification Plugin - Metadata: Results for the question “Are you satisfied with the results provided by the Metadata feature?”

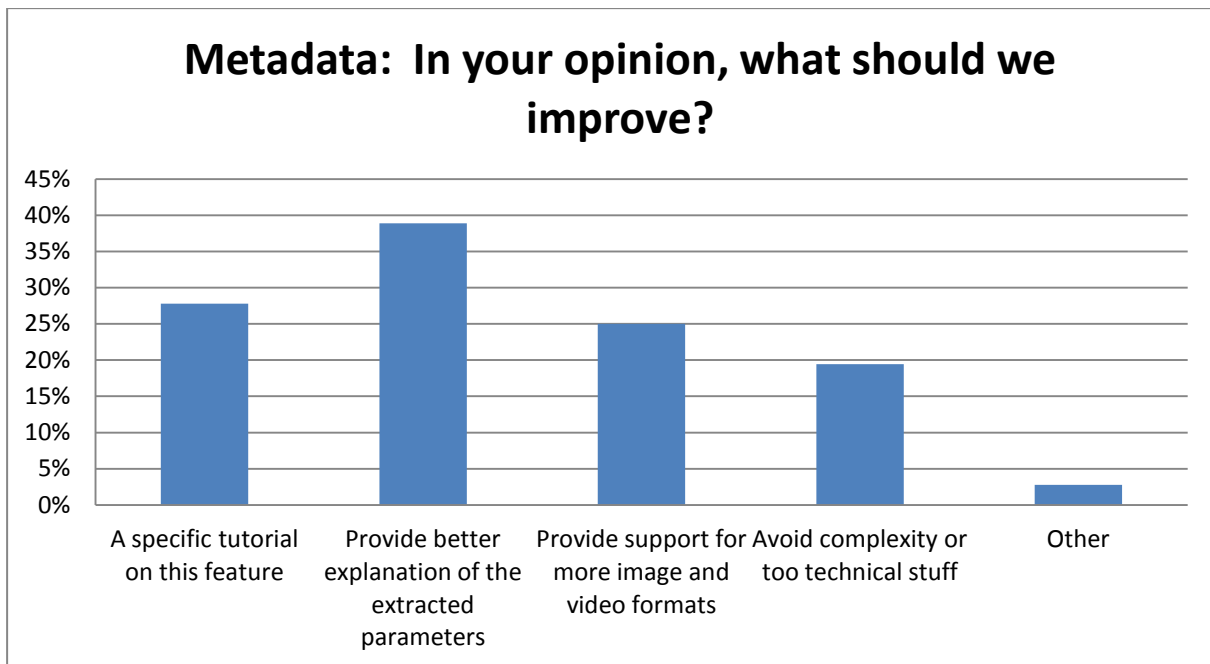


Figure 18: InVID Verification Plugin - Metadata: Results for the multiple response question “In your opinion, what should we improve?”

Results of the survey for questions regarding the picture forensic feature

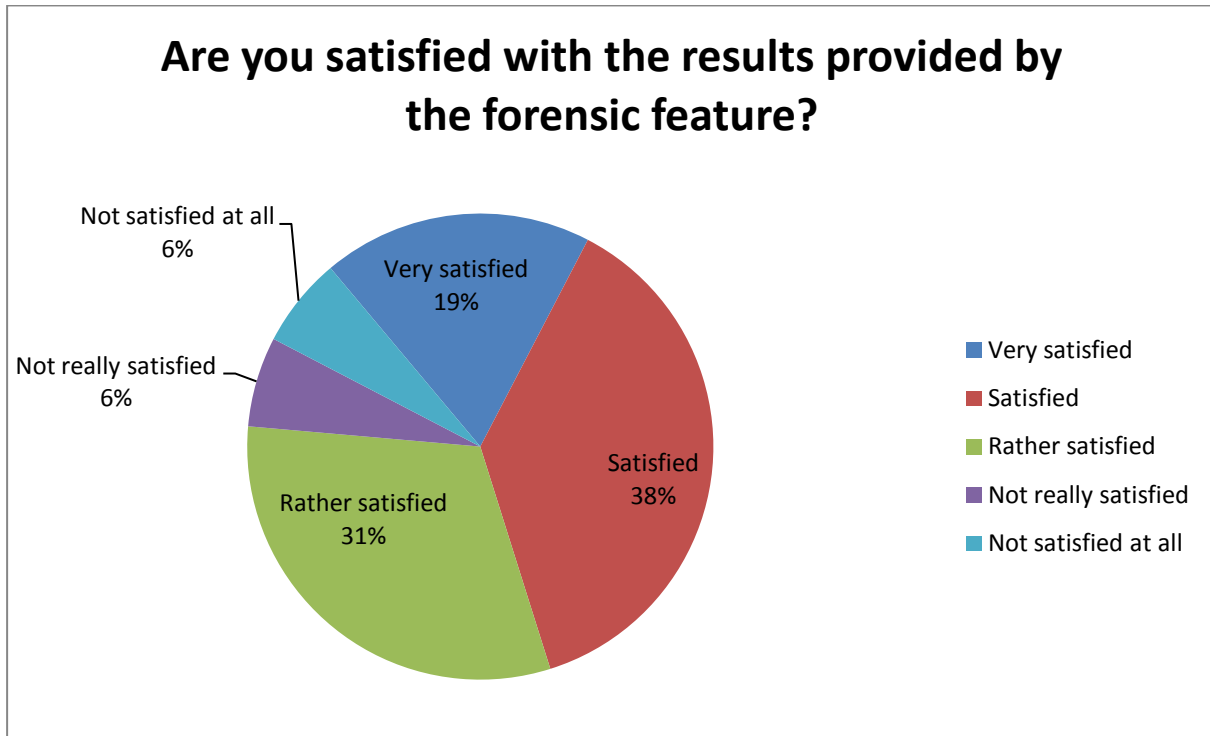


Figure 19: InVID Verification Plugin - Picture forensic: Results for the question “Are you satisfied with the results provided by the forensic feature?”

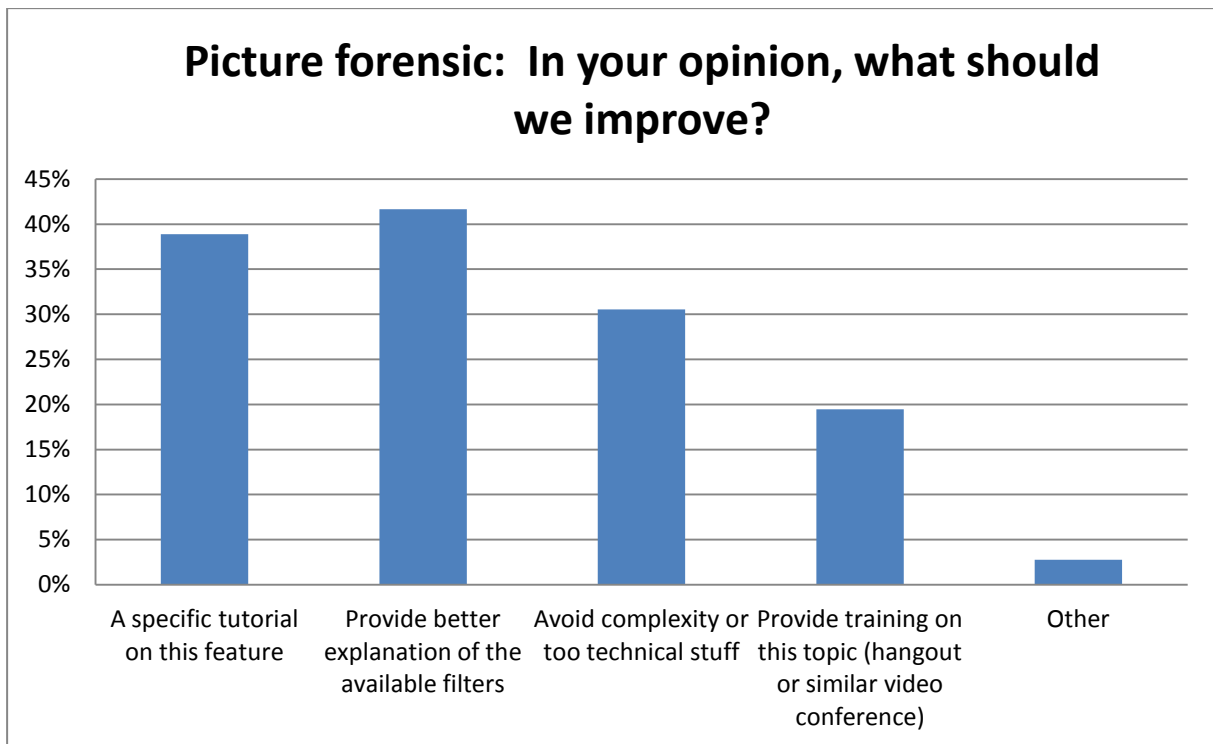


Figure 20: InVID Verification Plugin - Picture forensic: Results for the multiple response question “In your opinion, what should we improve?”

3.11 InVID Verification Application

3.11.1 Description of the service

The InVID Verification Application is a complete web-based technology that enables journalists to assess the reliability and trustworthiness of User Generated Videos, with the help of several integrated analysis components.

3.11.2 Tests

The InVID Verification Application was tested in test cycles 4, 5 and 6. Members of the consortium tested the application in all three test cycles. Tests with users external to the consortium were done in test cycle 5. These external testers include experts from Deutsche Welle that do not participate to the InVID project, as well as journalists from regional newspapers, such as Tiroler Tageszeitung and Vorarlberger Nachrichten.

In test cycle 5 the users' feedback was collected with the help of an online survey that contained 17 questions regarding the usability and functionality of the application. The survey was filled in by 3 members of the consortium ("internal") and 7 testers from outside of the consortium ("external"). Two external testers filled in one survey entry together, hence, the number of survey responses for external testers is 6.

In test cycle 6 a dedicated full-day session between the developers of this integrated technology and the internal testers of it, was held to give structured feedback for improvements on the functionality and usability of the application.

Table 16: Number of received feedback comments for the InVID Verification Application

Test cycle	Feedback comments	Survey responses
Test cycle 4	5 items	
Test cycle 5	42 items	9 survey responses from 10 testers
Test cycle 6	17 items	

3.11.3 Major outcomes of the survey

The following diagrams give an overview of the findings of the survey on the testers of the application. The result for each question is separated in two diagrams; one for the external testers (7 testers with 6 survey responses) and one for the internal testers (3 testers).

The outcomes about the efficiency of this technology are really positive with an exception about the near duplicate component of it. This, not so good, assessment was due to the relatively small index of videos that is utilized for near duplicate search (compared to the vast amount of videos available on the YouTube and other video sharing platforms). To address this, an additional feature was implemented, that automatically performs a video-title-based

search on YouTube, and indexes similar videos for further processing and assessment by the near duplicate component.

Results of the survey for general questions

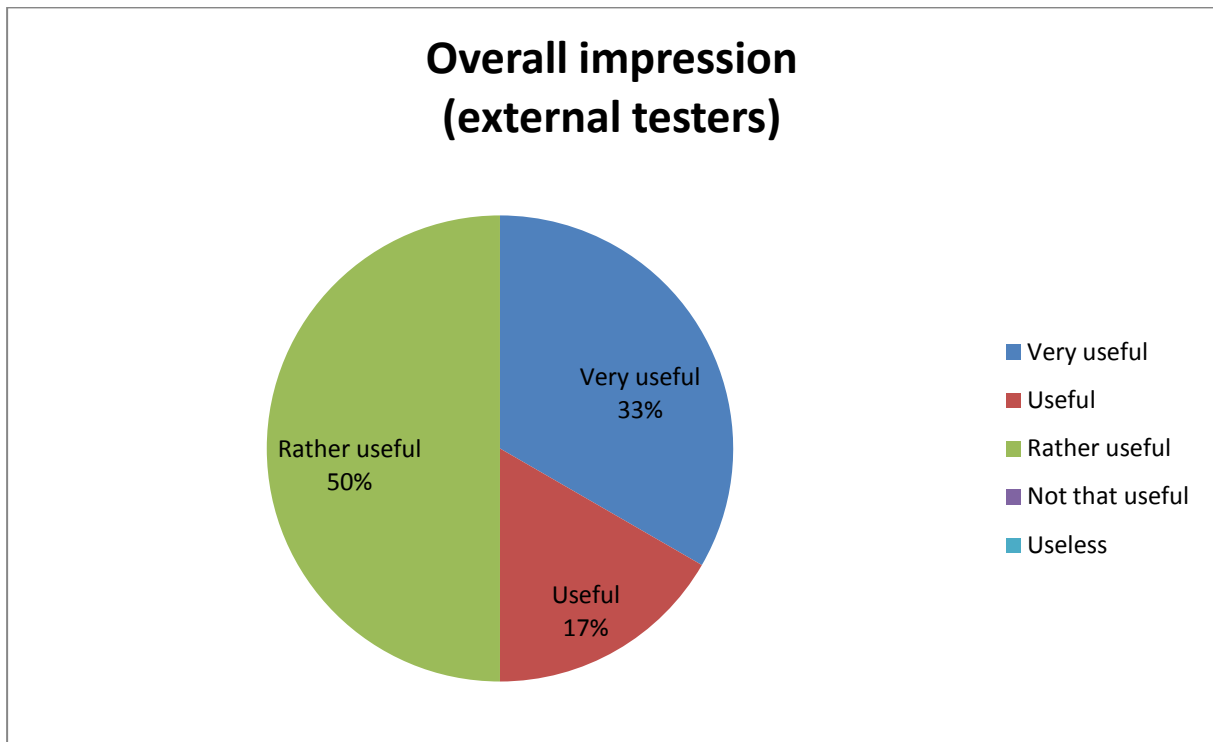


Figure 21: InVID Verification Application – external testers: Overall impression

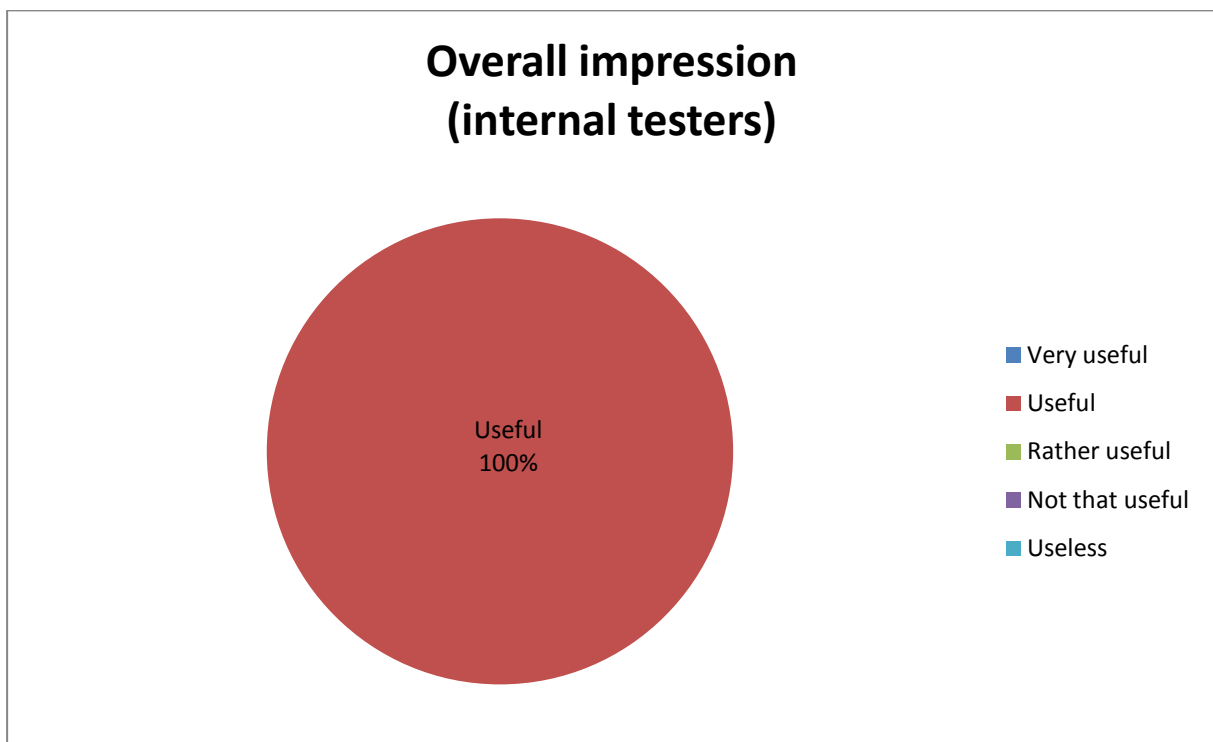


Figure 22: InVID Verification Application – internal testers: Overall impression

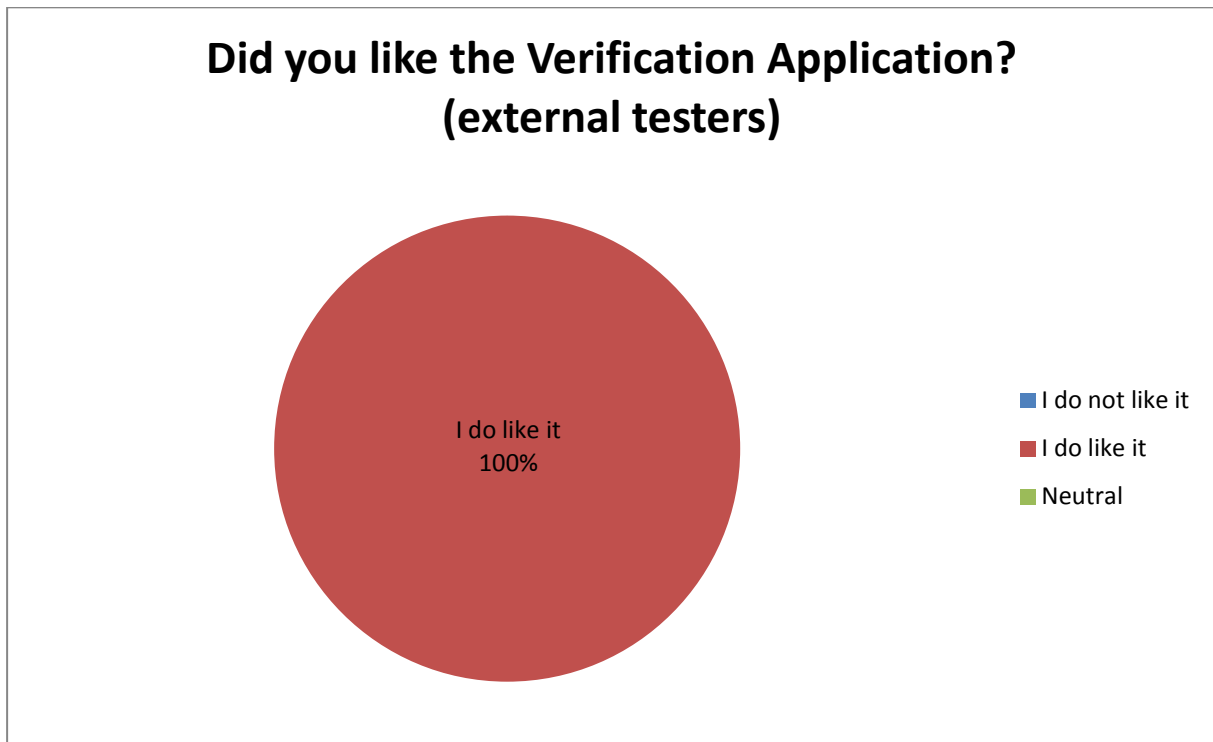


Figure 23: InVID Verification Application – external testers: Results for the question “Did you like the InVID plugin?”

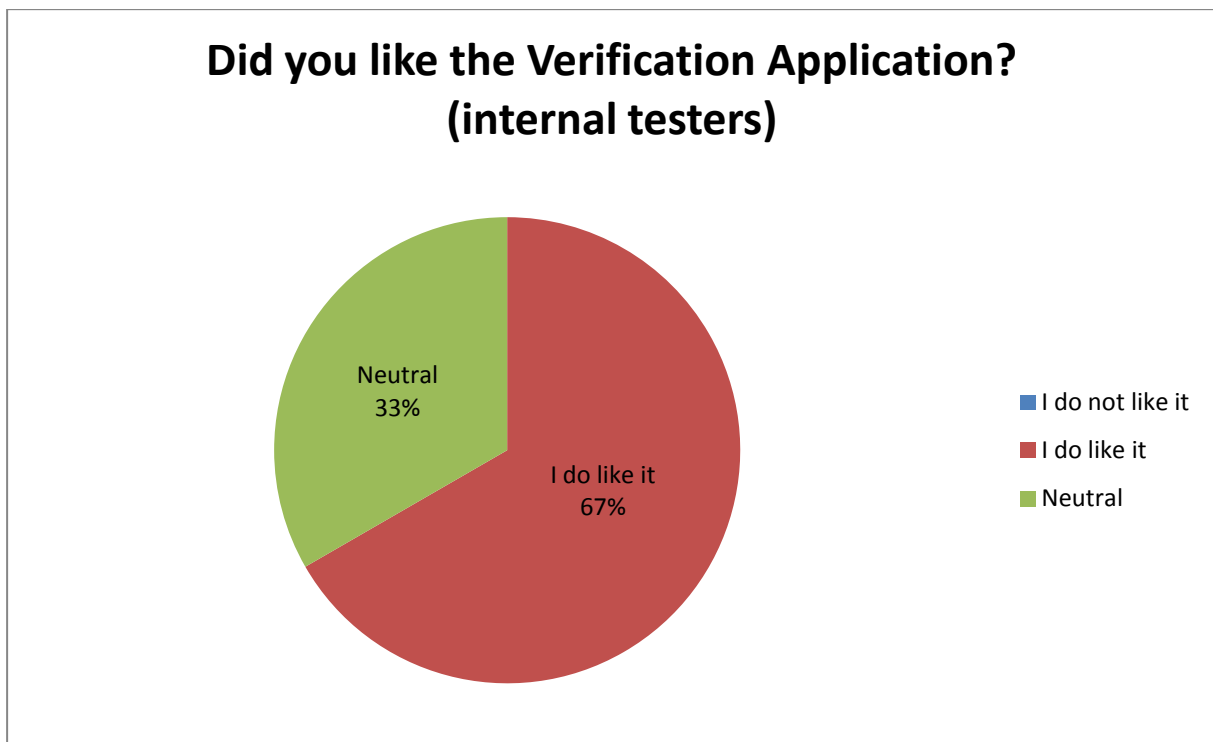


Figure 24: InVID Verification Application – internal testers: Results for the question “Did you like the InVID plugin?”

Results of the survey for questions about usability

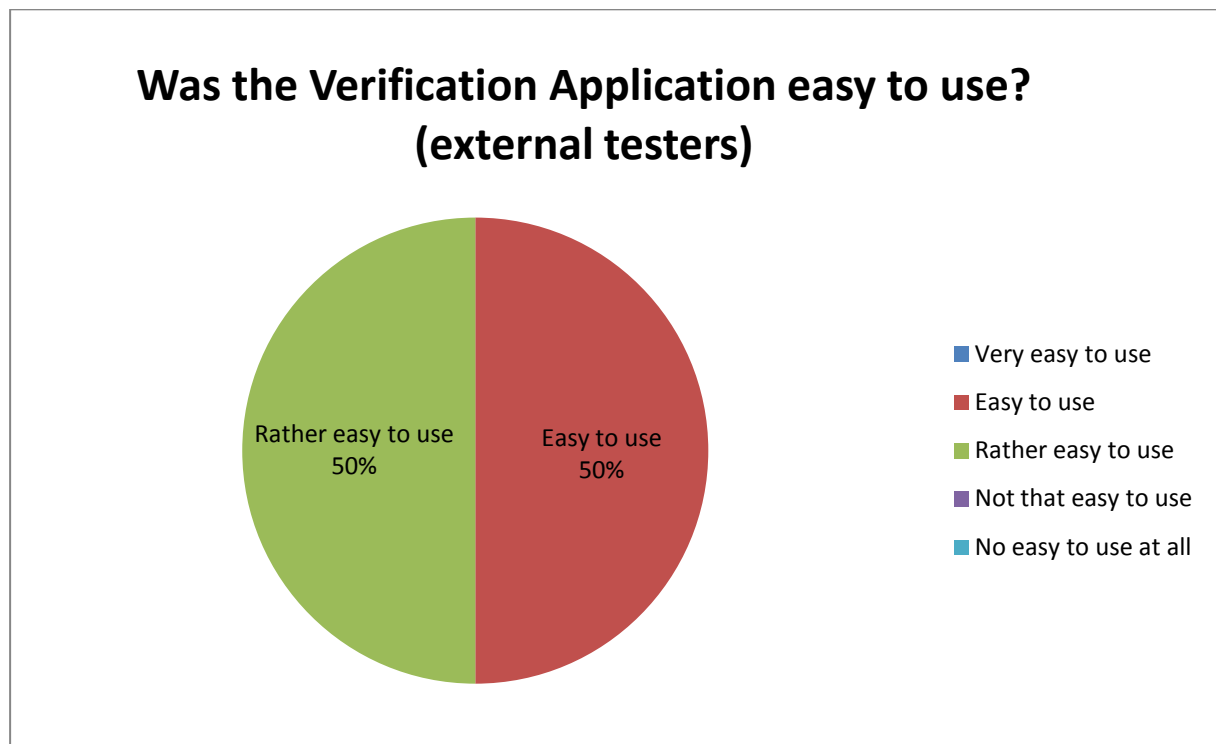


Figure 25: InVID Verification Application – external testers: Results for the question “Was the Verification Application easy to use?”

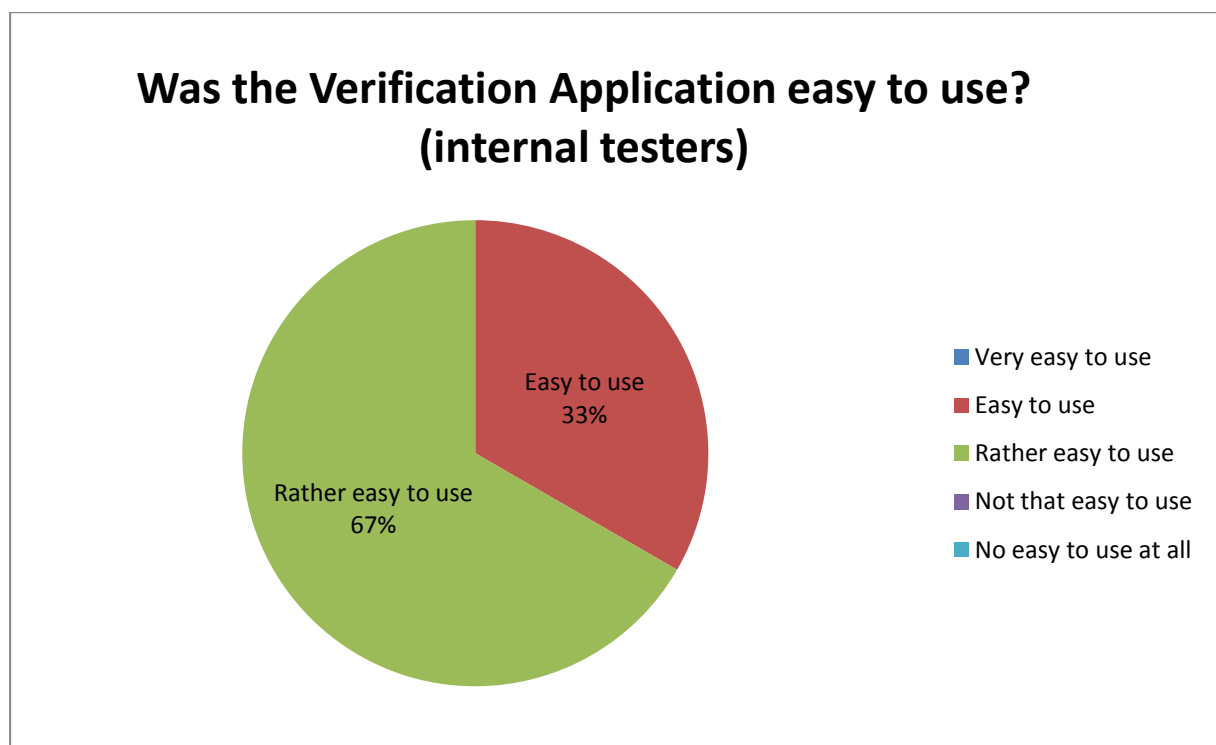


Figure 26: InVID Verification Application – internal testers: Results for the question “Was the Verification Application easy to use?”

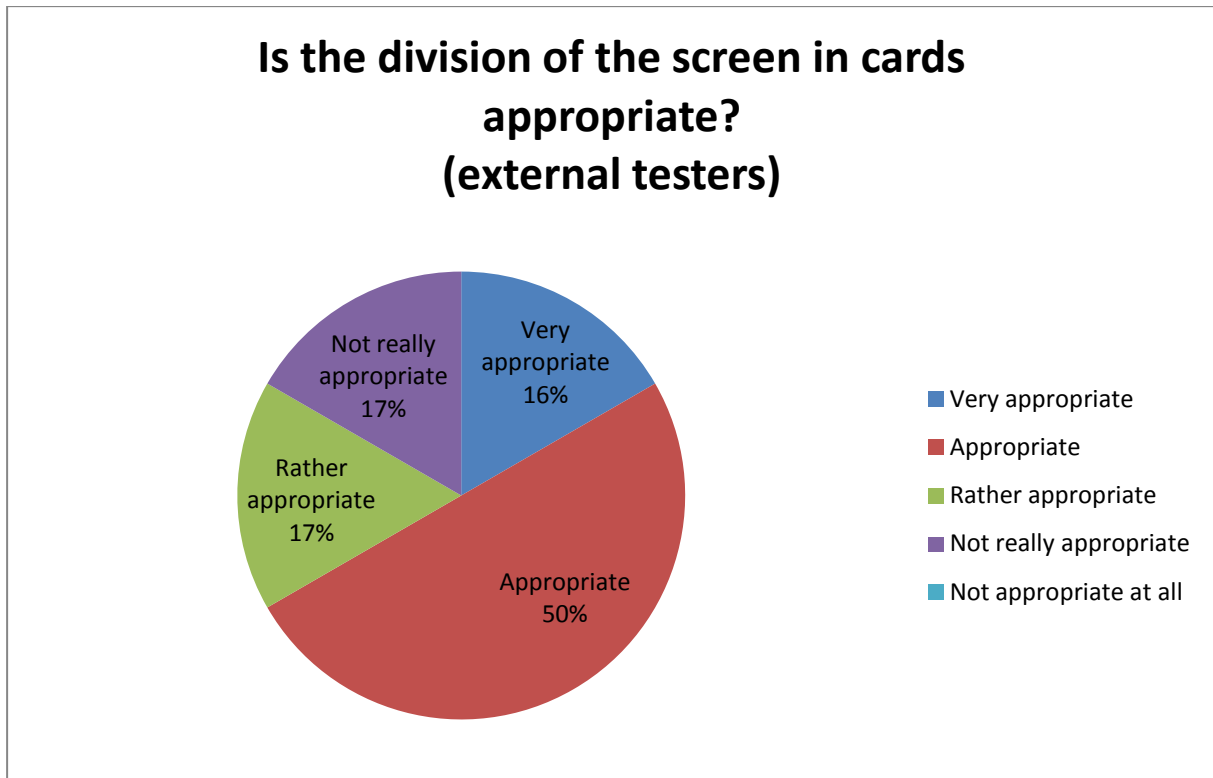


Figure 27: InVID Verification Application – external testers: Results for the question “Is the division of the screen in cards appropriate?”

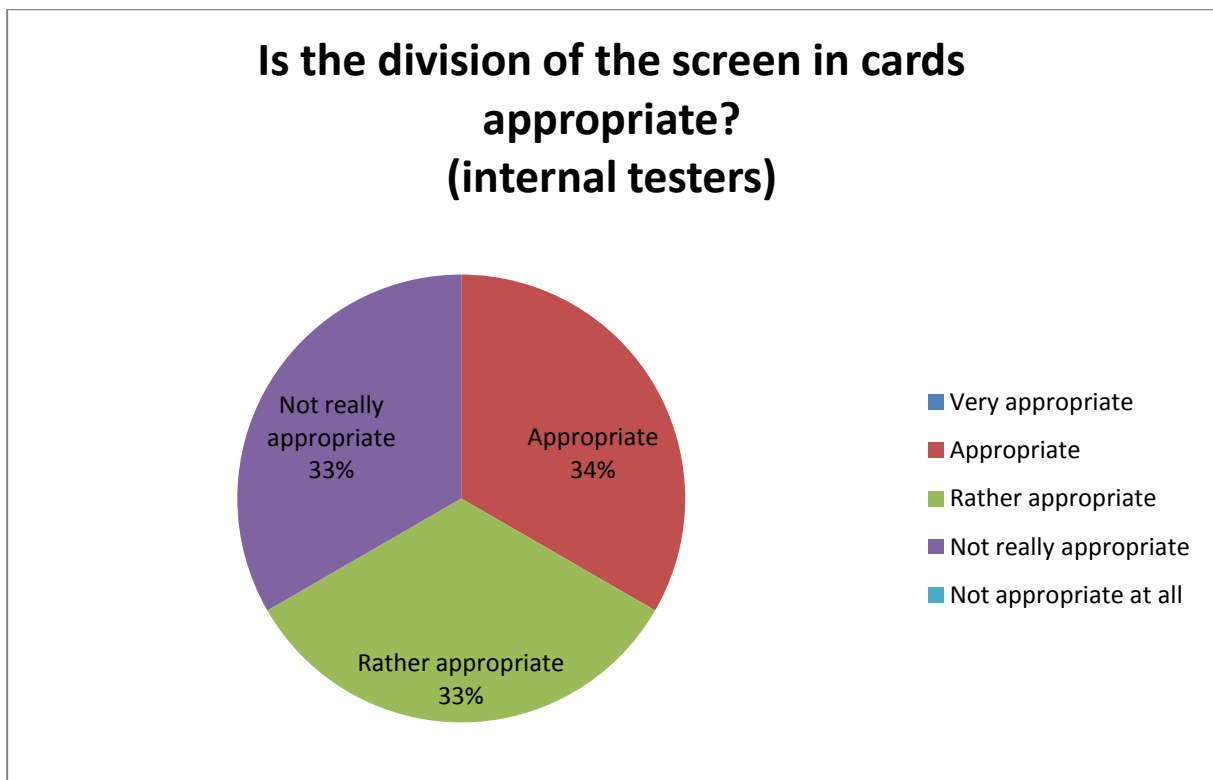


Figure 28: InVID Verification Application – internal testers: Results for the question “Is the division of the screen in cards appropriate?”

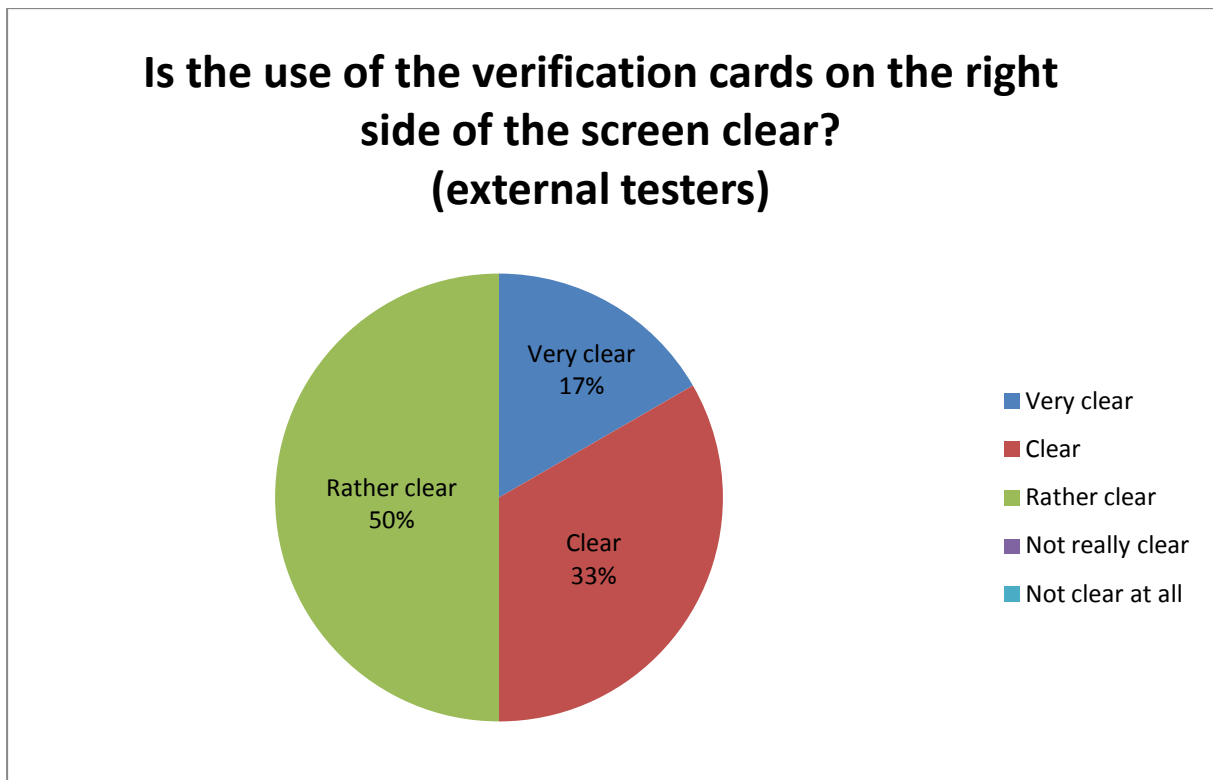


Figure 29: InVID Verification Application – external testers: Results for the question “Is the use of the verification cards on the right side of the screen clear?”

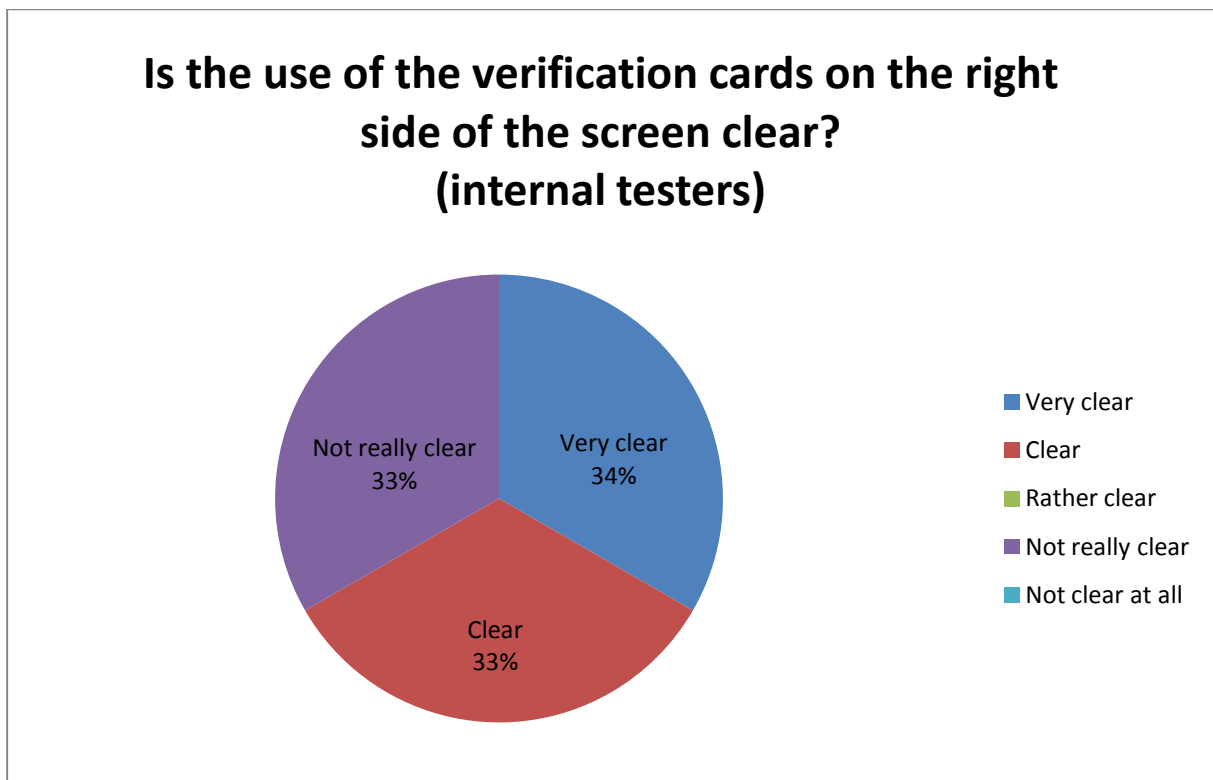


Figure 30: InVID Verification Application – internal testers: Results for the question “Is the use of the verification cards on the right side of the screen clear?”

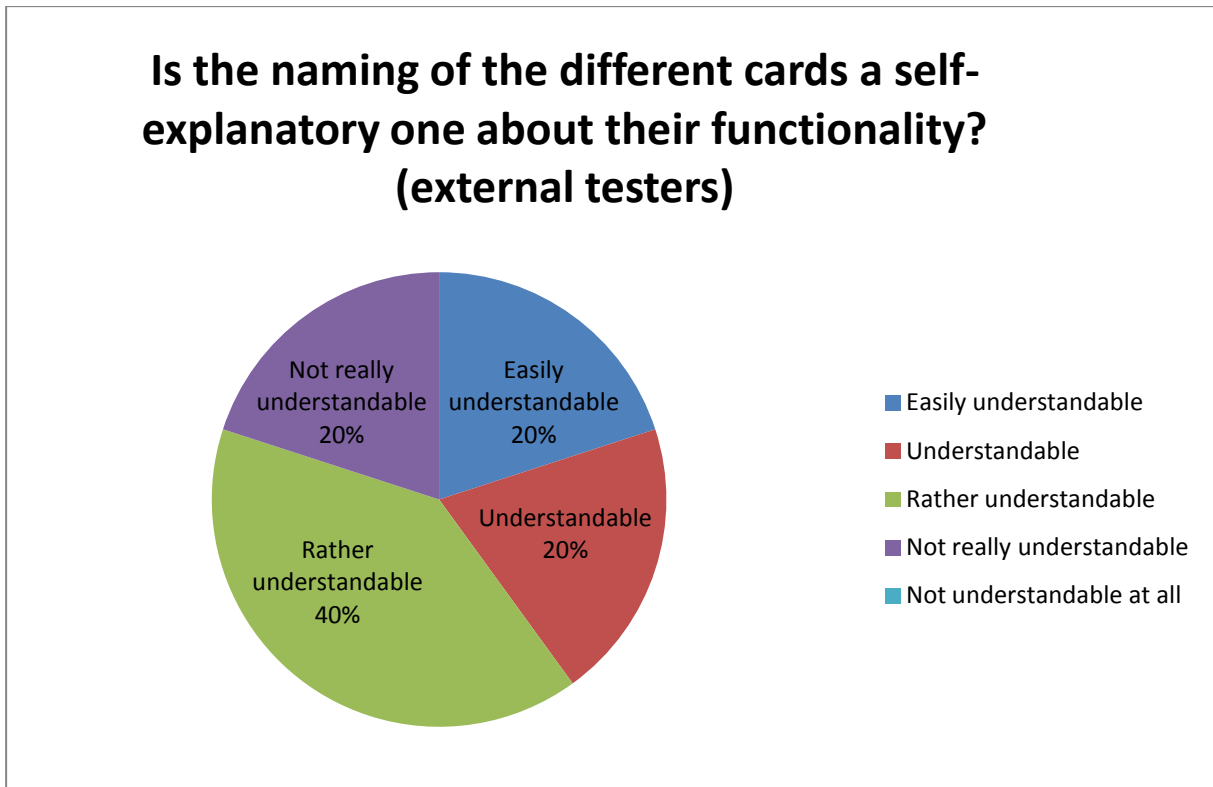


Figure 31: InVID Verification Application – external testers: Results for the question “Is the naming of the different cards self-explanatory about their functionality?”

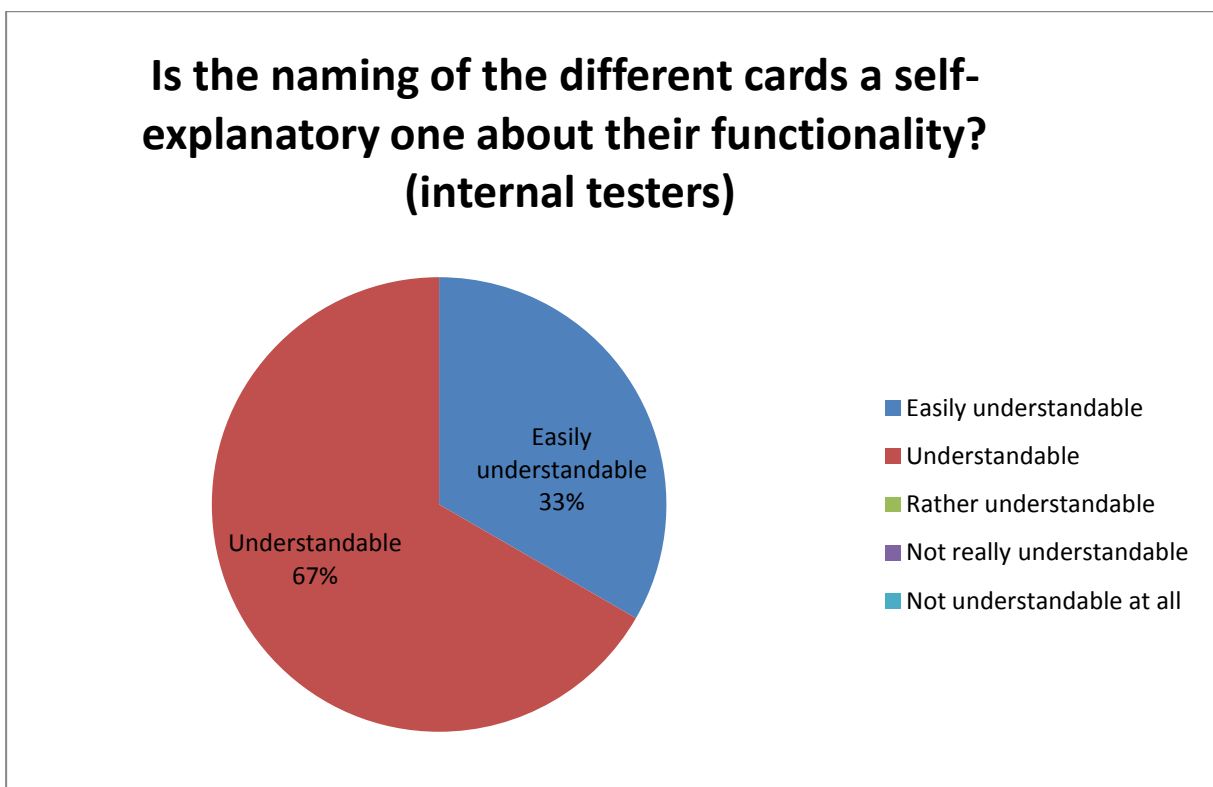


Figure 32: InVID Verification Application – internal testers: Results for the question “Is the naming of the different cards self-explanatory about their functionality?”

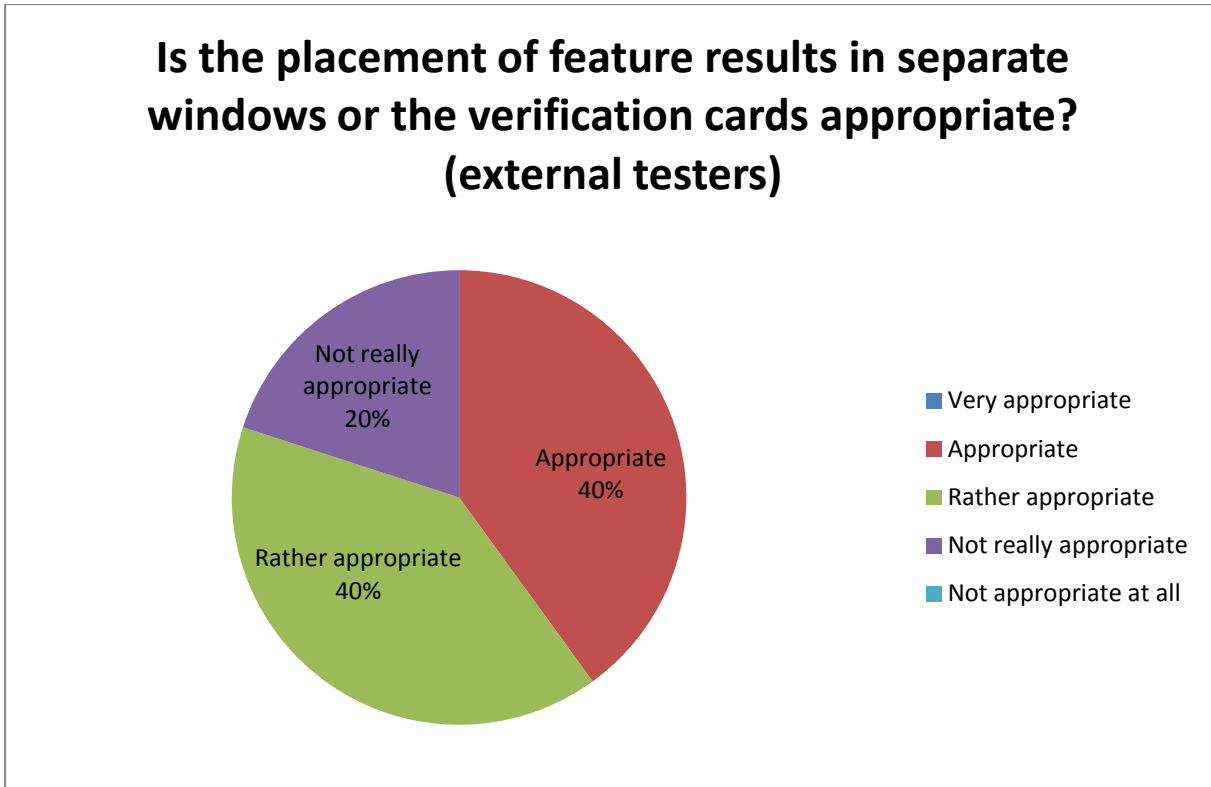


Figure 33: InVID Verification Application – external testers: Results for the question “Is the placement of feature results in separate windows or the verification cards appropriate?”

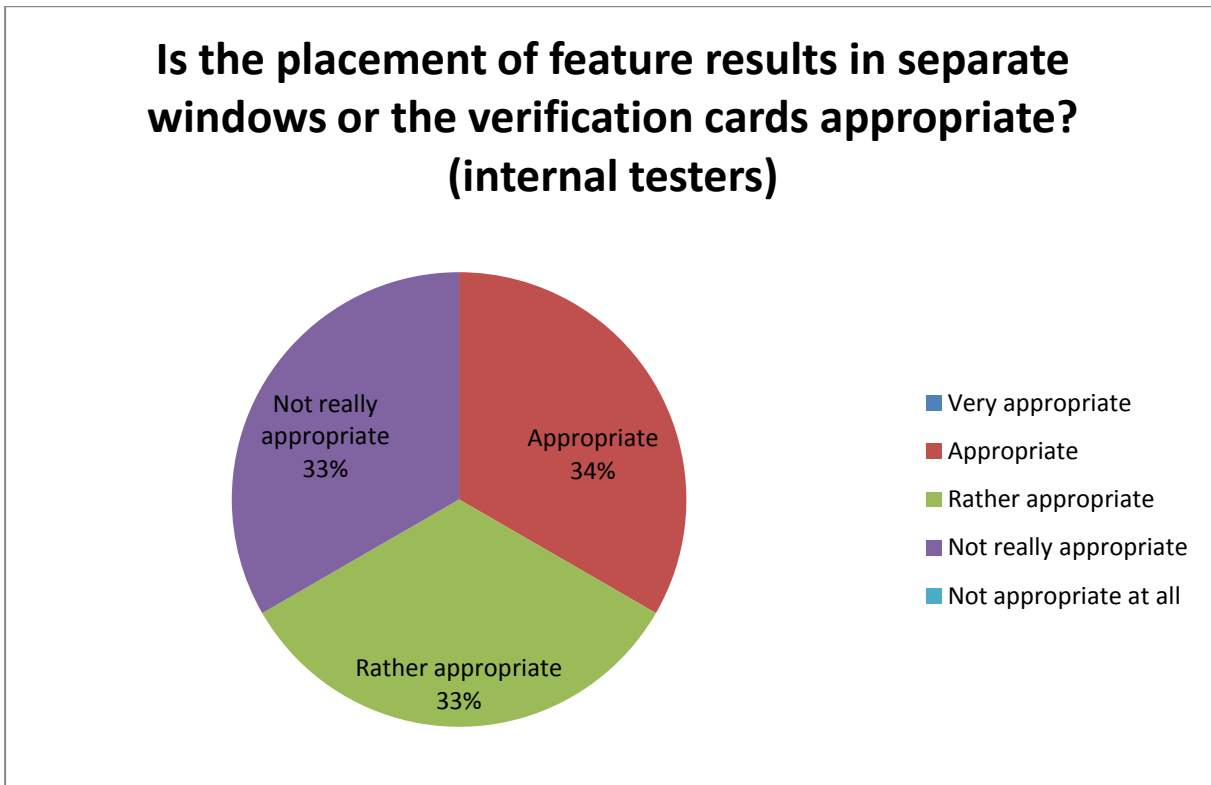


Figure 34: InVID Verification Application – internal testers: Results for the question “Is the placement of feature results in separate windows or the verification cards appropriate?”

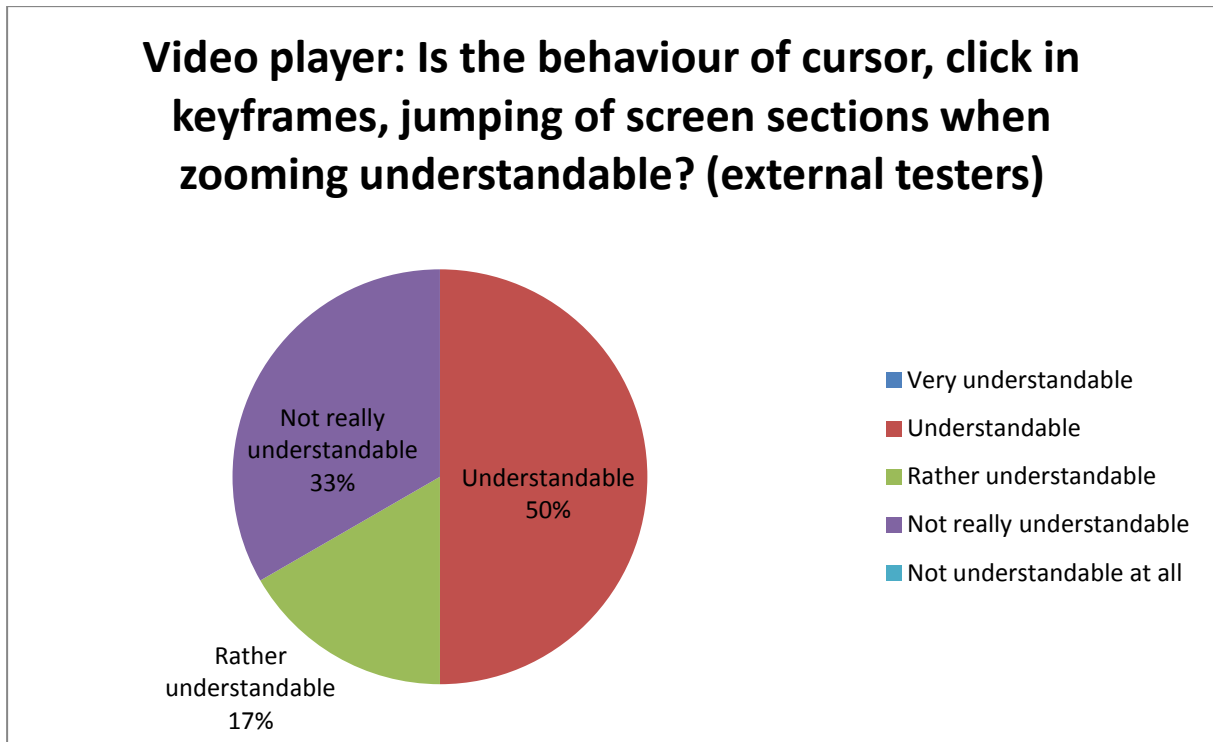


Figure 35: InVID Verification Application – external testers: Results for the question “Video player: Is the behaviour of cursor, click in keyframes, jumping of screen sections when zooming understandable?”

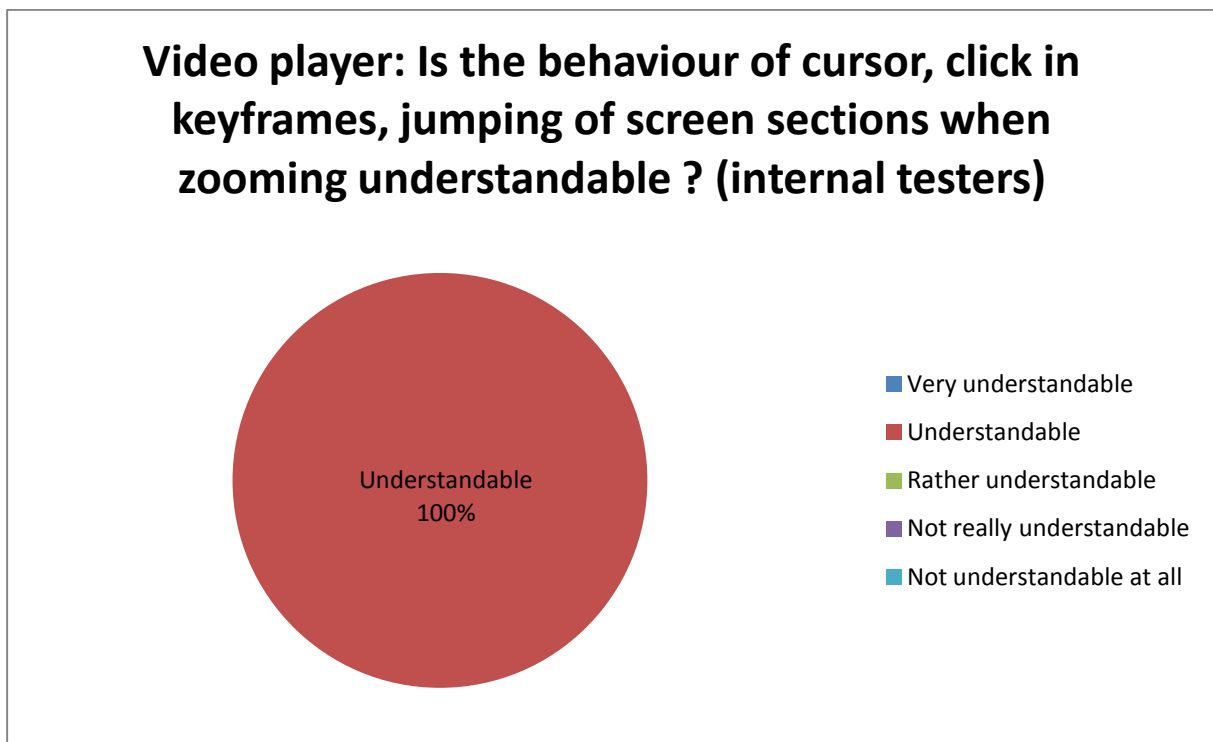


Figure 36: InVID Verification Application – internal testers: Results for the question “Video player: Is the behaviour of cursor, click in keyframes, jumping of screen sections when zooming understandable?”

Results of the survey regarding the near duplicates feature

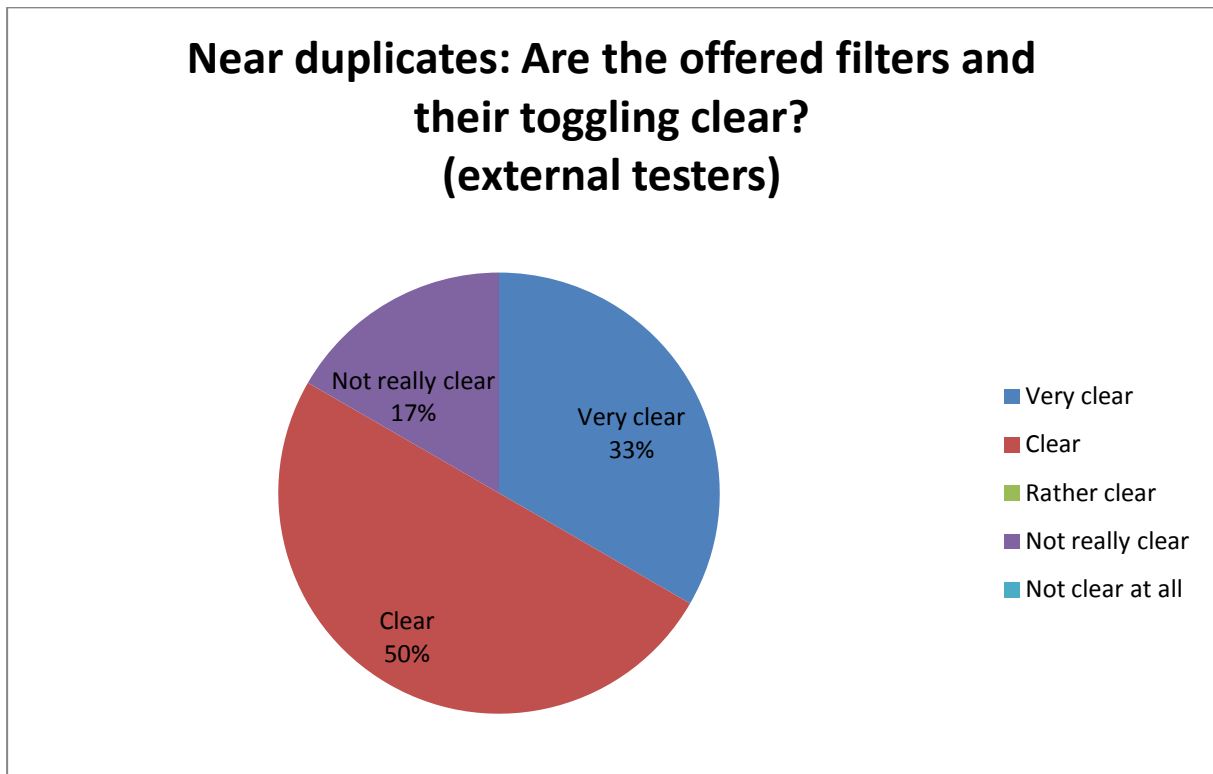


Figure 37: InVID Verification Application – external testers: Near duplicates: Results for the question “Are the offered filters and their toggling clear?”

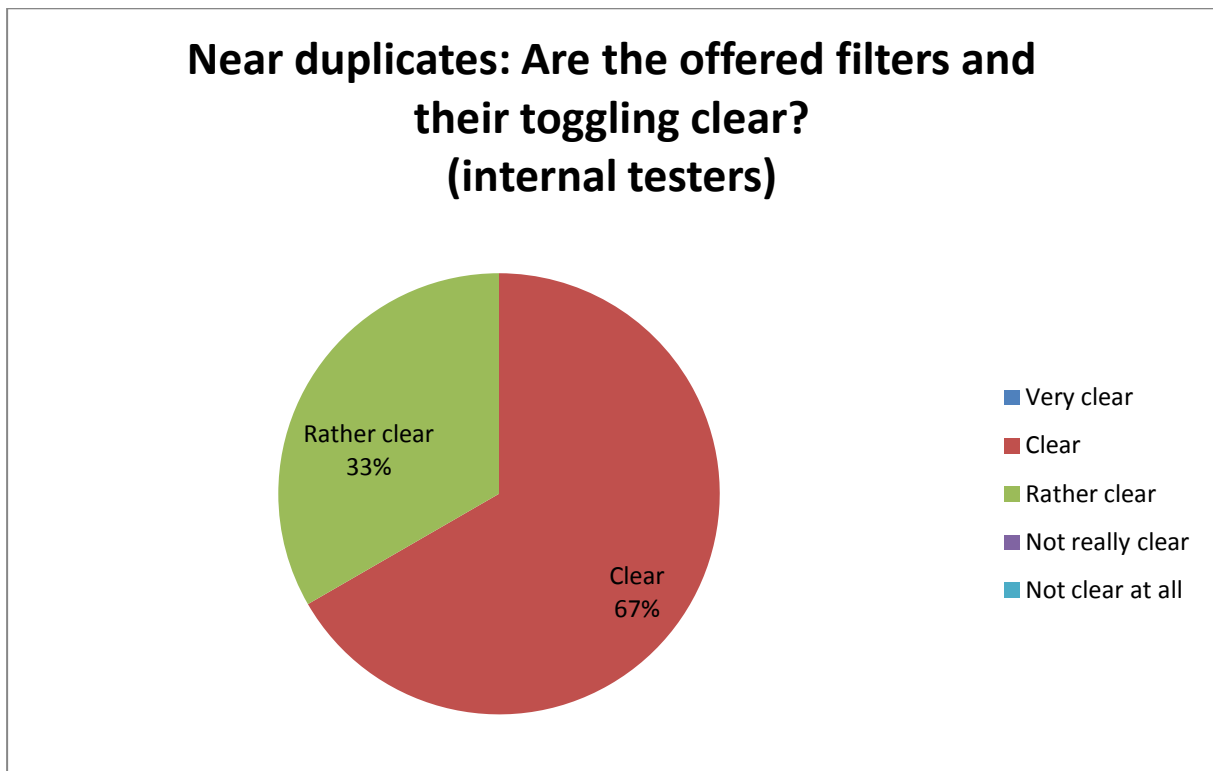


Figure 38: InVID Verification Application – internal testers: Near duplicates: Results for the question “Are the offered filters and their toggling clear?”

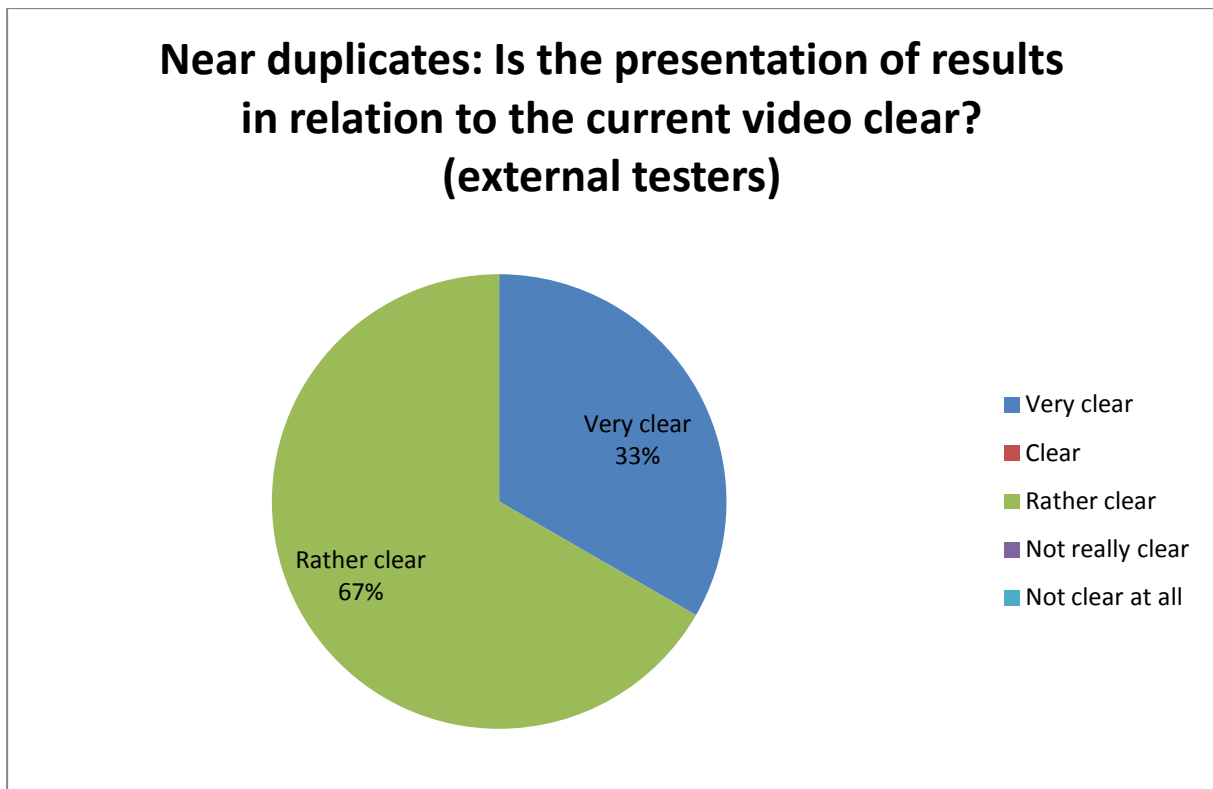


Figure 39: InVID Verification Application – external testers: Near duplicates: Results for the question “Is the presentation of results in relation to the current video clear?”

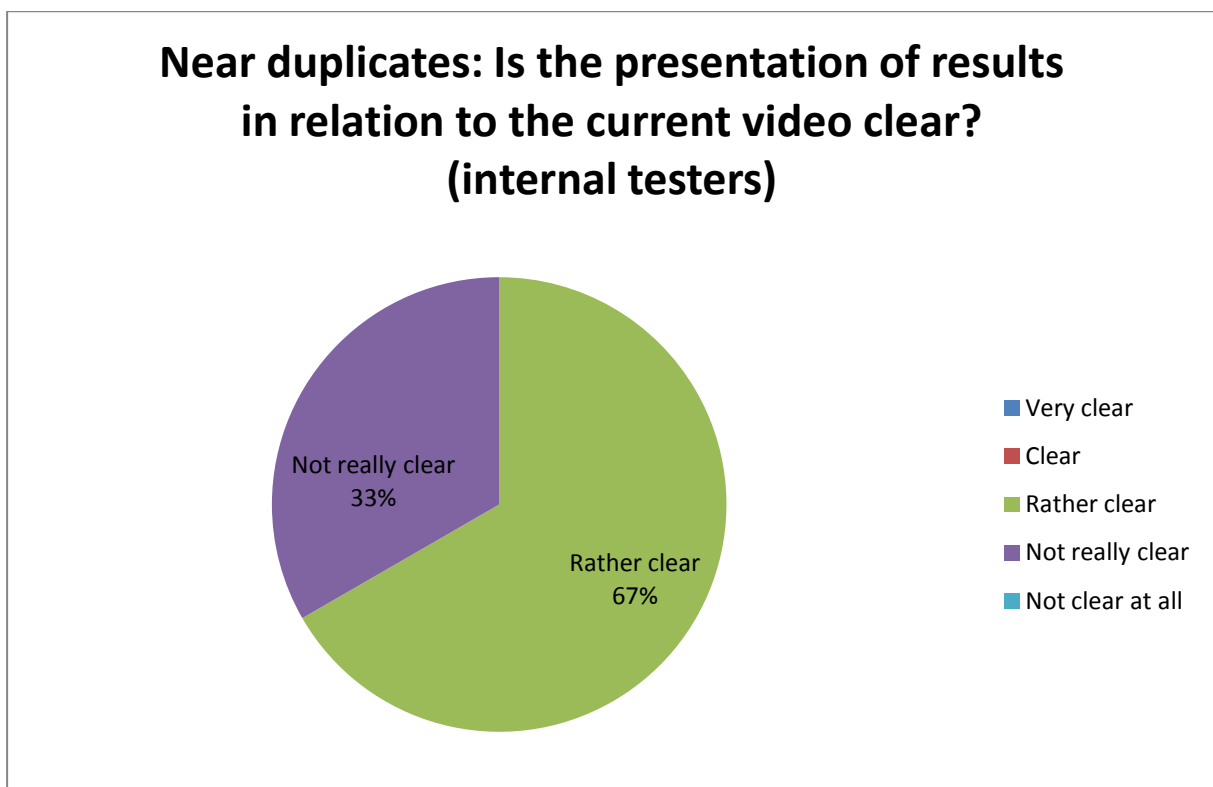


Figure 40: InVID Verification Application – internal testers: Near duplicates: Results for the question “Is the presentation of results in relation to the current video clear?”

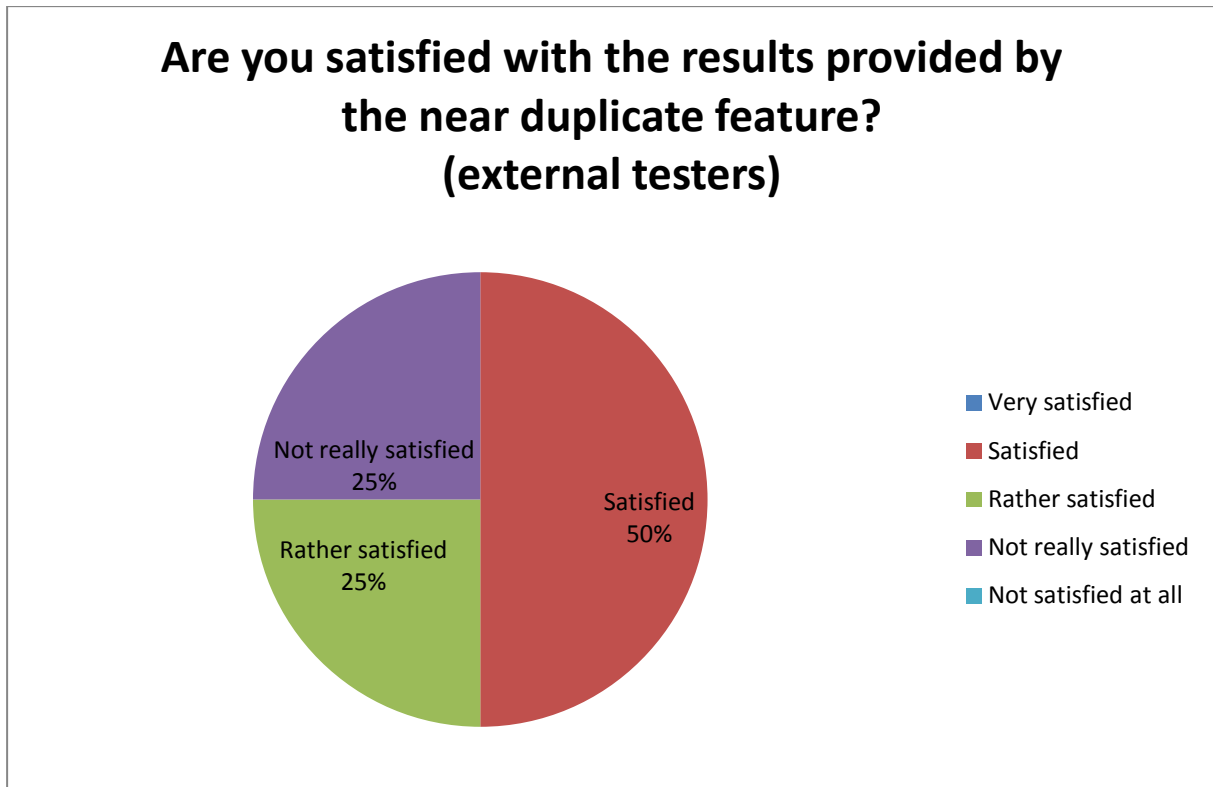


Figure 41: InVID Verification Application – external testers: Results for the question “Are you satisfied with the results provided by the near duplicate feature?”

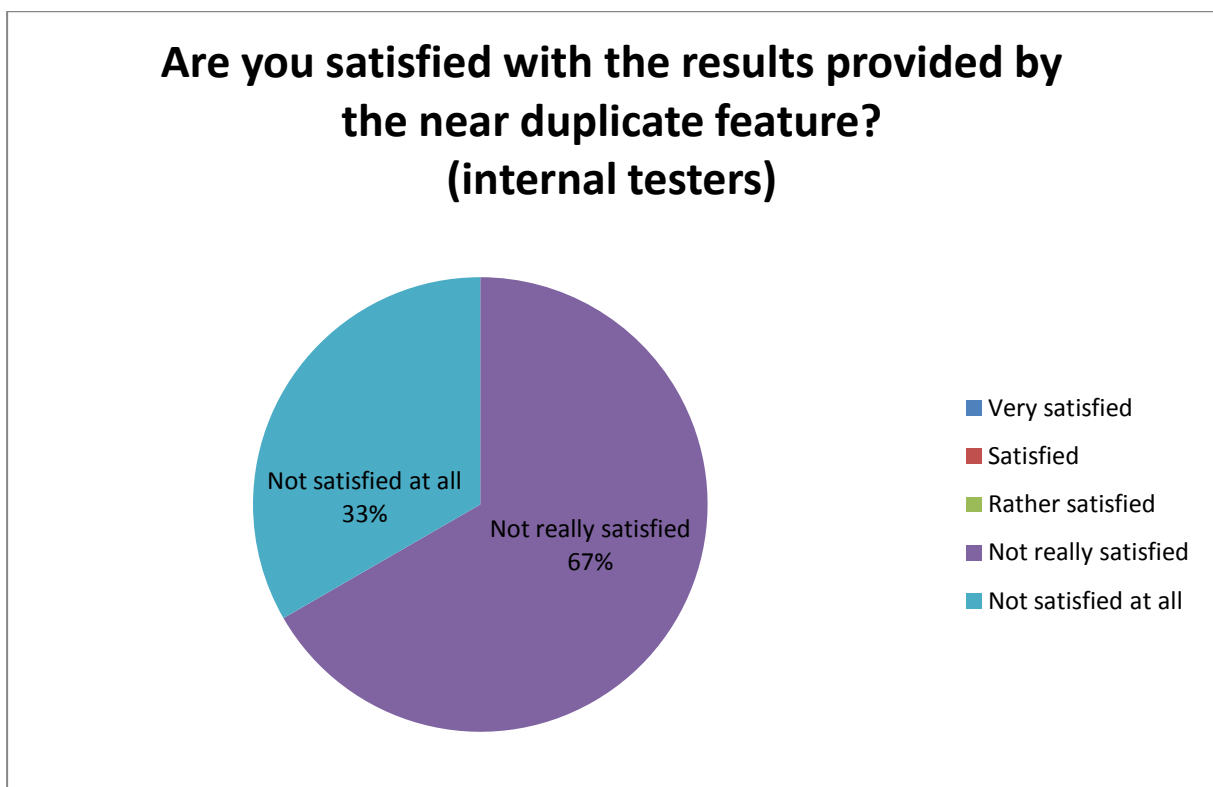


Figure 42: InVID Verification Application – internal testers: Results for the question “Are you satisfied with the results provided by the near duplicate feature?”

3.12 InVID Mobile Application

3.12.1 Description of the service

The Mobile Application is developed to allow communities of non-journalist users of a journalistic service (e.g. registered users of a news web portal or the web-edition of a newspaper) to contribute User Generated Videos directly to the News Organization that provides this service. Uploaded videos are forwarded to the editorial UGC management system which provides also integration with the InVID Verification Application.

3.12.2 Tests

The InVID Mobile Application was tested in test cycles 4 to 6. In test cycle 4 first tests with the prototype of the application were done by testers of the consortium. In test cycle 5 the Mobile Application was tested by journalists from regional newspapers (Tiroler Tageszeitung and Vorarlberger Nachrichten). In test cycle 6 the functionality of the Mobile Application was evaluated by testers of the consortium using a range of different smartphones.

Table 17: Number of received feedback comments for the InVID Mobile Application

Test cycle	Feedback comments
Test cycle 4	14
Test cycle 5	2 survey responses from 3 journalists 7 feedback items
Test cycle 6	6

3.12.3 Major outcomes of test cycles four to six

Tests by users of the consortium

The tests of the prototype in test cycle 4 found, as expected, a wide range of different bugs. These bugs were fixed before the beginning of test cycle 5. The performance tests on different devices during test cycle 6, showed certification problems on Android 6.0. Also suggestions for the improvement of the registration process were made by the participants to this test cycle.

Tests by external testers

Overall impression:

The overall impression about the InVID Mobile Application ranged from "Rather useful" to "Very useful".

Feedback on the ease of use:

The usability of the application was rated as "Easy to use".

Feedback on the integration with the UGC Management System:

The integration of the InVID Mobile Application with the APA-OnlineManager was evaluated as "Very useful". An additional integration with the APA-IT Video service was requested.

Additional feedback revealed a crash of the application when a user looks at the list of the already stored video files and tries to store a time-lapse video in the device.

3.13 InVID Core Platform API

3.13.1 Description of the service

The InVID Core Platform is the central metadata repository that acts as a data exchange platform for all InVID applications and components. Also it provides analysis services, such as the ones supported by the Video Fragmentation and Annotation Service and the Recognize module, via an API.

3.13.2 Tests

A dedicated test of this technology was performed in test cycle 4. The testing of the Core Platform API was done by IT technicians from AFP and APA-IT. The conducted tests focused on the extend to which recommendations from previous test cycles were addressed, and the new functionality of Recognize.

Table 18: Number of received feedback comments for the InVID Core Platform API

Test cycle	Feedback comments
Test cycle 4	12
Test cycle 5	--
Test cycle 6	--

3.13.3 Major outcomes of the test cycles

The major feedback for this API included reports about temporary proxy problems of the API, missing functions in the Swagger documentation of this technology, and suggestions for further improvements of the error handling and error reporting mechanism of the platform.

4 Conclusions and outlook for the next test cycles

The results of test cycles 4 to 6 have shown that the development of the overall InVID platform is in a very good progress, and that the individual InVID applications and analysis components are very useful to journalists. The test cycles have provided the developers with a lot of valuable feedback on various aspects of the evaluated technologies, both from internal and external testers.

Already in test cycles 1 to 3, the tests were done from a journalistic point of view. But the main focus in test cycles 1 to 3 was on the functionality of the different applications and components. In test cycles 4 to 6, this has shifted to an assessment of how the tools help a journalist in the task of verification and rights management. Of course feedback on a detailed level has also been provided in the test cycles for functional and non-functional (e.g. performance, usability) features of the evaluated technologies. Extended tests of the exposed APIs have ensured a stable connection between the different components and services of the InVID platform.

Testing with external users has been very important in the test cycles covered by this document. This allowed to gather feedback from people that are not connected to the project, and therefore have an independent and more subjective opinion on the usefulness, efficiency and stability of the applications. Both external and internal testers with a journalistic background, evaluated the exposed technologies as being very helpful for the video verification tasks of journalists.

The results of the test cycles have also shown proof of the progress of the applications with the increasing development and test cycles. Major feedback of the test cycles has been considered in the development of the applications and components and has led to a better assessment by the testers in the following test cycles.

In the upcoming test cycles it is planned to expand testing with external users. This will ensure that the development of the applications will perfectly meet the needs of the market for video verification. The testing of the components and applications by internal users with a journalistic background will also be continued to gather additional suggestions for improvements of these technologies. The main goal of the next test cycles 7 to 9 is to provide another round of detailed and focused feedback for the performance of the developed applications, that will allow us to apply the needed adjustments and improvements and make these technologies fully compatible to the market needs for efficient video verification and rights clearance.