

Multimodal news authentication as a service: The «True News» Extension

Anastasia Katsaounidou¹

Nikolaos Vryzas²

Rigas Kotsakis³

Charalampos Dimoulas⁴

Abstract

The current work focuses on the problem of misinformation. Filtering and blocking every unreliable source is impossible even to consider. Thus, discovering the dedicated steps to indicate fake content, according to the fact-checking procedures, and utilize them in automated/semi-automated mechanisms, is the key to defend the truth. Despite the availability of various authentication applications/services, there is a lack of integrated systems supporting media veracity in real-world scenarios. There are useful tools and practices for detecting processed/altered content, usually investigating a single manipulation or relying on the credibility of one source, which turns to be inadequate. Based on the above, a browser extension is presented, aiming at evaluating news authenticity in a multimodal, integrating and collaborative way. Operating unobtrusively in the background until needed, the extension is a solution transparent to the user. Without having to open a new browsing tab or to switch navigation /media environment, the user can identify relevant information regarding the five (5) clues, which frame a news story (Title, Date, Creator, Source, Containing Images). In essence, with the aid of the True News Plugin, a user reveals information from trusted sources, classifying them by the accuracy of their domain names, while also spotting possible misinformation through doctored images. The main novelty of the proposed module is that it will help users determine whether they can trust an article or not. By going through five (5) essential steps and by answering ten (10) questions, the proposed methodology attempts to introduce a valuable free tool in the field of Digital Forensics.

Keywords: Misinformation, Fake-news, Multimodal authentication, Plugin, Browser extension

¹PhD Researcher, Aristotle University of Thessaloniki, Greece, email: akatsaoun@jour.auth.gr

²PhD Researcher, Aristotle University of Thessaloniki, Greece, email: nvryzas@jour.auth.gr

³Tenured Senior Teaching Fellow and Instructor, Aristotle University of Thessaloniki, Greece, email: rkotsakis@gmail.com

⁴Associate Professor, Aristotle University of Thessaloniki, Greece, email: babis@jour.auth.gr

1. Introduction: Information can fit into the “palm of our hand” allowing everyone to see and hear everything in real time. Social Media are mining peoples’ reality by holding a critical role in the way they perceive information. Digital processes have become visible by shaping peoples’ experiences and algorithms from scientific papers suddenly emerge as objects of newspaper articles and conversations during coffee (Dourish, 2016; Katsaounidou, Dimoulas & Veglis, 2018). Observing what is happening on the Internet, we will face with a situation where Google holds 87% of worldwide online searches and Facebook has surpassed 2.2 billion monthly users. The two companies absorb more than 60% of the global cost of digital advertising and of course, they are the greatest owners of information (Statista, 2018).

Facebook's business model heavily relies on ads, as the majority of the social network's revenue comes from advertising (Budak, Agrawal, & Abbadi, 2011). The Internet that once looked like a rich range of blogs and web sites has been squeezed out by a few platforms that "manage" what ideas and opinions will be seen and shared, but also by ever stronger digital gatekeepers, whose information distribution technologies can easily be used by manipulators whose motives are in doubt (Marwick & Lewis, 2017).

Widespread misinformation on the Internet is a cause of concern, making everyone suspicious or prone to disagree with everything (Owen, 2019). Entirely made-up or manipulated content that looks like real journalistic report is disseminated, and designed headlines go viral. More often than not, due to its influential and emotional symbolism, people are encouraged to share them. Surprisingly, creators of content, trying to increase their profitability, support attempts to reproduce false content (emotional, political, etc.), which most of the time cause much more reactions than ordinary news (Bakir & McStay, 2018).

The explosion of misinformation spread and internet propaganda is partially a result of how the advertising platforms of the major digital platforms, such as the above are designed to keep people's attention (DiFranzo & Gloria, 2017). Until today, the scientific community has been divided on whether a regulatory framework needs to be established. However, the debate should revolve around what kind of regulatory framework is appropriate to avoid the complete disappearance of freedom of speech on the Web (Lazer, Baum, Benkler, Berinsky, Greenhill, Menczer, ...& Schudson, 2018). To answer the above question, we must first think about how much profit Social Network Services (SNSs) are willing to lose and how much freedom users are willing to sacrifice. From the above, it is clear how important the next day's solutions are, including automated and semi-automated techniques to detect and identify inaccurate information (Katsaounidou, Dimoulas & Veglis, 2018).

Misinformation forms an exceptionally complicated research field, where multiple scientific and applied disciplines are involved. It is a fertile ground for scientific articles, innovative technological proposals, startups, and newspaper articles. Although there are plenty of approaches trying to solve the problem, the real societal impact is difficult to be identified. This is partly due to the fact that the problem is discussed behind the closed doors of Verification Industry, a multidisciplinary fact-checking and Digital Forensics (DF) industry which includes debunking sites, researchers among the fields, projects, institutes and consortiums trying to analyze and propose solutions (Katsaounidou, Dimoulas & Veglis, 2018). It goes without saying that in the case of such particularly complex and multidisciplinary questions, more than one “rapporteur” must be appointed. Thus, in fighting false information, the number of initiatives, web environments,

organizations, platforms and tools for information verification has enormously increased (Wardle & Derakhshan, 2017).

By studying related research, it is easy to observe that most attempts to face the problem of misinformation through automated solutions seem sole and/or incidental. More specifically, examples of false news have been identified, discussed analyzed and used, to train classification algorithms that recognize inaccurate information in favor of humans (Reid & Sands, 2016; Katsaounidou & Dimoulas, 2018b). For this reason, most of the already implemented web-services are specialized in one type of multimedia content, i.e. text, image, audio, video, URLs.

Nevertheless, based on human-operated validation of sites and stories, most of the existing environments try to automate every aspect of the encompassed tasks and operations. Following the above approach, people learn to depend on software to come to conclusions and make decisions, a habit that has never proved to be adequate and beneficial for them in the long-run (Dourish, 2016). Therefore, the significance of semi-automated solutions, suitable for facilitating media authentication, while also supporting digital media literacy and life-long education emerges.

The goal of the verification field is that, eventually, users will acquire the knowledge and know-how to identify the realness of information with partial machine assistance, to be capable of transmitting accurate information instead of propagating misinformation (Wardle & Derakhshan, 2017). In this direction, the ultimate objective of the present project is to investigate the possibility of creating a reliable application, namely a browser extension, which will bear the latest forensic verification developments, without disregarding the well-known traditional principles of doubting about every little piece of information and answering to the five Ws questions (who, what, where, when, why) plus “how” (Katsaounidou & Dimoulas, 2018a). Hence, the interface will guide the users step by step through all the deployed validation processes, helping them to become familiar with the tasks they should follow in evaluating the truthfulness of an article, therefore cultivating their necessary verification skills.

Social Media (SM) and web pages have been proved unable to address the phenomenon of misinformation spread within their platforms, in the first place (Metzger & Flanagin, 2013). A browser extension, a software component that adds specific features to an existing application, has been chosen as the most appropriate solution for the task. SNSs like Facebook, Twitter, and YouTube are used by members to spread “knowledge” and to seed topics for journalists (Marwick & Lewis, 2017). Moreover, SNSs are the favorite places and/or tools of news organizations to broaden their audience. Unfortunately, due to their widespread nature and the lack of efficacious censorship, without suppressing rightful freedom of speech, SNSs are the central space for spreading misinformation online (Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012). Essentially, trying to overcome the above lack of control, the True News extension offers a solution transparent to the user, operating unobtrusively in the background until needed. Without having to open a new tab in the browser or to move away from each environment, the user can collect useful information regarding the nature of the news items.

Undoubtedly, it is necessary for the verification field to develop, modernize, update and optimize applications in the “war” against disinformation. A brief review of related works (applications, web pages, extensions, etc.) and their theoretical background is presented and analyzed, to allow the reader to become acquainted with the current state of the art.

Literature review:

As most researchers tend to agree, the most important initiative regarding fact checking is the International Fact Checking Network (IFCN), which aims at bringing together fact-checkers worldwide (International Fact Checking Network, 2018). First Draft⁵ forms another innovative network, initiated in June of 2015, with objectives to raise awareness, perform research, and address challenges relating to trust and truth in media in the digital era. “First Draft” undertakes practical journalism projects in the field, investigating effective methods for tackling information disorder online. It also provides practical and ethical guidance on how to find, verify and publish content, sourced from the social web (First Draft News, 2019). Integrated environments like the web-based collaboration platform Truly Media⁶ and its Artificial Intelligence (AI) utilities, powered by Truth Nest⁷, help individuals discover, analyze and verify any information through predefined operations. As already mentioned, most of the existing online services are unimodal, i.e., specializing in a single content entity, with the visual element being the most popular subject of investigation. Due to their common use as evidence in News, images have also become the most usual form of digital misinformation, i.e., tampered and re-used photos in the context of a new story (Katsaounidou, Dimoulas & Veglis, 2018). The simplest way to check the history of a picture on the Web is by reverse image search, a content-based search driven by the visual data and not by text. The most popular reverse image search engines are Google Image, TinEye, Bing, Yandex and Baidu. Significantly, Google has already tons of stored pictures, and anyone can upload an image file or paste its corresponding URL into the search bar for retrieving similar content (Reid & Sands, 2016). Existing open source tools offer verification solutions that can adequately support typical authentication tasks, although they do not feature the entire desired functionality and comprehensiveness for the average user. Two popular and free web services are the Image Verification Assistant⁸ (Zampoglou, Papadopoulos & Kompatsiaris, 2017) and Forensically⁹, which contain online toolkits for fundamental forensic investigation. Similarly, Ghiro¹⁰ is an open-source web application for inspecting image metadata.

Photo Detective¹¹, recently renamed to Axon Detect, is another profitable forensics toolset. The platform offers useful authentication insights, by taking into account lighting direction calculation, compression consistency estimation and metadata analysis. Likewise, system Pizarro¹² performs standard forensic procedures, along with algorithmic image reconstruction capabilities (Kamenicky, Bartos, Flusser, Mahdian, Kotera, Novozamsky, & Zitova, 2016). According to Korus (2017), Amped Authenticate¹³ is the most comprehensive commercial platform, assembling tools for the manual investigation of visual data, as well as automatic indicators of several forensic trails. Finally, the JPEG Snoop¹⁴ application allows the retrieval of the full JPEG compression settings, taking advantage of an extended database of identified capturing and storing signatures for many cameras.

⁵First Draft News <https://firstdraftnews.org/>

⁶Truly Media <http://www.truly.media/>

⁷TruthNest <https://www.truthnest.com/>

⁸Image Verification Assistant <http://reveal-mklab.it.gr/reveal/>

⁹Forensically <https://29a.ch/photo-forensics/#forensic-magnifier>

¹⁰Ghiro <http://www.getghiro.org/>

¹¹PhotoDetective <http://metainventions.com/photodetective.html>

¹²Pizarro <http://pizarro.utia.cas.cz/>

¹³Amped Authenticate <https://ampedsoftware.com/authenticate>

¹⁴JPEG Snoop, <http://www.impulseadventure.com/photo/jpeg-snoop.html>

Video content is more difficult to manipulate and also harder to verify, thus, the available video authentication platforms are outnumbered by the image-oriented ones (Papadopoulou, Zampoglou, Papadopoulos & Kompatsiaris, 2019; Teyssou, Leung, Apostolidis, Papadopoulos, Zampoglou, ... & Mezaris, 2017, October). The verification process becomes heavier as the visual information increases, i.e. checking every frame of a video file is hugely time-consuming. One solution is to reveal the video key-frames and the corresponding thumbnail pictures, utilizing them in reverse image search. This can flag up other videos that contain similar footage. Complementary keywords-based search can expedite the process.

In all cases, one can reveal whether a seemingly new image or video is reused, as long as he/she is equipped with patience and the correct tools. Amber Video¹⁵ is a related platform that uses signal processing and artificial intelligence to identify tampered audio and video, designed to detect /stop misinformation, therefore to eliminate distrust. Amber Video is also useful for individuals who need to investigate the accuracy of videos, the source of which is unknown. In VID¹⁶ is another solution that aims at detecting, checking and verifying newsworthy video material, spread through social media, thus exporting credibility marks (Papadopoulou et al., 2019; Teyssou et al., 2017).

A company that specializes in real-time video verification to protect customers, business and profits, offers information regarding its services through the site Iverify¹⁷. Moreover, Amnesty International¹⁸ has introduced a new web service to support journalists in checking YouTube videos. Additionally, to help address these kinds of issues, Amnesty International has also launched a website, the Citizen Evidence Lab¹⁹, providing journalists and human-rights advocates with tools and learning material on validating user-generated video. Furthermore, Storyful²⁰, in collaboration with Google, has created the Montage²¹, a product that allows users to team up on verifying or analyzing YouTube videos.

Tweet Verification Assistant²² is the only dynamic text centered application that evaluates the integrity of a tweet, by analyzing multiple (textual mainly) parameters, i.e. language, punctuation, number of hashtags, mentions and external links, as well as multimedia content (attached or connected) (Boididou, Papadopoulos, Zampoglou, Apostolidis, Papadopoulou, & Kompatsiaris, 2018).

Regarding Forensic Audio analysis, no open source platforms can be found. Among the commercial solutions, the IKAR Lab²³: Forensic Audio Suite and the Forensic Audio Workstation are listed, forming professional software suites for speech signal analysis, both provided by the Speech Technology Center. Moreover, the authors of the current paper have recently presented a framework (in its infancy) for delivering supervisory tools for audio-driven multimedia Content Authentication as a Service (CAAAS) (Vryzas, Katsaounidou, Kotsakis, Dimoulas, Kalliris, 2019). We have presented the existing web-based services and their implications in the related field

¹⁵ Amber Video <https://ambervideo.co/>

¹⁶ InVID <http://invid.condat.de/>

¹⁷ Iverify <http://www.iverifysecurity.com/solutions/video-verification.html>

¹⁸ Amnesty International <https://www.amnesty.org/en/>

¹⁹ Citizen Evidence Lab <https://citizenevidence.org/>

²⁰ Storyful <https://storyful.com/>

²¹ Montage <https://montage.meedan.com/welcome>

²² Tweet Verification Assistant <http://reveal-mklab.iti.gr/reveal/fake/>

²³ IKAR Lab https://speechpro.com/product/forensic_analysis/ikarlab

of interfaces, intending to validate multimodal content. The existing browser extensions are discussed below, providing some analysis insights that will facilitate the presentation and analysis of the proposed “True News” extension.

Browser extensions:

The most vanguard approach is the B.S. Detector²⁴ browser extension with 19.074 users worldwide, which is powered by OpenSources²⁵, a professionally curated list of unreliable or otherwise questionable sources (Zimdars, 2016). By domain classifications, B.S. Detector categorizes the sites as fake news, satire, extreme bias, conspiracy theory, rumor mill, state news, junk science, hate group, click bait and proceeds with caution, displaying a warning screen when someone enters a site known to publish false news stories. The main problem is that the repository of Open Sources has a tiny number of entries (834), compared to the billions of websites online. The optimal use of the B.S. Detector could be succeeded by checking and labeling all the webpages worldwide, a process which is considered unfeasible. Moreover, classifying a specific article as fake does not mean that all the items provided by the same site are unreliable (Fan, 2017).

Another interesting approach is the Fake News Guard²⁶ extension with 289 users, which checks every visited page and every link that ends up on the user’s Facebook feed against its blacklist. If the source of the article is blacklisted, the user gets a particularly detectable warning in the browser. Moreover, this approach introduces the element of interaction, allowing users to report on anything they consider unreliable, thus helping the creators of Fake News Guard to get valuable feedback to set further improvements. One drawback of this tool is that, although it is an extension, its full interface is unfolded in a separate browser tab.

In the same line, there are two similar name extensions, the Fake News Detector²⁷ with 1.398 users and the Fake News Detector²⁸ with 651 users. The first one marks fake news in the browsing pages in red color and the clickbait links in orange color. The second one allows users to detect and label news directly from their Facebook and Twitter accounts, using the following flags: Legitimate, Fake News, Click Bait, Extremely Biased, Satire or Not news. After flagging an item, this becomes visible to the rest of the users, so that they can be more cautious. Human-provided information is used to update a database, in order to train algorithms that can automatically classify news into the above-listed categories in the long-term. Through this method, even recently published stories that no one has seen, may be rapidly flagged in an automated way. The critical innovation of this specific plugin is its collaborative nature, since it displays on Facebook users, the decision of peoples and algorithms. The problem, in this case, is that anyone can flag information, regardless of his /her abilities in the information validation process and the skills on deciding what news is right and what wrong.

Regarding image authentication, a Google Chrome plugin, called RevEye²⁹ (23.608 users), checks databases at Google, TinEye, Bing, Yandex, and Baidu. The search engines are useful in the cases

²⁴ B.S Detector <https://github.com/selfagency/bs-detector>

²⁵ Open Sources Repository <http://www.opensources.co/>

²⁶ Fake News Guard <https://chrome.google.com/webstore/detail/fake-news-guard/pmebnfmgcgnpmecdcopidnjdlnggbech>

²⁷ Fake News Detector <https://chrome.google.com/webstore/detail/fake-news-detector/aebaikmeedenaijgjcfnmdfknobahep>

²⁸ Fake News Detector <https://chrome.google.com/webstore/detail/fake-news-detector/alomdfnfpbaagehmdokilpbjchacabk>

²⁹ RevEye <https://chrome.google.com/webstore/detail/reveye-reverse-image-sear/keaaccljhehbbapnphnmpiklalfhelgf?hl=en>

of re-used visual content, allowing users to find out the origin and the context /surrounding behind a picture and its use, or even to retrieve higher resolution versions.

Another popular approach for video verification is the InVID³⁰ extensions with 10.789 users (Papadopoulou et al., 2019; Nixon, Apostolidis, Markatopoulou, Patras & Mezaris, 2019) which aims at detecting, checking and verifying newsworthy video material, spread through social media to export credibility marks. In this case, also, a major weakness is that, although it is a plugin, it enables the full interface in a separate browser tab.

Prioritizing metadata as the most critical factor, the Send to Exif Viewer³¹ extension by Jose Tomas Tocino with 7.073 users adds an element to the contextual menu that opens images in a metadata viewer. Furthermore, some tools utilize the lists in the debunking databases of the fact-checking websites. This is the case of Instant Snopes Checker³² (Unofficial) with 596 users, which offers a quick and simple way to check if the current page title and keywords are listed on Snopes.com. This approach relies on the idea that click bait news posts are frequently recycled, hence the recorded titles and keywords of the debunked articles on Snopes.com may provide useful indications.

Closer to the True News approach, in terms of functions, is First Draft News Check³³ with 1.380 users, an interactive version of First Draft Visual Verification Guide for Images and Videos. The extension allows users to have an image or video open in a browser and then work through a checklist that prompts to investigate: 1) If they are looking at an original piece of digital material; 2) How confident they are about who created the content; 3) How confident they are about the date of the capture (i.e. when); and 4) how confident they are about the place of content creation (i.e. where). The extension calculates a verification score and creates a button that can be embedded on each website.

As already noticed, all the mentioned above tools are intended for users familiar with the English language but also for content analysis written in English. Nevertheless, there is a Greek Hoaxes Detector³⁴ extension with 2.061 users offered by the Greek debunking site “Ellinika Hoaxes³⁵”. Most of the code comes from the aforementioned BS Detector, under the LGPL-3.0 open source license. In essence, this extension checks a list of unreliable or questionable Greek sources and informs the user when he/she visits the blacklisted sources.

In summary, apart from some exceptions, most of the above-listed tools are fully automated, keep their internal procedures hidden, are usually unmoral (specializing in one content type), and do not offer collaboration. Hence, a browser extension aiming at evaluating the authenticity of news (and generally posts) in a multimodal, integrated and collaborative way, is a necessity (Katsaounidou & Dimoulas, 2018a). This is precisely the targeted innovation of the True News approach, guiding the user through concrete steps and questions that need to be answered, taking into consideration the traditional cross-validation rules.

Another limitation of the above-presented tools is that they do not produce a final decision, but they only give users some approximate indices. The True News approach seeks to formulate a decision-making process and extracting a conclusive answer as to whether the article is true or false. The

³⁰ InVID <https://chrome.google.com/webstore/detail/fake-video-news-debunker/mhccpoafgdgbhnjfhkcmgknnndkeenfhe?hl=en>

³¹ Send to Exif Viewer <https://chrome.google.com/webstore/detail/send-to-exif-viewer/gogiienhpmfmodmlnhdljokkjapfck?hl=en>

³² Instant Snopes Checker <https://goo.gl/U1vzih>

³³ First Draft News Check <https://chrome.google.com/webstore/detail/firstdraftnewscheck/japockpeaaanknlkhagilkgcledilbfk?hl=en>

³⁴ Greek Hoaxes <https://goo.gl/ryBNGA>

³⁵ Ellinika Hoaxes <https://www.ellinikahoaxes.gr/>

above is achieved through the semi-automated transparent processes that it contains. This process targets also to support users by enhancing digital literacy. The user, guided by the True News interface, answers with crisp Yes or No choices to the questions that each step includes. In the end, all these answers are fused into a final decision that is displayed to users, providing an overall truthfulness estimation. More detailed information about the True News system architecture is presented in the following sections.

2. True News Extension Presentation: As in most cases of project implementation, the final software service emerges after iterative design, debugging and improvements. The objectives of this subsection include the detailed presentation of the "True News" extension, the description of the fact-checking steps concerning news verification, and specifically the questions that the user has to answer to get the final result. Accordingly, User Experience (UX) Design aspects are analyzed and clarified to optimize usability with emphasis on the tools that are integrated in the plugin.

The application consists of five (5) different sub-modules that are successively deployed along the end-to-end chain. Though the original plan contained five (5) tabs, incorporating the required validation clues (Article Title, Date, Creator, Source, Containing Images), it turned up to an eventual screen setup of thirteen (13) tabs. The first panel corresponds to the welcome page. The next ten (10) displays contain the questions addressed to the users and the remaining two (2) form the areas to display the final pages. An analysis of similar applications was conducted, leading to the adoption of similar /consisted Graphical User Interface (GUI) design, following the modern UX trends and aesthetics. Representative screenshots are provided to help the illustration of all the intermediated processes and authored functionality.

Following the installation process, the welcome page appears on the right side of the browser, providing quick guidelines about the plugin usage (Figure 1). Careful dimensioning and positioning have been selected to place the add-on User Interface (UI) in the sidebar area of the Facebook platform, usually reserved for advertisements or other communication elements, avoiding to hide parts of the article or other useful information. Actually, the True News window has been set to a size a bit larger than the usual format of these hosted services, aiming at offering higher resolution and overall application visibility.

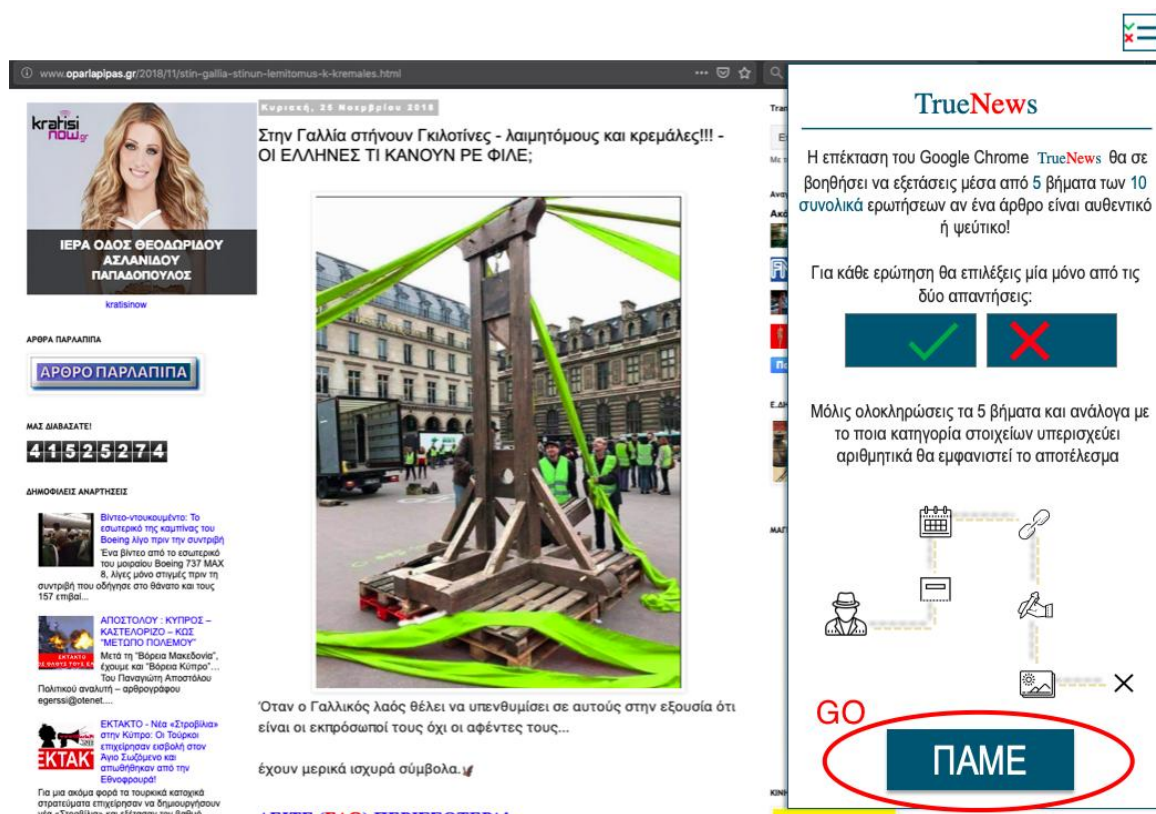


Figure 1. The Welcome Page of True News extension

When the plugin is activated through the selection of the GO (ΠΑΜΕ) button at the bottom of the window, the first question unfolds: “Does the title contains overloaded language, excessive punctuation (!!!), a lot of capital letters to emphasize?” (Figure 2 left), introducing users to the step-by-step validation process.

Following the user’s answer (Yes | No) the cross-checking proceeds to the second inquiry: “Does the title claims that it contains “a secret” or something that is hidden from the mainstream media /informing streams?”(Figure 2 center). In both of the above two cases, the title of the article is automatically repeated at the bottom of the UI, near the user's interaction buttons, to serve usability, expediting the analysis in a more straightforward manner. In this context, users are learning and becoming familiar with the evaluation of title-related features, therefore to detect potential propaganda articles.

Based on the provided feedback and instructions, the guide continues to the third question, aiming at checking previous records on that (or similar) on a predefined list of unreliable sources: “Are there any unreliable sites that have also published this title?” (Figure2 right). In this step, a pop-up window with the corresponded search engine results emerges, and the user is asked to simply track down the number of unreliable webpages, answering (Yes | No) concerning if there a doubtful background on that title.

The figure shows three panels of the TrueNews interface, each with a 'close' button in the top right corner.

Panel 1 (Left): Question 1 asks if the title contains overloaded language, excessive punctuation (!!!), a lot of capital letters to emphasize. The title is: "ΣΤΟΙΧΕΙΟ: ΤΙΤΛΟΣ 1. Ο τίτλος περιέχει "υπερβολική γλώσσα", πολλά σημεία στίξης ή πολλά κεφαλαία γράμματα;". Below the question, there is a list of unreliable sites: "AN στον τίτλο του άρθρου εντοπίζουμε: Ορθογραφικά λάθη, Πολλά θαυμαστικά (!!! , ; ;) ΚΕΦΑΛΑΙΑ ΓΡΑΜΜΑΤΑ, Υπογραμμισμένες λέξεις, Τονισμένη γραμματοσειρά, Διαφορετικούς χρωματισμούς, Κίτρινο φόντο". At the bottom, there are two buttons: a green checkmark and a red X.

Panel 2 (Center): Question 2 asks if the title claims that it contains "a secret" or something that is hidden from the mainstream media /informing streams. The title is: "ΣΤΟΙΧΕΙΟ: ΤΙΤΛΟΣ 2. Ο τίτλος περιέχει κάποιο "μυστικό" ή υποστηρίζει πως λέει κάτι που τα παραδοσιακά μέσα ενημέρωσης "δεν θέλουν" να μάθεις;". Below the question, there is a list of unreliable sites: "AN ο τίτλος αναφέρει ξεκάθαρα ότι μας ενημερώνει για κάποιο μυστικό και /ή, AN στον τίτλο ανιχνεύσουμε λέξεις με υπερβολικό περιεχόμενο όπως: ΔΕΙΤΕ, ΕΔΩ, ΦΑΣΙΣΤΕΣ, ΑΡΙΣΤΕΡΟΙ, ΔΕΞΙΟΙ, ΚΛΕΦΤΕΣ, ΚΑΤΕΒΑΣΟΥΝ, ΡΕ, ΠΡΟΔΟΤΕΣ, ΑΛΗΤΕΣ και άλλες λέξεις/ φράσεις με συναισθηματικό φορτίο ή απευθείας ερωτήσεις". At the bottom, there are two buttons: a green checkmark and a red X.

Panel 3 (Right): Question 3 asks if there are any unreliable sites that have also published this title. The title is: "ΣΤΟΙΧΕΙΟ: ΤΙΤΛΟΣ 3. Υπάρχουν άλλες ΜΗ αξιόπιστες ιστοσελίδες οι οποίες έχουν δημοσιεύσει τον ίδιο τίτλο;". Below the question, there is a list of unreliable sites: "AN στα πρώτα 10 αποτελέσματα που θα μας επιστρέψουν οι μηχανές αναζήτησης δούμε ότι το άρθρο έχει δημοσιευτεί και σε άλλες μη αξιόπιστες ιστοσελίδες όπως τις παρακάτω. Υποψιαζόμαστε ότι είναι προϊόν παραπληροφόρησης. Επέλεξε σε ποιες από τις παρακάτω ιστοσελίδες το εντόπισες". Below the list, there is a button: "Δες τα αποτελέσματα σε νέο παράθυρο". At the bottom, there are two buttons: a green checkmark and a red X.

Figure 2. Question 1: Does the title contain overloaded language, excessive punctuation (!!!), a lot of capital letters to emphasize (left). Question 2: Does the title claims that it contains “a secret” or something that is hidden from the mainstream media /informing streams? (center). Question 3: Are there any unreliable sites that have also published this title? (right)

The fourth question is similar to the previous one but in a reverse perspective: “Has the “title” of the article been already investigated by the debunking sites?” (Figure 3 left). As Figure 3 (left) depicts, the answer to this question is presented in an inline frame, offering also the option to visit the debunking sites and read more information regarding the specific article. Hence, if the title has already been evaluated by the debunking site “Ellinika Hoaxes”, it would probably be unreliable (implying the YES answer and vice versa).

True News extension also provides substantial evidence concerning the date and the source of the news item, two essential clues that everyone has to examine before believing an article. Hence, at the tab of the fifth question (Figure 3 center) the user can find information about previously published history (if any).

Likewise, at the display of the sixth question, users can identify, in real time, if the visiting page has been already blacklisted by debunking sites (Figure 3 right). In this step, a pop-up window with the corresponded search engine results is shown, and the user is asked to choose the specific page among a list of unreliable webpages.

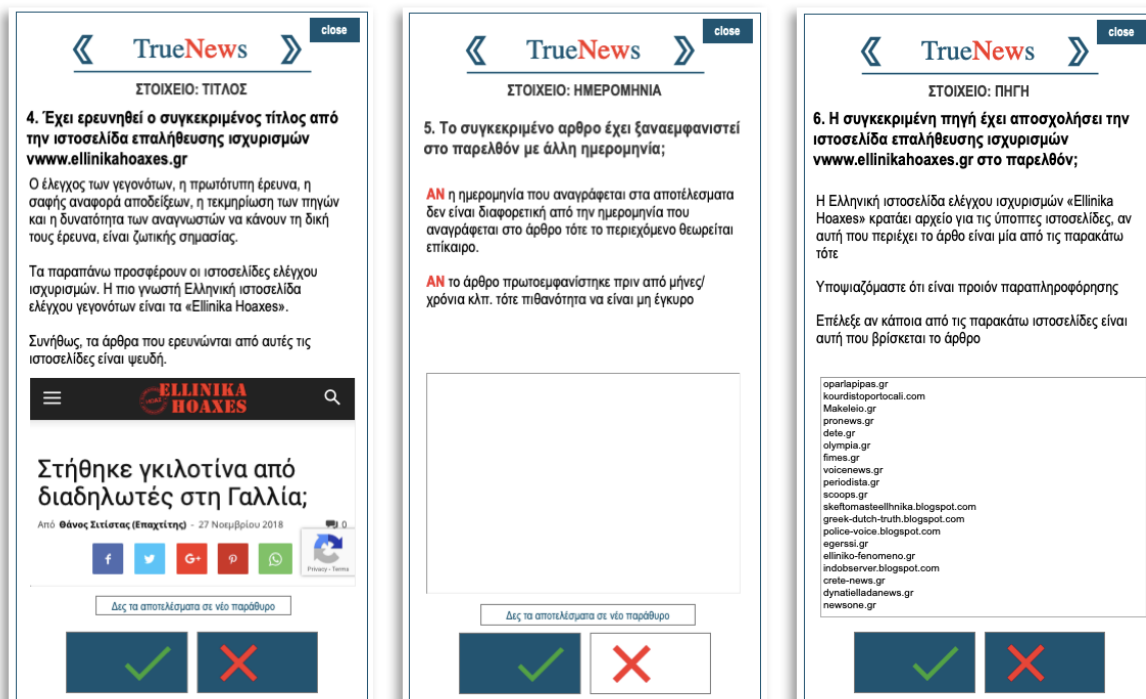


Figure 3. Question 4: Has the “title” of the article been already investigated by the debunking sites? (left). Question 5: Has this article been republished again in the past? (center). Question 6: Has this source been blacklisted from debunking sites? (right)

The seventh question offers guidelines regarding the identification of the author of the article (Figure 4 left). As mentioned before, articles containing images receive 94% more readability than plain text (Katsaounidou, Dimoulas & Veglis, 2018), a fact that makes them the most usual form of digital misinformation. However, the technique of Reverse Image Search (Google, TinEye, Bing, Yandex, Baidu) can help us discover the truth. Thus, the next three questions aim at identifying the nature of the containing images (if any).

In specific, the eighth question "Can you locate the ‘image’ elsewhere on the Internet (different version/context)?" inquires the above issue, returning publication dates, sources, and context of articles using similar visual documents /elements (Figure 4 center). If the results are consistent concerning the frame of the story, the article is considered true, otherwise, it is assumed probably a recycled item.

The next step takes advantage of Jeffrey's Image Metadata Viewer³⁶ and/or the Exif Info³⁷ tools to reveal the basic meta-information of the image (date, time, creation location, logging device, etc.), so that users would be able to answer the question "Is the metadata information of the image in relevance with the article information?" (Figure 4 left). For instance, if the dates do not match with each other, the article is questioned for its reliability.

The tenth and last question investigates the presence of a potential tampering operation: “Is the ‘image’ manipulated?” (Figure 5 left). Users can benefit from the free-to-use Image Verification

³⁶ Jeffrey's Image Metadata Viewer <http://exif.regex.info/exif.cgi>

³⁷ Exif Info <https://exifinfo.org/>

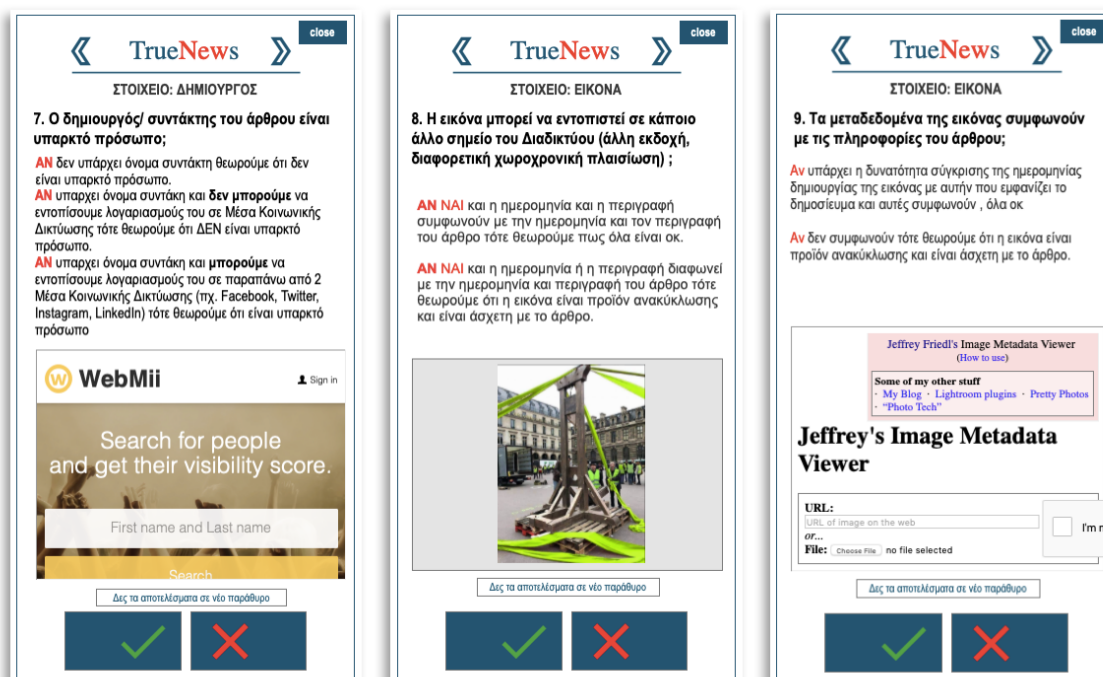


Figure 4. Question 7: Is the author of the article a real person? (left). Question 8: Can you locate the 'image' elsewhere on the Internet (different version/context) (center). Question 9: Is the metadata information of the image in relevance with the article information? (right).

Assistant tool to find out if the visual content is intact. If the tool detects that the image is edited, the article is documented on doctored data, therefore it is very likely to fall to the spectrum of Misinformation. After analyzing and integrating the answers, the system proceeds to a fused decision-making operation, combining all of the unimodal estimates to display the final result (Figure 5 center) and (Figure 5 right).

At this point, it is worth mentioning that the questions and their categorization emerged after thorough analysis. To verify that this corpus of items is a safe and effective way to reach conclusions regarding the nature of an article, many original and fake stories were analyzed from the research team with the use of these ten questions. The results of the above procedure were the promising vehicle on the way of their final determination. Of course, we are consistently trying to identify reasons for false positives to modify the guide and make it more efficient. The above will be achieved in the future by keeping records of the news being checked from the users and by adding/offering the ability to receive feedback.

As already mentioned, the main goal of the plugin is to help people become familiar with the tasks they should follow in evaluating the truthfulness of an article and cultivating their necessary verification skills. While the field of information verification is rapidly evolving, yet, besides the fact-checking principles, no unmistakable method of identifying false information exists. Thus, especially at this time when even first-generation fact-checking is no longer enough, the cooperation between humans and machines (algorithmic techniques) should be intensified.

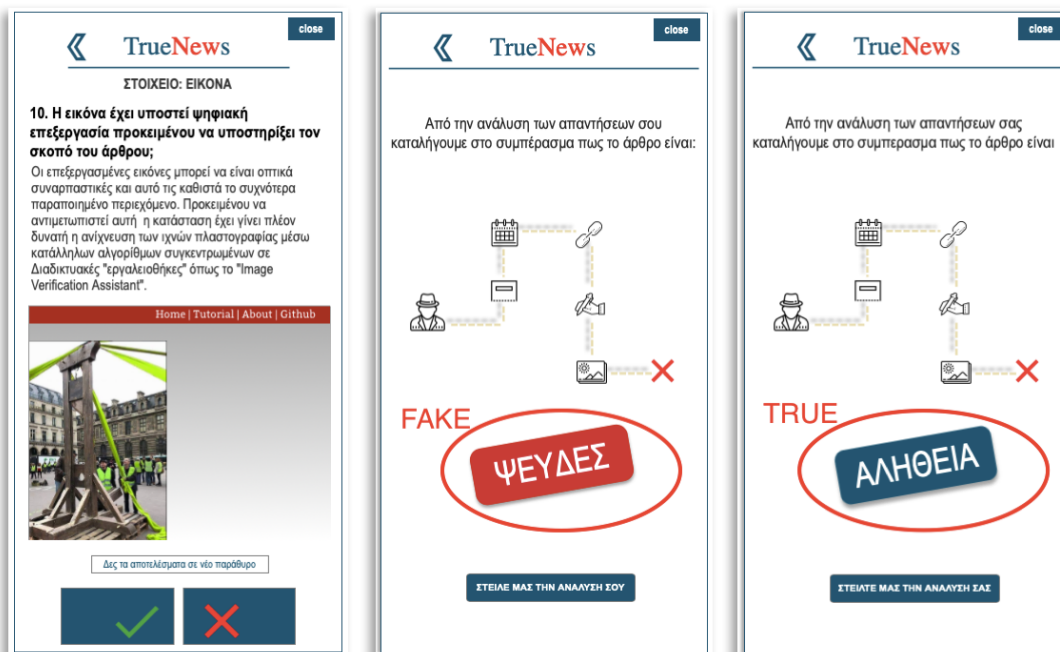


Figure 5. Question 10: Is the ‘image’ manipulated? (left). The result page for fake articles (center). The result page for real articles (right).

3. System Architecture: As mentioned above the True News Service offers a variety /combination of crosschecking and validating operations to estimate the credibility or truthfulness of an article. These processes have to be analyzed into the modules of a multimodal architecture that would serve the whole decision-making implementation. Specifically, Figure 6 presents the proposed architecture for the True News extension in a connected block-diagram flow chart.

Starting from the beginning, when an article is given to the system as input, the first step is to conduct a web parsing operation, to disintegrate the web page into stand-alone elements for further exploitation, namely the modalities of text and images. The textual information is then categorized into the conceptual fields that refer to the title, author, source, and date of the published article elongated by the main plain text area of the article. The above fields are investigated through (meta) search engines, dedicated services and debunking sites for data consistency, debunking sites logging, etc. before exposing the results to the user for the semi-automatic /driven operation of the article validity checking.

While more and more users engage to the service, their feedback is stored and subsequently exploited towards the formulation of generalized rules for more accurate and more automatic checking operations via Natural Language Processing techniques (NLP), like pattern recognition of a Fake Article via the style analysis of its title.

Next, as Figure 6 exhibits, the second modality of the proposed architecture is responsible for the processing of the parsed images. Specifically, the images of the input article are used for reverse search via the aforementioned services to investigate whether they appear in other sources as well. In addition, they are further analyzed by specialized algorithms for the detection of implicated potential (deliberate) manipulation.

The designed scheme is based on a decision-making process at each step according to the questions that are posed to the users, while a final combined/ weighted validation derives from all the results at each control module. At this point, the low fidelity prototype of the whole True News services

has been designed, while on the other hand almost all the modalities have been algorithmically deployed in Matlab 2018b software, except for the part of the dedicated image processing modules. However, the future plan is to implement the algorithms with Python scripting, to combine them with JavaScript web programming for a platform installation that would take the entries/ articles for all of the above functionalities.

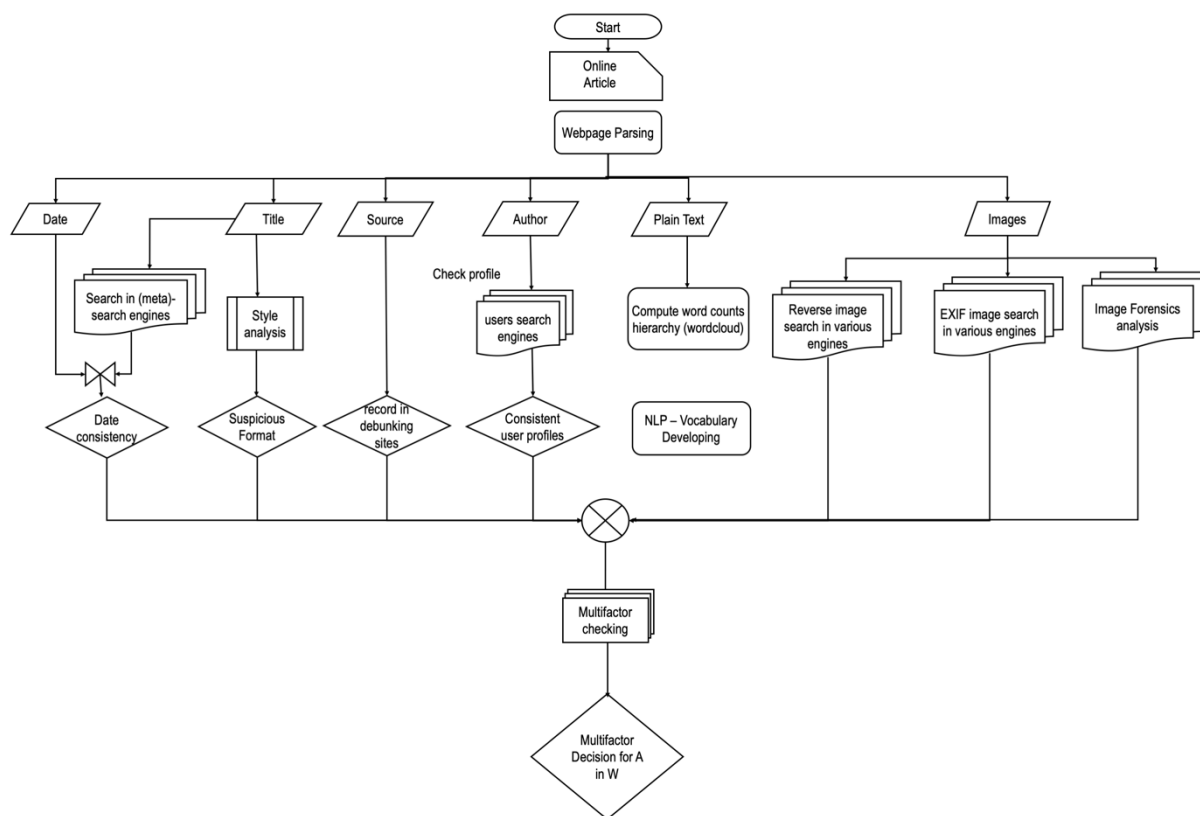


Figure 6. The architecture of True News extension

4. Conclusions: As mentioned previously, the verification industry has tried to offer solutions and to understand the challenges that exist in a variety of ways. Practices, methods, and tools, policies and procedures that adequately contribute to ensuring the quality of the transmitted information have been established trying to cultivate globally informing awareness and digital literacy. And yet, even now, the field of verification industry seems unable to restore the paradigm of factual information. Thus, the paper discusses the significance of semi-automated solutions and presents the True News plugin as the most suitable approach for facilitating media authentication, supporting digital media literacy and life-long education. The current paper, by introducing step by step all the deployed validation processes, hopes to help users become familiar with the steps they should follow in evaluating the truthfulness of an article, therefore cultivating their necessary verification skills. Moreover, it hopes to support the users who lacked critical thinking skills as well as sufficient knowledge of logic, history, etc. to be able to filter this information successfully. Hence, the ultimate goal of True News extension is to gather all the existent knowledge regarding verification in an easily accessible interface offered as an online service. We expect that the proposed framework will eventually ensure users' ability to transmit accurate information and to prevent misinformation propagation.

5. Acknowledgments: This research is co-financed by Greece and the European Union (European Social Fund-ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning» in the context of the project “Strengthening Human Resources Research Potential via Doctorate Research” (MIS-5000432), implemented by the State Scholarships Foundation (IKY) which supports A. Katsaounidou PhD Research. Moreover, this research has been partially supported by General Secretariat for Research and Technology (GSRT), the Hellenic Foundation for Research and Innovation (HFRI) (Scholarship Code: 1900) which supports N. Vryzas' Ph.D. research.

6. References

- Bakir, V., & McStay, A. (2018). Fake news and the economy of emotions: Problems, causes, solutions. *Digital Journalism*, 6(2), 154-175.
- Berghel, H. (2017). Lies, damn lies, and fake news. *Computer*, (2), 80-85.
[25](https://doi.org/10.1109/Boididou, C., Papadopoulos, S., Zampoglou, M., Apostolidis, L., Papadopoulou, O., & Kompatsiaris, Y. (2018). Detection and visualization of misleading content on Twitter. <i>International Journal of Multimedia Information Retrieval</i>, 7(1), 71-86.</p>
<p>Budak, C., Agrawal, D., & El Abbadi, A. (2011, March). Limiting the spread of misinformation in social networks. In <i>Proceedings of the 20th international conference on World wide web</i> (pp. 665-674). ACM</p>
<p>DiFranzo, D., & Gloria, M. J. K. (2017). Filter bubbles and fake news. <i>ACM Crossroads</i>, 23(3), 32-35.</p>
<p>Dourish, P. (2016). Algorithms and their others: Algorithmic culture in context. <i>Big Data & Society</i>, 3(2), 2053951716665128.</p>
<p>Kamenicky, J., Bartos, M., Flusser, J., Mahdian, B., Kotera, J., Novozamsky, A., ...& Zitova, B. (2016). PIZZARO: Forensic analysis and restoration of image and video data. <i>Forensic science international</i>, 264, 153-166.</p>
<p>Katsaounidou A., Dimoulas C. (2018b). The Role of media educator on the age of misinformation Crisis. In EJTA Teachers' Conference on crisis reporting 19-19 October 2018. Thessaloniki, Greece</p>
<p>Katsaounidou, A. N., & Dimoulas, C. A. (2018a). Integrating Content Authentication Support in Media Services. In <i>Advanced Methodologies and Technologies in Digital Marketing and Entrepreneurship</i> (pp. 395-408). IGI Global.</p>
<p>Katsaounidou, A., Dimoulas, C., & Veglis, A. (Eds.). (2018). <i>Cross-Media Authentication and Verification: Emerging Research and Opportunities: Emerging Research and Opportunities</i>. IGI Global.</p>
<p>Korus, P. (2017). Digital image integrity—a survey of protection and verification techniques. <i>Digital Signal Processing</i>, 71, 1-26</p>
<p>Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ...& Schudson, M. (2018). The science of fake news. <i>Science</i>, 359(6380), 1094-1096.</p>
<p>Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. <i>Psychological Science in the Public Interest</i>, 13(3), 106-131.</p>
<p>Marwick, A., & Lewis, R. (2017). <i>Media manipulation and disinformation online</i>. New York: Data & Society Research Institute.</p>
</div>
<div data-bbox=)

- Metzger, M. and A. J. Flanagin (2013) Credibility and trust of information in online environments: The use of cognitive heuristics, *Journal of Pragmatics*, 59 pp. 210-220
- Nixon, L., Apostolidis, E., Markatopoulou, F., Patras, I., & Mezaris, V. (2019, January). Multimodal Video Annotation for Retrieval and Discovery of Newsworthy Video in a News Verification Scenario. In *International Conference on Multimedia Modeling* (pp. 143-155). Springer, Cham
- Owen, L. (2019). 2019: A year when fake news gets intimate and everyone disagrees on everything. Retrieved from <http://www.niemanlab.org/>
- Papadopoulou, O., Zampoglou, M., Papadopoulos, S., & Kompatsiaris, I. (2019). A corpus of debunked and verified user-generated videos. *Online Information Review*, 43(1), 72-88.
- Reid, A., & Sands, P. (2016). Tools and tricks for truth seekers: Why people need to learn verification techniques to combat hoaxes and misinformation on social media. *Index on Censorship*, 45(1), 83-87.
- Statista (2018). Facebook's advertising revenue worldwide from 2009 to 2017, Retrieved from: <https://www.statista.com/statistics/271258/facebooks-advertising-revenue-worldwide/>
- Teyssou, D., Leung, J. M., Apostolidis, E., Apostolidis, K., Papadopoulos, S., Zampoglou, M., ... & Mezaris, V. (2017, October). The InVID plug-in: web video verification on the browser. In *Proceedings of the First International Workshop on Multimedia Verification* (pp. 23-30). ACM.
- Vryzas, N., Katsaounidou, A., Kotsakis, R., Dimoulas, C. A., & Kalliris, G. (2019, March). Audio-Driven Multimedia Content Authentication as a Service. In *Audio Engineering Society Convention 146*. Audio Engineering Society.
- Wardle, C., & Derakhshan, H. (2017). Information Disorder: Toward an interdisciplinary framework for research and policymaking. Council of Europe report, DGI (2017), 9.
- Zampoglou, M., Papadopoulos, S., & Kompatsiaris, Y. (2017). Large-scale evaluation of splicing localization algorithms for web images. *Multimedia Tools and Applications*, 76(4), 4801-4834.
- Zimdars, M. (2016). False, misleading, click bait-y, and satirical “news” sources. Google Docs.