In Video Veritas – Verification of Social Media Video Content for the News Industry

Fake news based on video reuse and how to deal with it: video fragmentation and reverse image search

Evlampios Apostolidis, CERTH-ITI

Thessaloniki, December 2017

www.invid-project.eu

Fake news based on video reuse

- How does this fake work?
  - A previously existing video is reused under a different and irrelevant context, aiming to deliberately mislead the viewers about a fact/event

Fake news based on video reuse

- How does this fake work?
  - A previously existing video is reused under a different and irrelevant context, aiming to deliberately mislead the viewers about a fact/event

5/9/2017; Claim: Hurricane "Irma" in the islands near the US

Fake news based on video reuse

- How does this fake work?
  - A previously existing video is reused under a different and irrelevant context, aiming to deliberately mislead the viewers about a fact/event

5/9/2017; Claim: Hurricane "Irma" in the islands near the l

24/11/2016; Claim: Hurricane Otto in Bocas del Toro, Panama

Fake news based on video reuse

- How does this fake work?
  - A previously existing video is reused under a different and irrelevant context, aiming to deliberately mislead the viewers about a fact/event

5/9/2017; Claim: Hurricane "Irma" in the islands near the l

24/11/2016; C Otto in Bocas d

29/5/2016; Claim: Hurricane in Dolores, Uruguay

Fake news based on video reuse

- How does this fake work?
  - A previously existing video is reused under a different and irrelevant context, aiming to deliberately mislead the viewers about a fact/event

5/9/2017; Claim: Hurricane "Irma" in the islands near the l

24/11/2016; C Otto in Bocas d

29/5/2016; Claim: Hurricane in Dolores, Uruguay

Fake news based on video reuse

- How does this fake work?
  - A previously existing video is reused under a different and irrelevant context, aiming to deliberately mislead the viewers about a fact/event

19/3/2017; Claim: Migrant seeking free healthcare in a public hospital in France

Fake news based on video reuse

- How does this fake work?
  - A previously existing video is reused under a different and irrelevant context, aiming to deliberately mislead the viewers about a fact/event

19/3/2017; Claim: Migrant seeking free healthcare in a public hospital in France

26/2/2017; Claim: Drunk patient in a hospital in Novgorod, Russia

Fake news based on video reuse

- How does this fake work?
  - A previously existing video is reused under a different and irrelevant context, aiming to deliberately mislead the viewers about a fact/event

19/3/2017; Claim: Migrant seeking free healthcare in a public hospital in France

26/2/2017; Claim: Drunk patient in a hospital in Novgorod, Russia

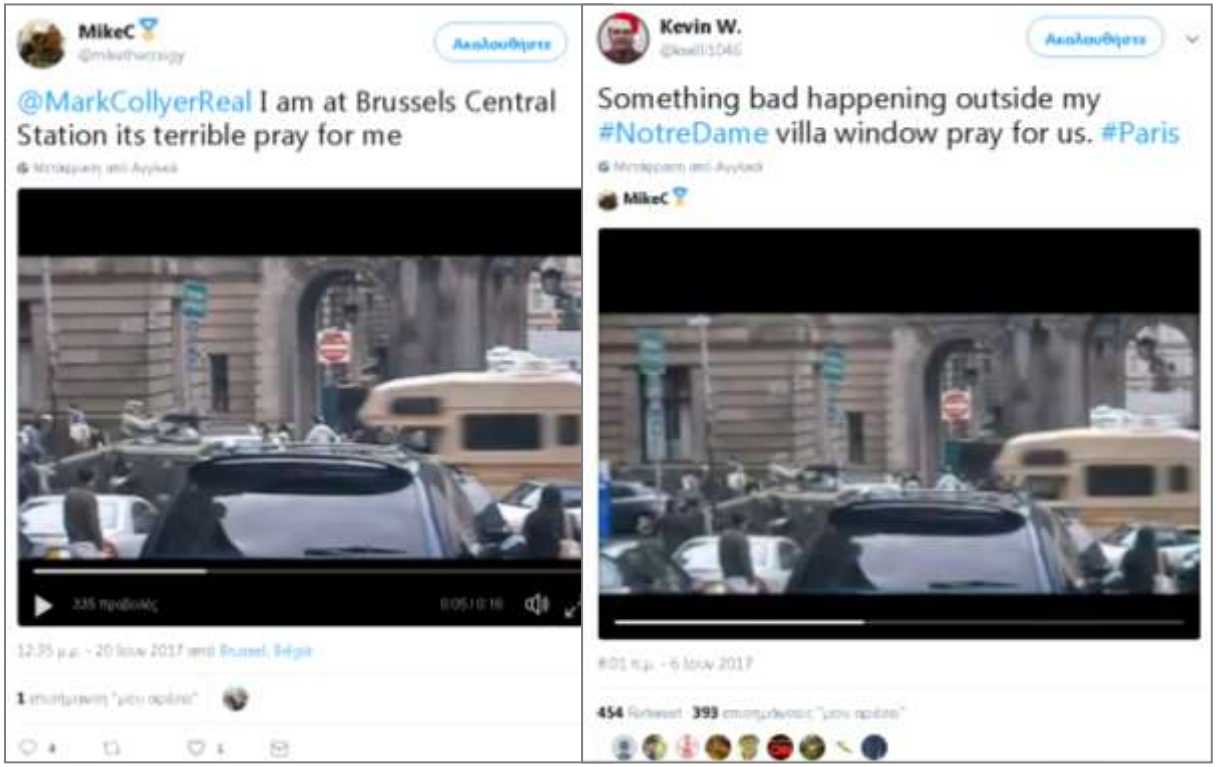
Fake news based on video reuse

- How does this fake work?
  - A previously existing video is reused under a different and irrelevant context, aiming to deliberately mislead the viewers about a fact/event

20/6/2017; Claim: Attack at Gare Centrale, Brussels, Belgium

Fake news based on video reuse

- How does this fake work?
  - A previously existing video is reused under a different and irrelevant context, aiming to deliberately mislead the viewers about a fact/event



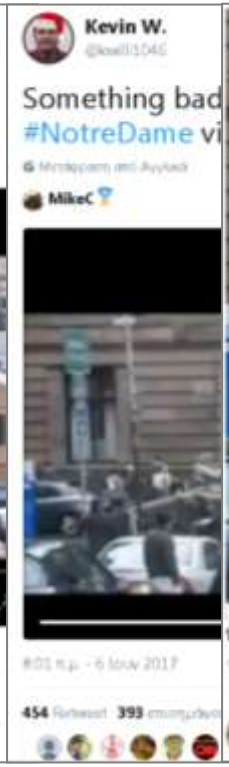
20/6/2017; Claim: Att Centrale, Brussels,

6/6/2017; Claim: Hammer attack against police in Notre-Dame, Paris, France

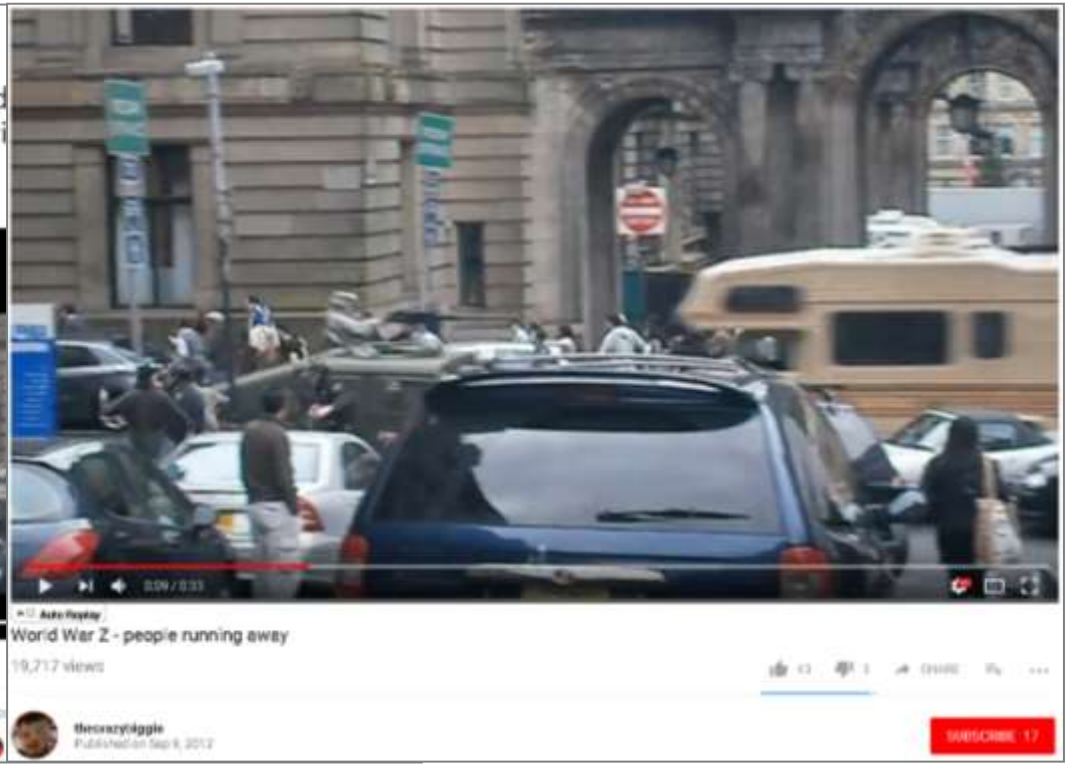
Fake news based on video reuse

- How does this fake work?
  - A previously existing video is reused under a different and irrelevant context, aiming to deliberately mislead the viewers about a fact/event

20/6/2017; Claim: Att Centrale, Brussels,



6/6/2017; Claim: Hamme police in Notre-Dame,

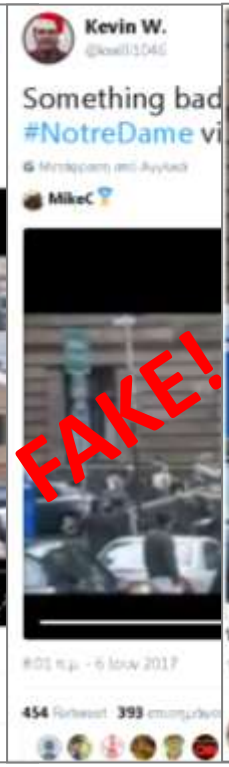


9/9/2012; Claim: Making of the movie "World War Z"

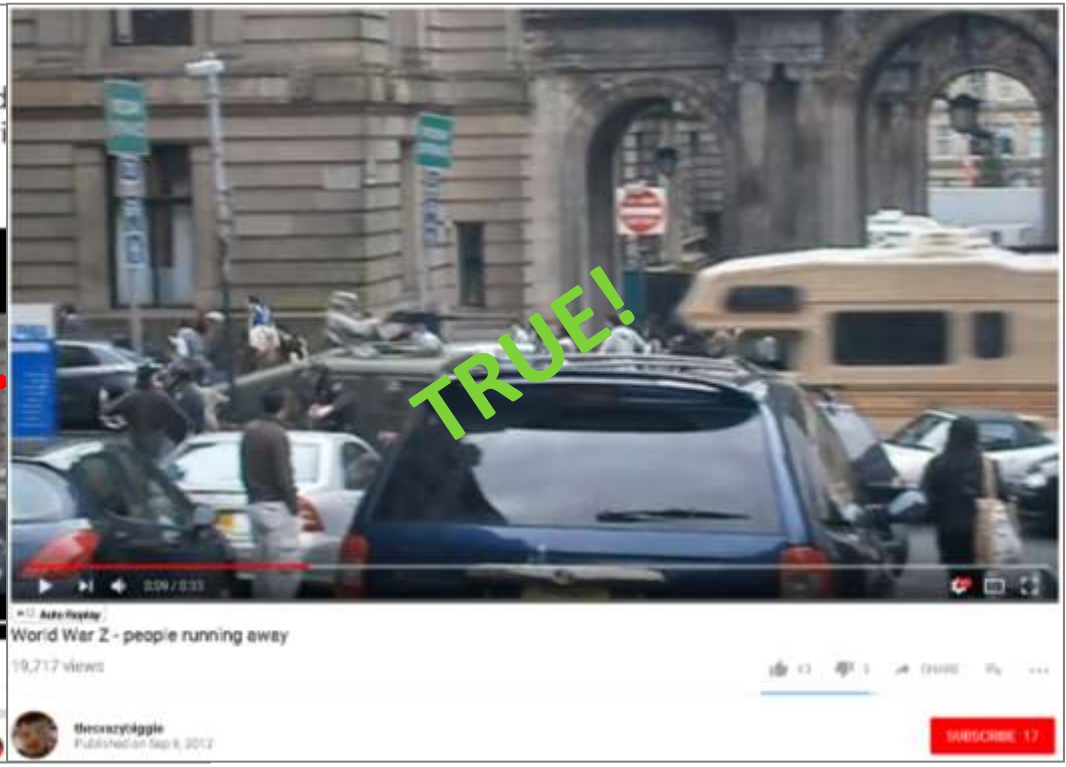
Fake news based on video reuse

- How does this fake work?
  - A previously existing video is reused under a different and irrelevant context, aiming to deliberately mislead the viewers about a fact/event

20/6/2017; Claim: Att Centrale, Brussels,



6/6/2017; Claim: Hamme police in Notre-Dame,



9/9/2012; Claim: Making of the movie "World War Z"

Fake news based on video reuse

- How does this fake work?
  - A previously existing video is reused under a different and irrelevant context, aiming to deliberately mislead the viewers about a fact/event

24/11/2017; Claim: Attack of al-Rawda mosque in Sinai, Egypt

Fake news based on video reuse

- How does this fake work?
  - A previously existing video is reused under a different and irrelevant context, aiming to deliberately mislead the viewers about a fact/event

24/11/2017; Claim: Attack of al-Rawda mosque in Sinai, Egypt

29/5/2015; Claim: Attack of Imam Hussein mosque in Dammam, Saudi Arabia

Fake news based on video reuse

- How does this fake work?
  - A previously existing video is reused under a different and irrelevant context, aiming to deliberately mislead the viewers about a fact/event

24/11/2017; Claim: Attack of al-Rawda mosque in Sinai, Egypt

29/5/2015; Claim: Attack of Imam Hussein mosque in Dammam, Saudi Arabia

Dealing with this type of fakes

- How journalists deal with this kind of fake news?
  - Try to find previous occurences of the video on the Web

Dealing with this type of fakes

- How journalists deal with this kind of fake news?
  - Try to find previous occurences of the video on the Web
- What processes/tools do they follow/use for performing reverse video search?

Dealing with this type of fakes

- How journalists deal with this kind of fake news?
  - Try to find previous occurences of the video on the Web
- What processes/tools do they follow/use for performing reverse video search?
  - Taking screenshots of the video and doing reverse search of these screenshots with the help of the Google Images engine

Dealing with this type of fakes

- How journalists deal with this kind of fake news?
  - Try to find previous occurences of the video on the Web
- What processes/tools do they follow/use for performing reverse video search?
  - Taking screenshots of the video and doing reverse search of these screenshots with the help of the Google Images engine
  - Using search-engine-based plug-ins, such as RevEye[1] and TinEye[2], that allow reverse image search of Web images related to a video

1. https://goo.gl/ZRHTDH
2. https://goo.gl/G5Rpu9

Dealing with this type of fakes

- How journalists deal with this kind of fake news?
  - Try to find previous occurences of the video on the Web
- What processes/tools do they follow/use for performing reverse video search?
  - Taking screenshots of the video and doing reverse search of these screenshots with the help of the Google Images engine
  - Using search-engine-based plug-ins, such as RevEye[1] and TinEye[2], that allow reverse image search of Web images related to a video
  - Using the YouTube DataViewer[3], which supports reverse search of YouTube video thumbnails

1. https://goo.gl/ZRHTDH
2. https://goo.gl/G5Rpu9
3. https://citizenevidence.amnestyusa.org/

Dealing with this type of fakes

- How journalists deal with this kind of fake news?
  - Try to find previous occurences of the video on the Web
- What processes/tools do they follow/use for performing reverse video search?
  - Taking screenshots of the video and doing reverse search of these screenshots with the help of the Google Images engine
  - Using search-engine-based plug-ins, such as RevEye[1] and TinEye[2], that allow reverse image search of Web images related to a video
  - Using the YouTube DataViewer[3], which supports

    reverse search of YouTube video thumbnails
  - Time-consuming and cumbersome processes that:
    - either involve manual generation and uploading of video screenshots
    - or rely on the use of a limited set of video thumbnails

The InVID solution

- Interactive tool for reverse video search on the Web
- Time-efficient process that requires minimum manual intervention
- Fine-grained search at the video-fragment-level, through:
  - segmentation of the video into visually coherent fragments
  - extraction of representative keyframes for each video fragment
  - reverse search of these keyframes via the Google search engine

The InVID solution

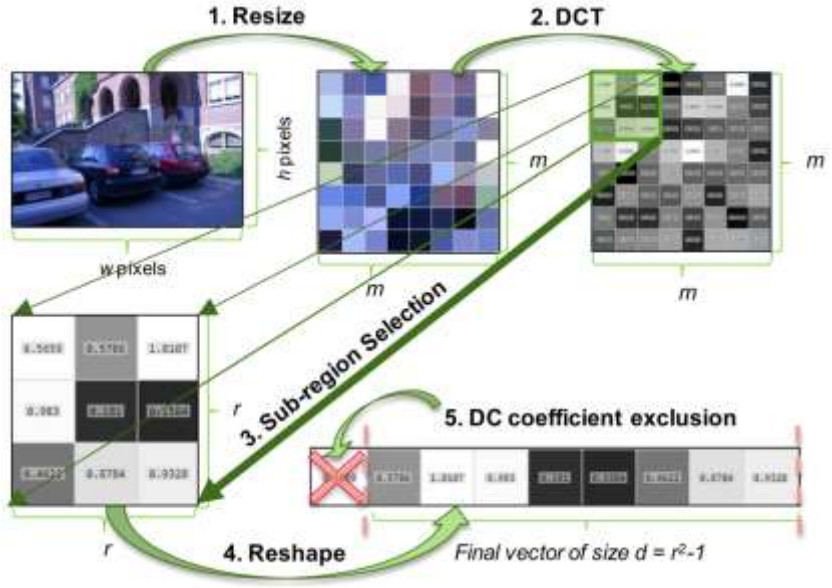
- Interactive tool for reverse video search on the Web
- Time-efficient process that requires minimum manual intervention
- Fine-grained search at the video-fragment-level, through:
  - **segmentation of the video into visually coherent fragments**
  - **extraction of representative keyframes for each video fragment**
  - reverse search of these keyframes via the Google search engine

Video fragmentation & keyframe selection

- User-generated videos (UGVs) are captured without interruption using a single camera, thus, being single-shot videos
- Algorithms for shot boundary detection (e.g. [1]) fail to reveal information about the structure of these videos
- A more fine-grained segmentation into sub-shots, is needed!
- Proposed approaches define video sub-shots, as:
  - sequences of frames with a small variation in their visual content, based mainly on pair-wise evaluation of frames' visual similarity/dissimilarity [3-7]
  - sequences of frames corresponding to different video recording actions (e.g. camera pan/tilt, camera zoom in/out), relying on motion extraction and classification using pre-defined motion models [8-10] or pre-trained systems [11-13]

Video fragmentation & keyframe selection

- The InVID approach [2]
  - The visual content of each frame is represented with the help of a 2D Discrete Cosine Transform (see figure)
  - Video fragmentation into sub-shots is performed by assessing the visual resemblance of neighboring frames using the cosine similarity
  - The algorithm indicates both sub-shots with minor or no activity, and sub-shots with gradually, but consistently, changing visual content
  - As representative keyframe:
    - for the former type of sub-shots the middle frame is selected
    - for the latter type of sub-shots the frame with the most pronounced change of visual content is selected
  - The analysis takes approx. 3% of the video's duration (being more than 30 times faster than real-time processing)



The reverse video search tool

- Available at: http://multimedia3.iti.gr/videofragmentation_v5/service/start.html
- Allows the analysis of both online and locally stored videos
  - Supported platforms: YouTube, DailyMotion, Facebook, Twitter, Dropbox
  - Supported video formats: mp4, webm, avi, mov, wmv, ogv, mpg, flv, mkv

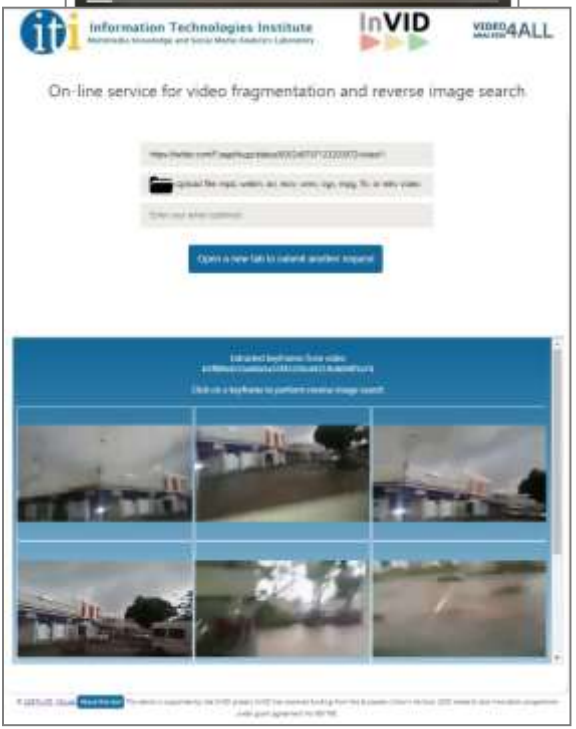
- The user can monitor the progress of the analysis, or close the browser and be notified (by e-mail) when the results are ready
- After the analysis ends, s/he gets the set of extracted keyframes and can perform reverse search by left clicking on any of them

Case study

- Debunking a fake video about hurricane Irma

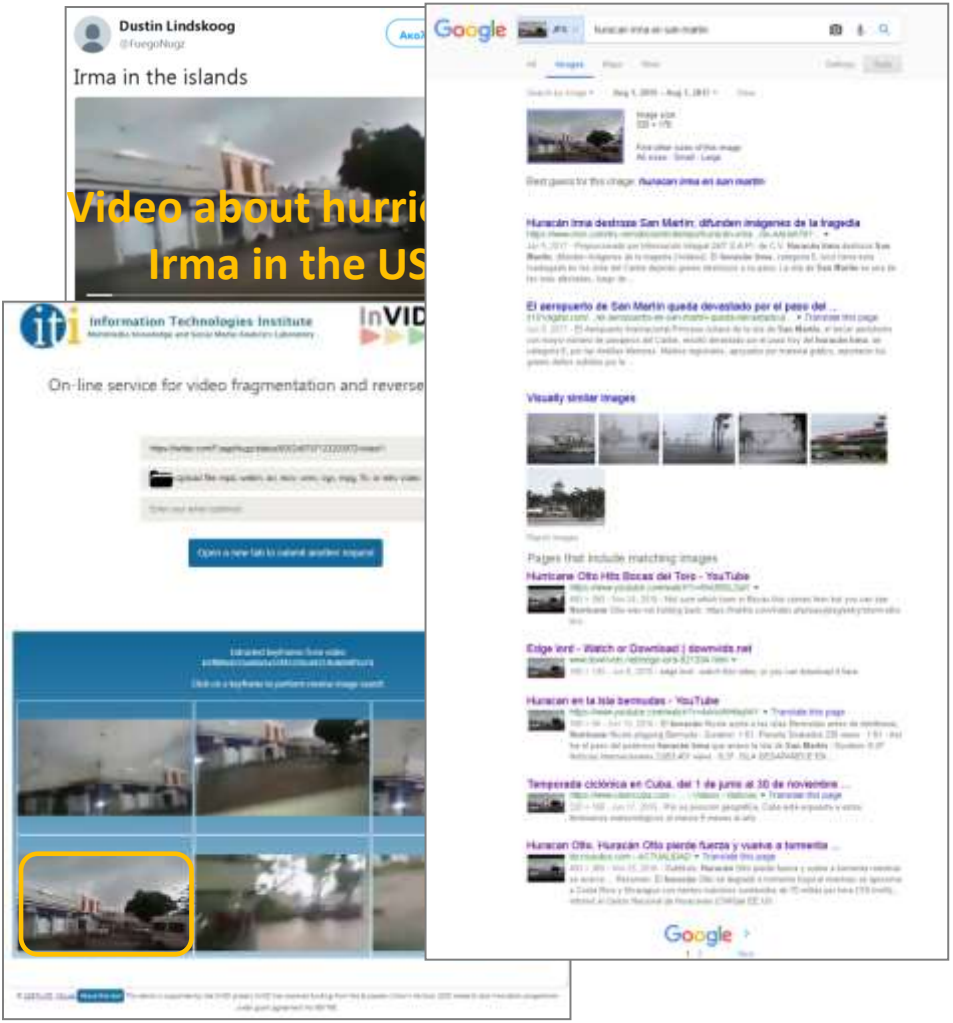
Case study

- Debunking a fake video about hurricane Irma



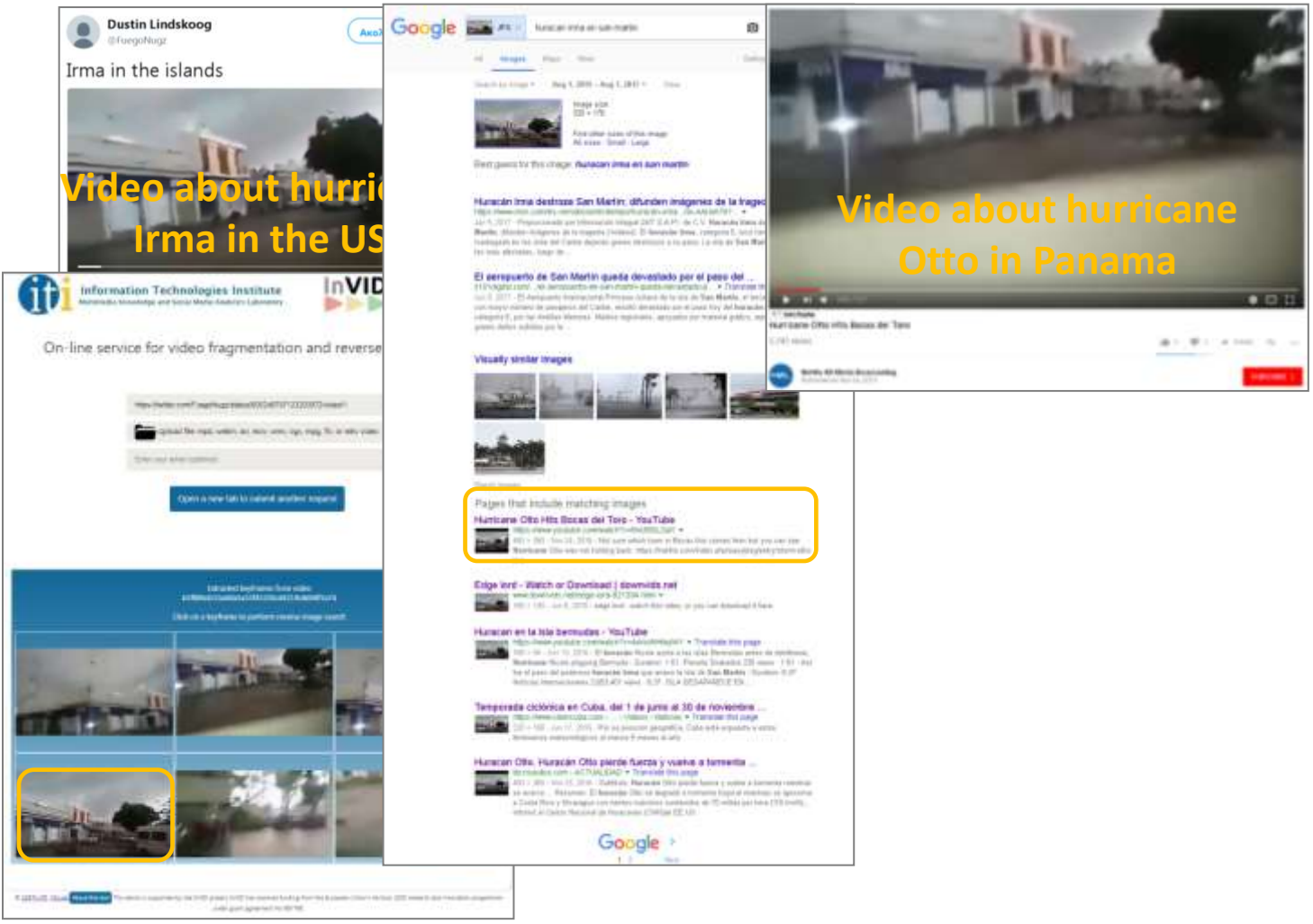
Case study

- Debunking a fake video about hurricane Irma



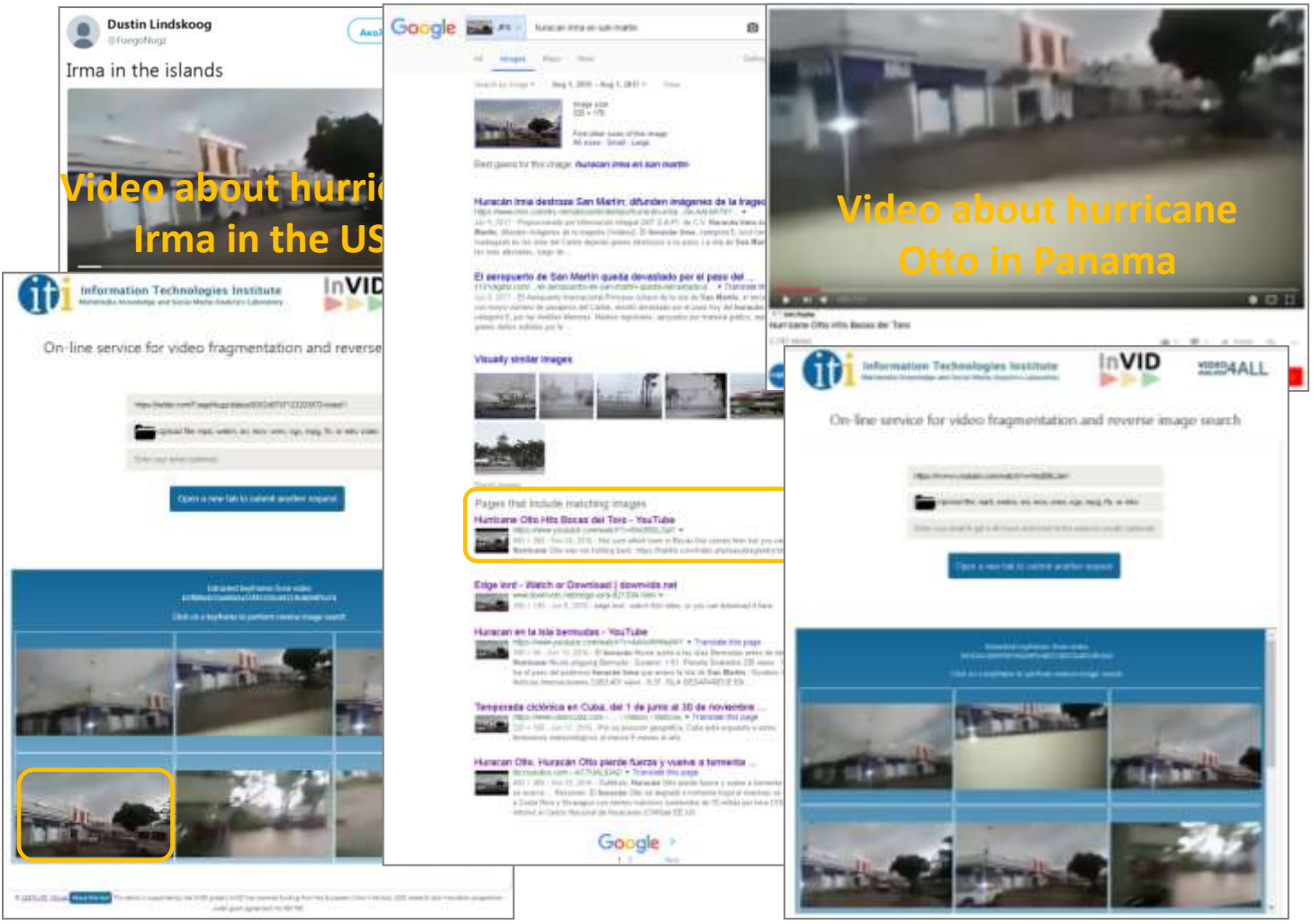
Case study

- Debunking a fake video about hurricane Irma



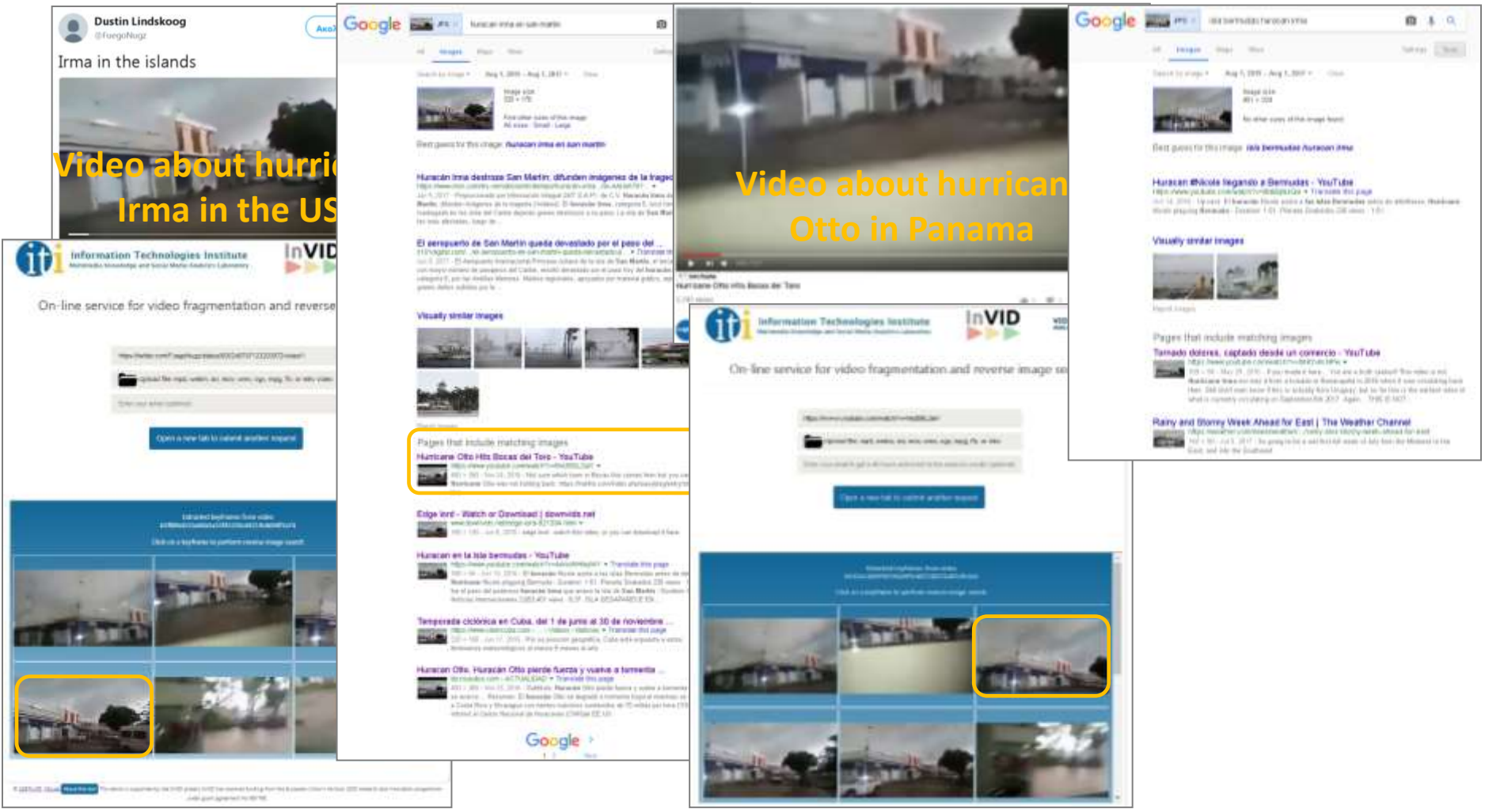
Case study

- Debunking a fake video about hurricane Irma



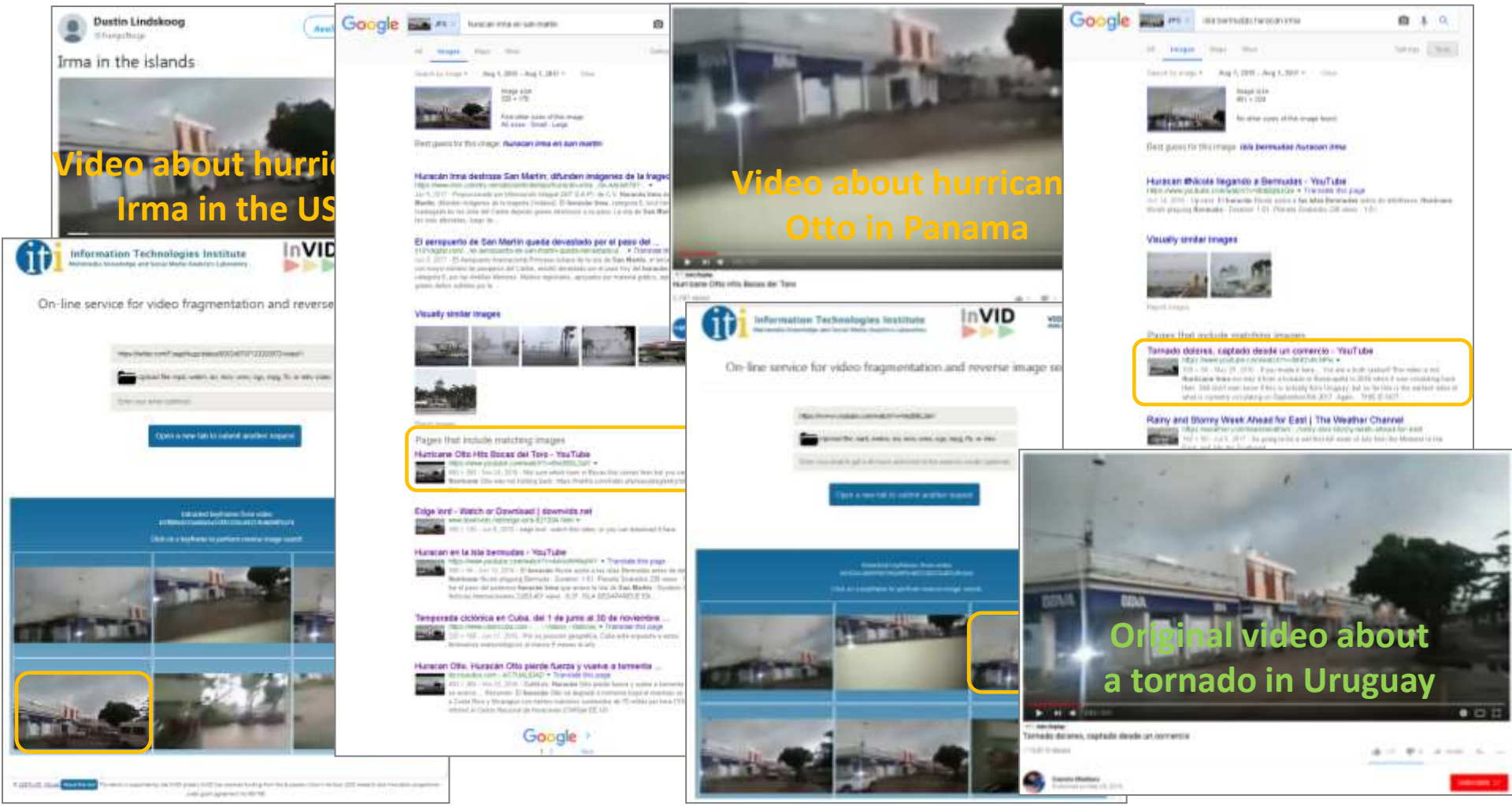
Case study

- Debunking a fake video about hurricane Irma



Case study

- Debunking a fake video about hurricane Irma



Summary and outlook

- The developed tool for video fragmentation and reverse image search facilitates the detection of previously existing occurrences of a published video on the Web
- Its interactive UI makes the detection of such videos a “few-clicks” process that requires minimum manual intervention
- The gathered feedback from journalists and media verification experts (through its integration into the InVID Verification Plugin) is very positive and encouraging
- Improvements are foreseen regarding:
  - the keyframe selection process
  - the compatibility with online video platforms
  - the detection of mirrored videos

Publications

1. Apostolidis, E., et al.: Fast shot segmentation combining global and local visual descriptors. In: Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing. pp. 6583-6587 (2014), software available at http://mklab.iti.gr/project/video-shot-segm
2. Teyssou, D. , et al.: The InVID Plug-in: Web Video Verification on the Browser. In: Proc. Int. Workshop on Mult. Verification at ACM Multimedia Conf. 2017, Mountain View, CA, USA, October 2017, software available at http://www.invid-project.eu/verify
3. Pan, C.M., et al.: NTU TRECVID-2007 fast rushes summarization system. In: Proc. of the 1st ACM TRECVID Video Summarization Workshop. pp. 74-78 (2007)
4. Dumont, E., et al.: Rushes video summarization using a collaborative approach. In: Proc. of the 2nd ACM TRECVID Video Summarization Workshop. pp. 90-94 (2008)
5. Bai, L., et al.: Automatic summarization of rushes video using bipartite graphs. Mult. Tools and Appl. 49(1), 63-80 (2010)
6. Liu, Y., et al.: Rushes video summarization using audio-visual information and sequence alignment. In: Proc. of the 2nd ACM TRECVID Vid. Summar. Workshop. pp. 114-118 (2008)
7. Ojutkangas, O., et al.: Location based abstraction of user generated mobile videos, pp. 295-306. Springer Berlin Heidelberg (2012)
8. Mei, T., et al.: Near-lossless semantic video summarization and its applications to video analysis. ACM Trans. Mult. Comput. Commun. Appl. 9(3), 16:1-16:23 (2013)
9. Cooray, S.H., et al.: An interactive and multi-level framework for summarising user generated videos. In: Proc. of the 17th ACM Int. Conf. on Mult. pp. 685-688 (2009)
10. Nitta, N., et al.: Content analysis for home videos. ITE Trans. on Media Tech. and Appl. 1(2), 91-100 (2013)
11. Abdollahian, G., et al.: Camera motion-based analysis of user generated video. IEEE Trans. on Mult. 12(1), 28-41 (2010)
12. Karaman, S., et al.: Hierarchical Hidden Markov Model in detecting activities of daily living in wearable videos for studies of dementia. Mult. Tools and Appl. 69(3), 743-771 (2014)
13. Gonzalez-Daz, I., et al.: Temporal segmentation and keyframe selection methods for user generated video search-based annotation. Expert Syst. Appl. 42(1), 488-502 (2015)

In Video Veritas!

Thank you for your attention!

Any questions?