Title: When automated analysis goes wrong

What do you do when your fancy neural network for image analysis tells you that one of your modern artworks is "a pizza sitting on top of a window"? Or that your seventeenth-century oil-painting of St. Francis clutching a crucifix is a "person using a phone"? And how do you cope with scikit-image telling you that a photo of the Eiffel Tower is more similar to one of the Statue of Liberty than another taken at a slightly different angle? And finally, how did it come to pass that an automated text-analysis programme quite literally said that the Met's collection was full of shit?

I will discuss various techniques for dealing with the uncertainties that can happen when machines analyse images and text in bulk, using real-life examples from the collections of various institutions including the Metropolitan Museum of Art, the Nationalmuseum of Sweden, the Barnes Foundation and the Qatar Digital Library. These include constraining results; altering probability thresholds; and deliberately withholding information. Plus the all-important strategies of knowing when to change your approach, when to re-introduce humans to the mix, and when to just give up and laugh at the silly computers.
Bio:
Tristan is a digital media professional with over ten years experience working with museums, libraries and archives. Since joining Cogapp in 2004 he has worked on a huge variety of projects, including online collections (Metropolitan Museum of Art, National Portrait Gallery, MoMA) and online archives (Qatar Digital Library, Endangered Archives Programme).

He loves data wrangling, server-side programming and IIIF, which he has been involved with since 2014.