What digital tools will reshape the discipline of history? Will it be topic models? Or geospatial analysis? Or is it maybe the subtler shift towards online, full-text searchable archives, which enable us to effortlessly track our actors across time and space through keyword searches? And if so, what are the downsides of this shift? How, for example, can we prevent large-scale archives from turning into echo chambers that only provide us with primary sources that confirm our assumptions instead of challenging them? In this talk, I will repurpose my primary source base—an online archive of more than 10 million formerly secret tobacco industry documents—into a laboratory for tackling these questions.

After exploring the curious material history of this now digital depository and highlighting the ways that access to these documents has changed public health historiography for both better and worse, I will demonstrate new quantitative tools that I have developed for Tobacco Analytics (www.tobacco-analytics.org) to query and visualize the millions of documents in this archive all at once. Digital archives provide us with a vantage point from which we can explore how the internet, OCR, and full-text search have decisively changed which sources we can and cannot find; how we access them; and ultimately how we write history. Digital history, I want to suggest, is not the future of our discipline but the present.