



E-DISCOVERY TRENDS 2017: WEB CONTENT COLLECTION

A background image showing a woman with curly hair speaking to a man in a suit who is looking at a document. The image is overlaid with a semi-transparent blue filter and a geometric pattern of white lines on the right side.



Table of Contents

Introduction	3
Who Collects Web Content and How They Search	4
Where Content is Being Found on the Internet	6
Social Media Platforms and Relevant Content	8
Web Evidence Admissibility Concerns	10
Safe Searching for Potentially Dangerous Content and Websites	11
Conclusion	12



INTRODUCTION

Many legal professionals today use the internet to research their cases. When interesting content is found, it becomes necessary to collect that web-based content. However, the specifics on how web content is collected, where it is located online, and what is captured, changes from case to case.

This report shares the latest web content collection trends identified by litigators. It is designed to help all legal professionals to not only see how and what their colleagues are collecting, but also to improve their own collection practices and processes.

We surveyed litigators from a variety of states, law firms and practice areas, including general litigation, intellectual property law, cybersecurity, insurance fraud, class action and commercial law. Our purpose was to look at how litigators and law firms search for, collect and use web-based content for their cases.



Web content collection is the act of saving and/or printing web content, such as social media profiles, websites, videos, images, or any other public content accessible on a web browser that is intended to be used for due diligence or as court evidence.

WHO COLLECTS WEB CONTENT AND HOW THEY SEARCH



A majority of the litigators we surveyed identified at least half of their cases involve collecting web content.

And because it is becoming more and more common to collect web content as evidence or for research, it is no surprise that over 90% of litigators consider it a best practice to capture relevant web content for a case matter, when available.

However, because web collection is prevalent in most cases today, the amount of time it takes to search through high volumes of relevant content, as well as collect it, means litigators are not the only ones collecting such content. Apart from themselves, litigators identified many other professionals who might need to capture web content: paralegals, litigation support, information technology, and library or research service professionals (*Figure 1*).

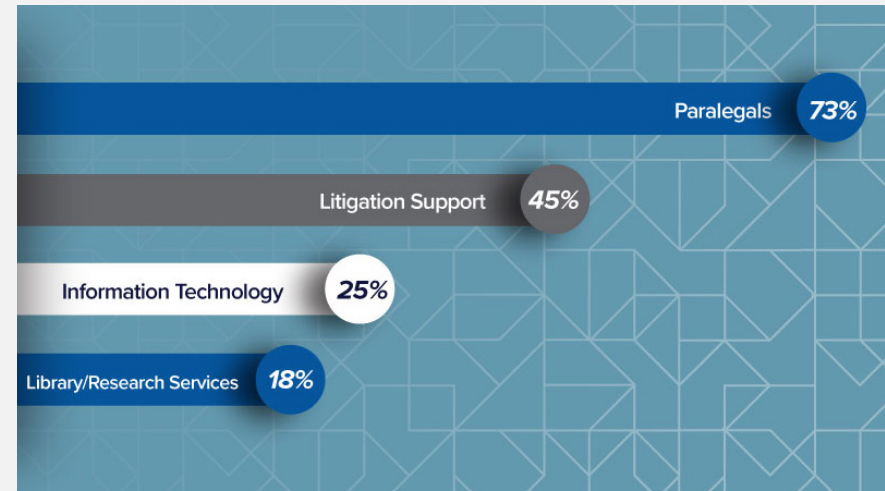


Figure 1 - Other professional roles that collect web content within the litigator's firm

When legal teams begin to research relevant parties online for a case, 82% of those surveyed said they first use a search engine like Google or Bing. Searches vary based on the case facts, but litigators identified the top nine types of information they typically look for when researching online for an individual (Figure 2).

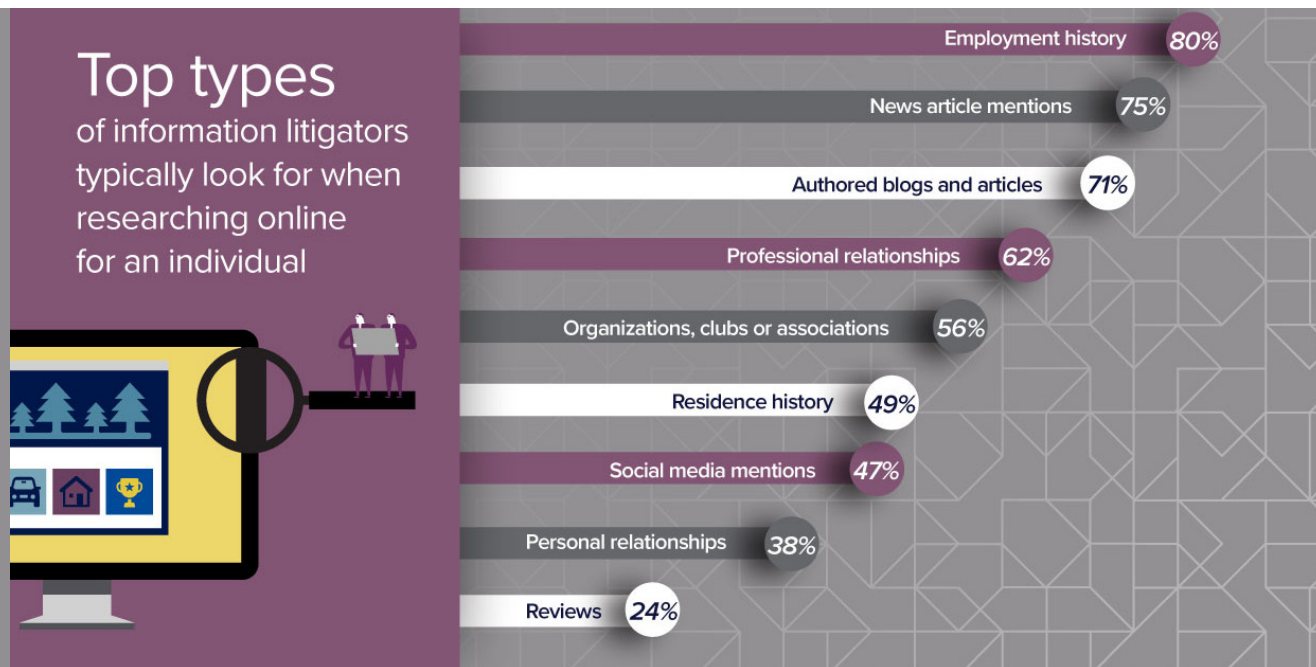


Figure 2

When relevant content is found, it is not always just a one-time capture, as most online content is dynamic and changes. 80% of litigators stated that they have needed to monitor and collect the same webpage or social profile over time to take

that into consideration. And that is only if they are quick enough to capture the content. **87% said that the content was later changed or even gone when they went back to it.**

WHERE CONTENT IS BEING FOUND ON THE INTERNET



The internet has many types of discoverable content, but there are certain types that are more frequently collected than others.

From the litigators surveyed, here is a look at some of the most common web content that is collected from the practice areas of general litigation, intellectual property law, cybersecurity, insurance fraud, class action and commercial law (*Figure 3*).

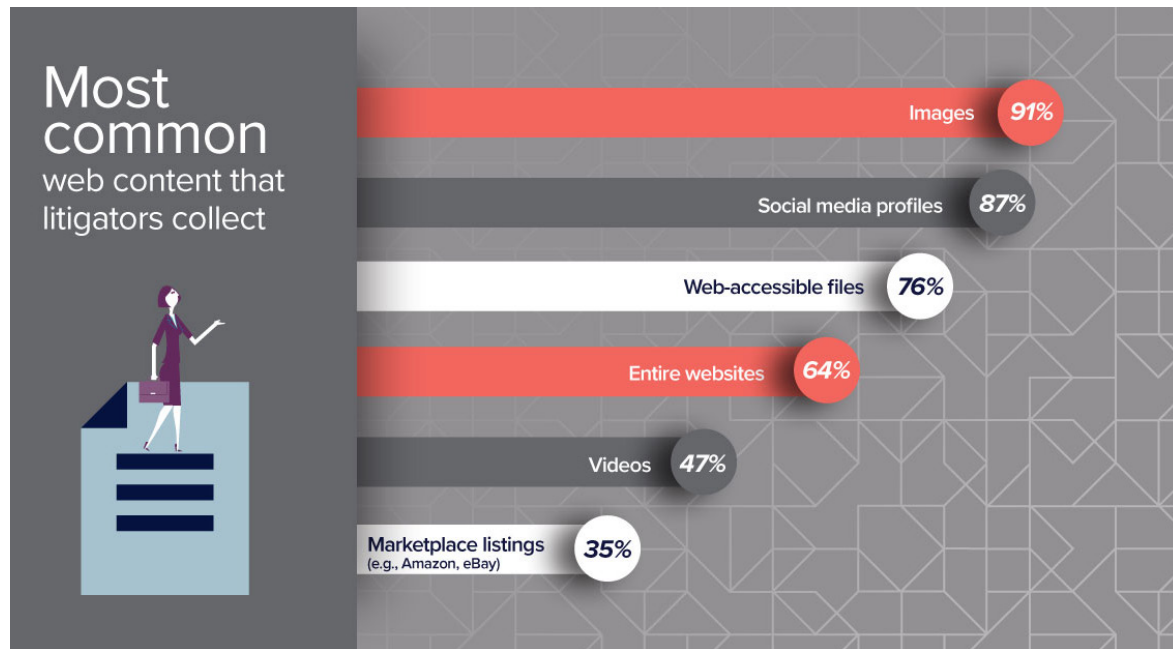


Figure 3

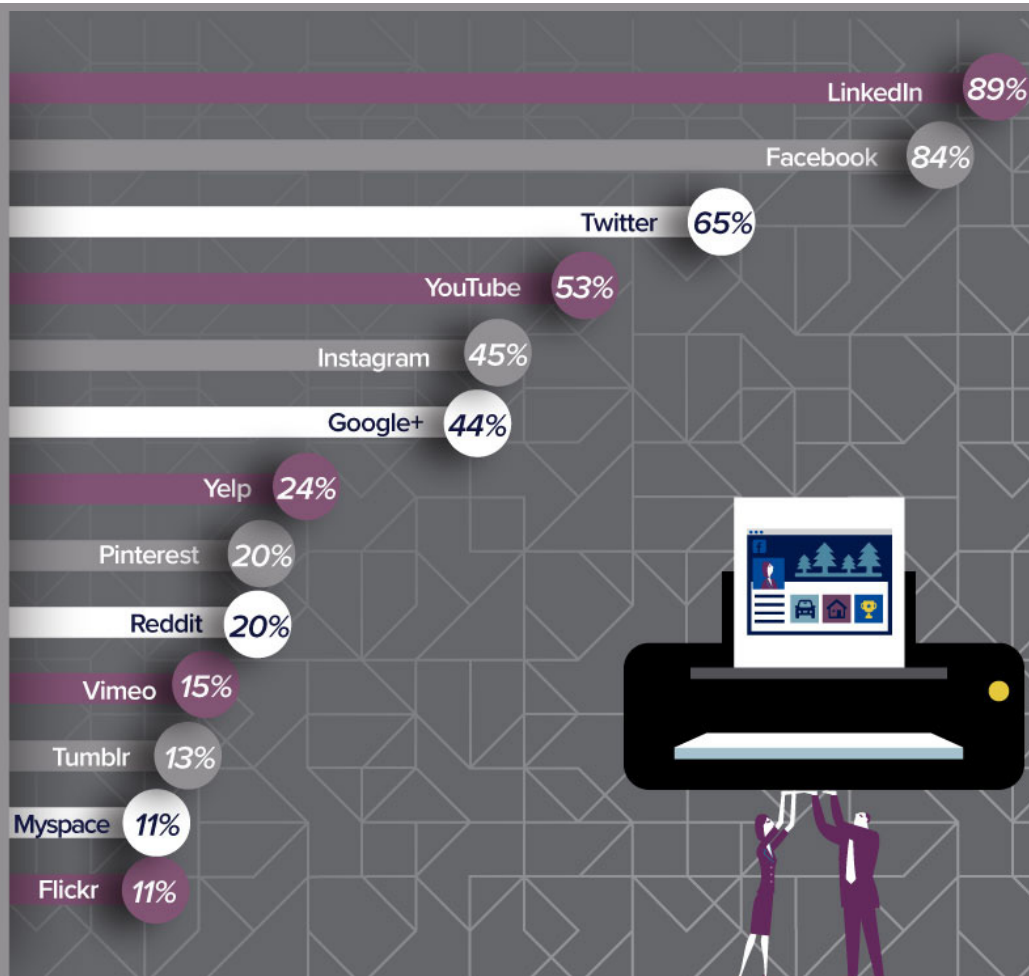


It is easy to assume that experienced litigators know where to search online for web content, but that is not always the case.

With a plethora of places to find information about certain parties, it is no wonder that **60% of surveyed litigators worry that they are not looking in the right place for relevant web content.** This statistic signifies relevant content may go undiscovered, which could be the difference between winning or losing a case. By having solutions designed for the legal field, litigators can easily capture complete content, such as all pages of a website or all posts and comments of a social media profile.

SOCIAL MEDIA PLATFORMS AND RELEVANT CONTENT

Legal professionals commonly struggle with searching through all of the relevant content on social media platforms. As those platforms continue to grow their user base, so does the amount of content that is published online. Because of this, social media can be a rich place to look for content and data that can either confirm or dispute case facts.



With this type of content on the rise, we thought it would be worth diving deeper into social media content collection by asking litigators where and what they are looking for on social media platforms.

When asked what social media platforms they use when searching for information, 89% of litigators said LinkedIn, 84% said Facebook, and 65% said Twitter. While we expected to see LinkedIn, Facebook and Twitter as the top social media platforms (Figure 4), we were surprised to find a high percentage of litigators still searching on less common platforms like Reddit, Yelp and Flickr.

Figure 4 - Social media platforms litigators use when searching for information



When searching any social media platform, content can hide in places you may not initially think to look. The saying “you don’t know what you don’t know” is certainly true when it comes to locating places within social media platforms that may have relevant content to capture.

With so many social media platforms and different features within each, it can be very difficult to know where to look. Here is a glance at the most common social media content that litigators review and capture during their web content collection process (*Figure 5*).

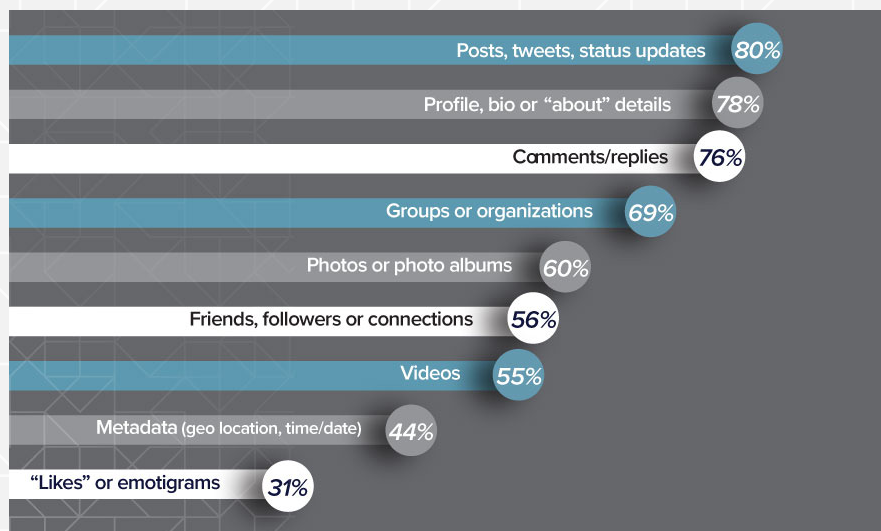


Figure 5 - Common social media content litigators review & capture

We discovered that only 44% of litigators look at social media metadata, which could include geodata like time, date and place. It is worth noting that the vast majority of litigators who are *not* looking at social media metadata could be missing vital content to support their cases. Social media metadata can be very valuable, especially in cases where time and place should be noted.

WEB EVIDENCE ADMISSIBILITY CONCERNS



Many litigators are concerned about admissibility of their web content captures when not using a tool designed for legal, and 73% feared that they were not collecting content correctly or were violating the chain of custody.

Including metadata with your web content collections will help to support admissibility and reduce concerns associated with authentic web evidence. When asked what metadata they find valuable to accompany collected web content, 96% of litigators said the time and date of the capture, 80% said the website URL of the content, 49% said IP addresses of the capturing computer and web server, 44% said the person making the capture, and 42% said the webpage source code (HTML).

Another concern is the accuracy of the electronic or hard copy of a web capture. 84% of litigators said that they had issues with web capture output because the appearance was not accurate. When trying to accurately capture online content, such as lengthy webpages or social media profiles, without a professional-grade service or tool, it can produce a capture that does not represent exactly what is seen online. For instance, images and content could be missing, or the page format could be misaligned – all working against the authenticity of the web capture.



SAFE SEARCHING FOR POTENTIALLY DANGEROUS CONTENT AND WEBSITES



With so many options available for capturing web content only a small fraction are designed specifically for the legal industry's unique needs.

Often times, cases require research or evidence collection on websites that could potentially harm or track your computer if capturing web content through a basic browser like Google Chrome, Internet Explorer, or Mozilla Firefox.

Most litigators use a basic browser as part of their capture method, which is not secure for legal use. Further, many of the search activities are documented and tracked, which means if a legal professional is trying to be anonymous, they are unintentionally increasing their chances of drawing attention to themselves.

It is no surprise to see then that 58% of litigators worried their investigation would somehow alert the other party of their search activity (Figure 6). In addition, 35% were

concerned about getting a virus or malware from an untrusted website that they needed to visit to gather relevant content for their case.

Alternatively, capturing web content through a [tool or service that uses a specialized, remote browser](#) is the most secure way to avoid these concerns. This type of browser will connect to a separate server when capturing content, versus your personal or firm's computer, thereby decreasing security risks and preserving the chain of custody.



Figure 6



Legal professionals in all practice areas are encouraged to implement the best practices laid out in this report to improve their web content collection process.

[Page Vault](#) helps litigators and other legal professionals stay on top of web content collection best practices. Our solutions not only increase content admissibility, but we also educate the legal industry on web content trends that apply to current and future cases.

Learn more at www.page-vault.com.