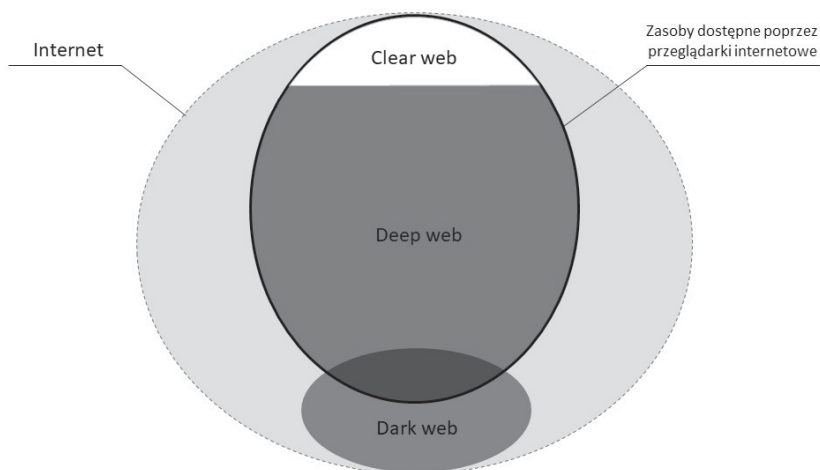




W rozważaniach na temat wyszukiwania danych w Internecie najczęściej domyślnie stosuje się uproszczenie, przypisujące sieci indeksowanej wszystkie lub niemal wszystkie dostępne dla użytkownika zasoby. Rzeczywistość jest jednak zupełnie odmienna, gdyż według różnych źródeł objętość sieci indeksowanej to jedynie od około 1 do 4%<sup>35</sup> (niektóre opracowania podają maksymalnie 10%<sup>36</sup>) zasobów całego Internetu. Według części badaczy *deep web* jest 500-krotnie lub nawet 5000-krotnie razy większa od sieci indeksowanej, a wyszukiwarki mają dostęp jedynie do 16% danych zawartych w tej ostatniej<sup>37</sup>. Pozostałą, ogromną część zasobów stanowi sieć nieindeksowana *deep web*, czyli głównie:

- tematyczne bazy danych o ograniczonym dostępie,
- strony wymagające opłaty (np. gazety z tzw. *paywallem*),
- fora dyskusyjne,
- zasoby dostępne z użyciem technologii stanowiącej barierę dla wyszukiwarek (np. sieci dostępne przez VPN lub wykorzystujące niestandardowe protokoły/formaty danych),
- strony, do których linki nie są nigdzie opublikowane,
- zasoby darknet.



### RYS. 3.

Podział Internetu z perspektywy możliwości przeglądania poszczególnych jego części przy wykorzystaniu przeglądarek internetowych

Źródło: opracowanie własne.

<sup>35</sup> *How Much of the Internet is the Dark Web in 2023?*, <https://techjury.net/blog/how-much-of-the-internet-is-the-dark-web>, dostęp online 10.11.2023 r.; [Saleem, Islam, Kabir, 2022, s. 33628–33660].

<sup>36</sup> [Coffey, 2020].

<sup>37</sup> [Król, 2019, s. 53–66].