



Authors' contribution/
Wkład autorów:
A. Study design/
Zaplanowanie badań
B. Data collection/
Zebranie danych
C. Statistical analysis/
Analiza statystyczna
D. Data interpretation/
Interpretacja danych/
E. Manuscript
preparation/
Przygotowanie tekstu
F. Literature search/
Opracowanie
piśmiennictwa
G. Funds collection/
Pozyskanie funduszy

THE LINK ROT PHENOMENON AND ITS INFLUENCE ON THE QUALITY OF THE WEBSITES OF RURAL TOURISM FACILITIES IN POLAND

ZJAWISKO LINK ROT I JEGO WPŁYW NA JAKOŚĆ WITRYN INTERNETOWYCH OBIEKTÓW TURYSTYKI WIEJSKIEJ W POLSCE

Karol Król^{1(B,F)}

¹University of Agriculture in Krakow
Uniwersytet Rolniczy im. Hugona Kołłątaja w Krakowie

Król, K. (2019). The link rot phenomenon and its influence on the quality of the websites of rural tourism facilities in Poland/ Zjawisko link rot i jego wpływ na jakość witryn internetowych obiektów turystyki wiejskiej w Polsce. *Economic and Regional Studies*, 12(1), p. 68-79. <https://doi.org/10.2478/ers-2019-0007>

ORIGINAL ARTICLE

JEL code: M15, M31

Submitted:
December 2018

Accepted:
January 2019

Tables: 4
Figures: 0
References: 32

ORYGINALNY ARTYKUŁ
NAUKOWY

Klasyfikacja JEL: M15, M31

Zgłoszony:
grudzień 2018

Zaakceptowany:
styczeń 2019

Tabele: 4
Rysunki: 0
Literatura: 32

Summary

Subject and purpose of work: The subject of the research were the profiles of links (internal and outbound ones) of websites of rural tourism facilities in Poland. The purpose of the study was to assess the scale of the link rot phenomenon and its impact on the quality of websites of rural tourism facilities in Poland.

Materials and methods: The assessment of the link profile was conducted using the Broken Link Checker application. The websites' quality assessment (the degree of their optimization for search engines) was conducted using three different applications which automate tests.

Results: A total of 919 websites were examined. Broken links were recorded on 464 sites, where approximately 65% of them had only a few broken links. Broken links that referred to the description of an offer or social media accounts were particularly unfavourable. No linear correlation was recorded between the number of broken links and the website's optimization for search engines.

Conclusions: The research demonstrated that the scale of the link rot phenomenon was relatively small in the group of the examined sites. It seems, however, that this was not the result of conscious design measures or optimization.

Keywords: broken links, dead links, link rot, rural tourism, quality, optimization

Streszczenie

Przedmiot i cel pracy: Przedmiotem badań były profile linków (wewnętrznych i wychodzących) witryn internetowych obiektów turystyki wiejskiej w Polsce. Celem pracy była ocena skali zjawiska link rot i jego wpływu na jakość witryn internetowych obiektów turystyki wiejskiej w Polsce.

Materiały i metody: Ocena profilu linków przeprowadzono za pomocą aplikacji Broken Link Checker. Ocena jakości witryn (stopnia optymalizacji dla wyszukiwarek) wykonano przy pomocy trzech różnych aplikacji automatyzujących testy.

Wyniki: Badaniom poddano łącznie 919 witryn internetowych. Niedziałające linki odnotowano na 464 witrynach, przy czym na około 65% z nich odnotowano jedynie kilka działających linków. Szczególnie niekorzystne były działające linki, które odsyłały do opisu oferty lub kont w mediach społecznościowych. Nie odnotowano zależności liniowej między liczbą działających linków a stopniem optymalizacji witryny dla wyszukiwarek.

Wnioski: Badania pokazały, że skala zjawiska link rot była w zbiorze badanych witryn relatywnie nieduża. Wydaje się jednak, że nie było to wynikiem świadomych zabiegów projektowych lub optymalizacji.

Słowa kluczowe: działające linki, martwe linki, rozpad linków, turystyka wiejska, jakość, optymalizacja

Address for correspondence/ Adres korespondencyjny: dr inż. Karol Król (ORCID 0000-0003-0534-8471), Uniwersytet Rolniczy im. Hugona Kołłątaja w Krakowie, Wydział Inżynierii Środowiska i Geodezji, al. Mickiewicza 24/28, 30-059 Kraków, Polska; tel. +48 12 662 40 16; e-mail: k.krol@onet.com.pl

Journal indexed in/ Czasopismo indeksowane w: AgEcon Search, AGRO, BazEkon, Index Copernicus Journal Master List, ICV 2017: 100,00; Polish Ministry of Science and Higher Education 2015-2018: 9 points/ AgEcon Search, AGRO, BazEkon, Index Copernicus Journal Master List ICV 2017: 100,00; Ministerstwo Nauki i Szkolnictwa Wyższego 2015-2018: 9 punktów. **Copyright:** © 2019 Pope John Paul II State School of Higher Education in Białą Podlaską, Karol Król. All articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) License (<http://creativecommons.org/licenses/by-nc-sa/4.0/>), allowing third parties to copy and redistribute the material in any medium or format and to remix, transform, and build upon the material, provided the original work is properly cited and states its license.

Introduction

Owners of rural tourism facilities should undertake such activities that will allow them to attract a satisfying number of customers. These activities may be essentially divided into internal ones - increasing the attractiveness of an offer, and external ones - consisting in particular of broadly understood promotion (Król, 2018a). Having created a tourist product and set the price, this information should be passed on to potential buyers (Krzyżanowska, 2014). The competitive advantage on the market of tourism services may be determined by using the opportunities offered by the Internet and global technological changes (Kachniewska, 2014a; Naramski et al., 2015; Król, 2016).

The global trend towards the development of information and communications technologies (ICT) has caused the technicalization and computerization of the tourism industry. This is not always understood by small entrepreneurs operating in rural areas and mostly using traditional business organization methods (Kachniewska, 2014b). This also applies to the distribution and promotion of a tourist offer, in particular, the more efficient use of online sales channels and social media. Despite the growing availability of the Internet, as well as computer techniques and tools, the awareness of the importance of Internet marketing and issues related to the optimization of websites among owners of rural tourism facilities is still too small.

The effectiveness of the promotion of rural tourism on the Internet is sometimes limited by, among others, the low level of marketing and IT competences of the owners of rural tourism facilities, their lack of trust in online tools and uncertainty about their own skills in this field (Tomczak-Woźniak, 2013). Nevertheless, many owners of rural tourism facilities, in order to make their offer stand out and to reach a wider audience with it, use the potential of websites, which largely depends on the broadly understood quality of the website (Król, 2018a).

Purpose and goals of the research

The attention of marketers and website owners is most often focused on the assessment of the quality of inbound links. The quality of inbound links and, more specifically – the quality of websites from which these links originate, is one of the attributes that testify to the usefulness (popularity) of a website (a given resource). Links from other websites constitute a form of recommendation (Król, 2017b). In addition, this attribute affects the placement of a site in search results. The quality of internal links and outbound links is examined less often, although their role is significant. Internal links determine the structure of a website. In turn, outbound links redirect users to other places on the web. In both of these cases, broken links prevent access to a given resource and reduce the quality of a site. A large accumulation of broken links is referred to as “link rot.” The aim of the work was to assess the scale of the link rot phenomenon and

Wstęp

Właściciele obiektów turystyki wiejskiej powinni podejmować takie działania, które pozwolą im pozyskać satysfakcjonującą liczbę klientów. Działania te można zasadniczo podzielić na wewnętrzne – zwiększające atrakcyjność oferty, i zewnętrzne – polegające w szczególności na szeroko pojętej promocji (Król, 2018a). Po stworzeniu produktu turystycznego i ustaleniu ceny, informacja o tym powinna być przekazana potencjalnym nabywcom (Krzyżanowska, 2014). O przewadze konkurencyjnej na rynku usług turystycznych może zadecydować wykorzystanie możliwości, jakie dają Internet oraz globalne zmiany technologiczne (Kachniewska, 2014a; Naramski i in., 2015; Król, 2016).

Światowy trend w kierunku rozwoju technologii informacyjno-komunikacyjnych (ICT) wywołał technicyzację i informatyzację branży turystycznej. Nie zawsze spotyka się to ze zrozumieniem małych przedsiębiorców działających na obszarach wiejskich i bazujących na tradycyjnych sposobach organizacji biznesu (Kachniewska, 2014b). Dotyczy to również dystrybucji i promocji oferty turystycznej, w szczególności lepszego wykorzystania internetowych kanałów sprzedaży i mediów społecznościowych. Pomimo rosnącej dostępności Internetu oraz technik i narzędzi komputerowych wciąż zbyt mała jest świadomość znaczenia marketingu internetowego i zagadnień związanych z optymalizacją witryn internetowych wśród właścicieli obiektów turystyki wiejskiej.

Efektywność promocji turystyki wiejskiej w Internecie bywa ograniczana m.in. przez niski poziom kompetencji marketingowych i informatycznych właścicieli obiektów turystyki wiejskiej, brak zaufania do narzędzi internetowych i niepewność, co do własnych umiejętności z tego zakresu (Tomczak-Woźniak, 2013). Mimo to wielu właścicieli obiektów turystyki wiejskiej, aby wyróżnić swoją ofertę i dotrzeć z nią do szerszego grona odbiorców wykorzystuje potencjał witryn internetowych, który w dużej mierze zależy od szeroko rozumianej jakości witryny (Król, 2018a).

Przedmiot i cel badań

Uwaga marketerów i właścicieli witryn internetowych najczęściej skoncentrowana jest na ocenie jakości linków przychodzących. Jakość linków przychodzących, a uściślając – jakość witryn internetowych, z których te linki pochodzą, jest jednym z atrybutów świadczących o przydatności (popularności) witryny (danego zasobu). Linki pochodzące z innych witryn stanowią formę rekomendacji (Król, 2017b). Ponadto atrybut ten wpływa na plasowanie witryny w wynikach wyszukiwania. Rzadziej badana jest jakość linków wewnętrznych i linków wychodzących, choć ich rola jest znacząca. Linki wewnętrzne ustanawiają strukturę serwisu. Z kolei linki wychodzące odsyłają użytkowników w inne miejsca sieci. W jednym i drugim przypadku niedziałające linki uniemożliwiają dostęp do określonego zasobu i obniżają jakość witryny. Duże nagromadzenie zepsutych linków określane jest terminem link rot. Celem pracy jest ocena skali zjawie-

its impact on the quality of websites of rural tourism facilities in Poland. A research hypothesis that the sites of rural tourism facilities have a large number of broken links has been put forth. The hypothesis was based on a survey of studies which demonstrated that the websites of rural tourism facilities are rarely updated and often need optimization for search engines (Król, 2018a). At the same time, the following research questions were posed: (1) what is the scale of the link rot phenomenon on the websites of rural tourism facilities in Poland? (2) Is there a relationship between the scale of the link rot phenomenon and the degree of a website's search engine optimization (SEO)?

The quality of websites

The term "quality" is derived from ancient times, from the Greek word *poiotes* (Latin *qualitas*). According to Plato, "the quality of things is the degree of perfection achieved by them" (Bielawa, 2011). Quality is a polysemous concept. From the point of view of a consumer, it means the meeting of needs that may be divided into functional and non-functional ones. In technical terms, quality is determined without taking the recipient into account. The benchmark for it may include a norm, a standard, a project (Sikorski, 2012). Extending the strictly technical approach to also include the buyer allows to define quality in the technical and economic sense as "a preference for specific properties that should be given to products so that they meet the users' expectations" (Prussak, 2006, p. 15).

The quality of a website is influenced by many factors. These factors may be divided into two main groups: (1) *on-site* ones, referring to the commercial quality and the content of a site (quality and usefulness of the content, the way it is presented), and (2) *off-site* ones, referring to the site's surroundings, e.g. the authority of a domain (Król, 2018a). All this comprises the so-called *search engine optimization* (SEO) of a website. SEO is the extent to which a website is adapted to the changing design standards and to the requirements set for websites by both users (*user experience*, UX) and the algorithms of search engines (*machine experience*). The quality of websites is also influenced by links, especially inbound and internal ones.

Types of links and "link authority"

Each online resource has its own unique URL (*Uniform Resource Locator*) address, consisting of the protocol and the host (the server's address), while more complex URLs may also include, among others, the port and the access path to the resource. Websites (network resources) are connected with each other by hyperlinks that are located in the structure of hypertext documents (Benbow, 1998; Taylor and Hudson, 2000). Hyperlinks, also referred to as links, may be divided into: (1) *inbound links*, those originating from other websites (*incoming links*), otherwise known as *backlinks*. These are

ska link rot i jego wpływu na jakość witryn internetowych obiektów turystyki wiejskiej w Polsce. Postawiono hipotezę badawczą, że witryny obiektów turystyki wiejskiej posiadają dużą liczbę niedziałających linków. Hipotezę postawiono na podstawie przeglądu badań, które wykazały, że witryny obiektów turystyki wiejskiej są rzadko aktualizowane i często wymagają optymalizacji dla wyszukiwarek internetowych (Król, 2018a). Jednocześnie postawiono następujące pytania badawcze: (1) jaka jest skala zjawiska link rot na stronach internetowych obiektów turystyki wiejskiej w Polsce? (2) czy istnieje związek między skalą zjawiska link rot a stopniem optymalizacji witryny dla wyszukiwarek internetowych (SEO)?

Jakość witryn internetowych

Pojęcie „jakości” wywodzi się z czasów starożytnych, z greckiego *poiotes* (łac. *qualitas*). Według Platona „jakość rzeczy to stopień osiągniętych przez nie doskonałości” (Bielawa, 2011). Jakość jest pojęciem wieloznacznym. Z punktu widzenia konsumenta to spełnienie potrzeb, które można podzielić na funkcjonalne i niefunkcjonalne. W ujęciu technicznym jakość jest określana z pominięciem odbiorcy. Wzorcem jest norma, standard, projekt (Sikorski, 2012). Rozszerzenie podejścia ściśle technicznego o nabywcę pozwala na określenie jakości w rozumieniu techniczno-ekonomicznym jako „preferowanie określonych właściwości, które należy nadać produktom, aby spełniały oczekiwania użytkowników” (Prussak, 2006, s. 15).

Na jakość witryny internetowej wpływa wiele czynników. Czynniki te można podzielić na dwie zasadnicze grupy: (1) wewnętrzne (*on-site*), odnoszące się do jakości użytkowej oraz do zawartości witryny (jakość i przydatność treści, sposób ich prezentacji), a także (2) zewnętrzne (*off-site*), odnoszące się do otoczenia witryny, np. autorytet domeny (Król, 2018a). Wszystko to składa się na tzw. optymalizację witryny dla wyszukiwarek internetowych (*search engine optimization*, SEO). SEO to stopień w jakim witryna jest dostosowana do zmieniających się standardów projektowych oraz do wymogów, jakie stawiają w trytonom zarówno użytkownicy (*user experience*, UX), jaki i algorytmy wyszukiwarek (*machine experience*). Na jakość witryn internetowych wpływają także linki, w szczególności przychodzące i wewnętrzne.

Rodzaje i „moc linków”

Każdy zasób internetowy ma swój niepowtarzalny adres URL (*Uniform Resource Locator*) złożony z protokołu i hosta (adresu serwera), przy czym bardziej rozbudowane adresy URL mogą zawierać także m.in. port i ścieżkę dostępu do zasobu. Witryny internetowe (zasoby sieciowe) są ze sobą połączone hiperłączami, które są umiejscawiane w strukturze dokumentów hipertekstowych (Benbow, 1998; Taylor i Hudson, 2000). Hiperłącza (hyperlinks), zwane też linkami, odnośnikami lub odsyłaczami, można podzielić na: (1) linki przychodzące (*inbound links*), pochodzące z innych witryn (*incoming link*), inaczej zwane linka-

links recommending a given website (*the referent*), originating from the website that refers to them (*the referrer*). *Inbound links* are included in the group of *external links*. Recommended sites assume the so-called "authority of the referrer," also known in SEO jargon as "*Link Juice*"¹. Inbound links constitute a kind of "*vote of confidence*" given by one site to another. Building a portfolio of incoming links constitutes the foundation of SEO activities; (2) *outbound links*, redirecting to other sites, and (3) *internal links*, connecting individual pages of a given site.

There are also textual and graphical links, as well as "natural" and advertising (paid) links. Links may be "good", i.e. those that constitute a form of recommendation and may have a positive impact on a site's visibility in search results or "bad" (the so-called penalty risking links), which may have a negative impact (Król, 2017b). Links may also be divided in terms of the manner in which the resource is opened, e.g. in a new or the same browser window (tab). In addition, *site-wide* links are distinguished. They are usually an integral part of a website (the website's template), thus they are located on every page and always assume an identical form (they have the same anchoring text and the same surroundings). These links are most often placed in the footer of a website.

Link rot phenomenon

Internet resources are subject to constant changes. Sites are modernized. New resources are published, others are deleted, files and directories are renamed. The ephemeral nature of the web causes the connections between sites to "be broken." This necessitates the continuous monitoring of the link profile (Goh and Ng, 2007).

"*Link building*" is of great importance in internet marketing. It should be preceded by the verification of *backlinks* in terms of their number and origin (potential value), attributes, as well as effectiveness (in generating traffic). At the same time, the optimization of internal links is often neglected. Internal linking shapes the structure of connections between the pages of a given site. Analyses of outbound links are overlooked just as often. *Broken or dead* outbound links lead to nowhere. Mary P. Benbow (1998) described the phenomenon of clicking on a broken link as "transferring the user to the cyber-no-man's-land." Travelling to nowhere is usually associated with

mi zwrotnymi (*backlinks*). Są to linki polecające daną witrynę (*the referent*), pochodzące z witryn odsyłających (*the referrer*). *Inbound links* zaliczają się do linków zewnętrznych (*external links*). Witryny polecane przyjmują tzw. „moc witryny odsyłającej”, zwaną też w żargonie SEO – „sokiem z linków” (*Link Juice*¹). Linki przychodzące stanowią pewnego rodzaju „wotum zaufania” od jednej strony dla innej (*vote of confidence*). Budowanie portfela linków przychodzących stanowi podstawę działań SEO; (2) linki wychodzące (*outbound links*), odsyłające do innych witryn oraz (3) linki wewnętrzne (*internal links*), łączące poszczególne strony danej witryny.

Wyróżnia się także linki tekstowe i graficzne oraz „naturalne” i reklamowe (kupione). Linki mogą być „dobre”, tj. takie, które stanowią formę rekomendacji i mogą mieć wpływ na wzrost widoczności witryny w wynikach wyszukiwania oraz „złe” (tzw. linki ryzykujące karę), które mogą oddziaływać negatywnie (Król, 2017b). Linki można podzielić także z uwagi na sposób w jaki jest przywoływany zasób np. w nowym lub w tym samym oknie (zakładce) przeglądarki. Ponadto wyodrębnia się linki typu *site-wide*. Są one przeważnie integralną częścią serwisu internetowego (szablonu witryny), przez co znajdują się na każdej stronie i przyjmują zawsze identyczną postać (mają jednakowy tekst zakotwiczenia i jednakowe otoczenie). Linki te najczęściej zamieszczane są w stopce serwisu.

Zjawisko link rot

Zasoby Internetu ulegają ciągłym zmianom. Witryny są modernizowane. Nowe zasoby są publikowane, inne usuwane, zmieniane są nazwy plików, katalogów. Efemeryczność sieci sprawia, że połączenia między witrynami „bywają zrywane”. Wymaga to ciągłego monitorowania profilu linków (Goh i Ng, 2007).

Budowa „zaplecza linków” (*link building*) ma istotne znaczenie w marketingu internetowym. Powinna być poprzedzona weryfikacją linków przychodzących (*backlinks*) pod kątem ich liczby i pochodzenia (potencjalnej wartości), atrybutów, a także skuteczności (w generowaniu ruchu). Jednocześnie często pomijana jest optymalizacja linków wewnętrznych. Linkowanie wewnętrzne kształtuje strukturę powiązań między stronami danej witryny. Równie często pomijane są analizy linków wychodzących. Linki wychodzące, które są „uszkodzone” (*broken links, dead links*) prowadzą donikąd. Mary P. Benbow (1998) opisała zjawisko kliknięcia w niedziałający link, jako „przeniesienie użytkownika na ziemię cyber-niczczą” (cyber-no-man's-land). Podróży donikąd towarzyszy

¹ Link Juice is an alternative (informal, non-technical) term for "*link equity*", which actually expresses the equity of the referring website. Greg Boser, an *SEO specialist* is thought to be the creator of the term Link Juice. If a site enjoys considerable authority (has a good reputation), i.e. it is characterized by a large number of visits and inbound links, it has more Link Juice to transfer. Link Juice defines the "SEO authority" (the value of SEO, the value of inbound links), which is transferred by the referrer in the form of (back)links to the target website. This term is based on the assumption that some links transfer value and authority from one site to the other. A large number of inbound links coming from high-quality websites may translate into the better placement of a site in search results.

¹ Link Juice to alternatywne (potoczne, nietechniczne) określenie dla „kapitału linku” (*link equity*), który w rzeczywistości określa kapitał witryny odsyłającej. Za twórcę terminu Link Juice uznawany jest Greg Boser (*SEO specialist*). Jeśli witryna cieszy się dużym autorytetem (ma dobrą reputację), tj. posiada m.in. dużą liczbę odwiedzin i linków przychodzących, to ma więcej Link Juice do przekazania. Link Juice określa „moc SEO” (wartość SEO, wartość linków przychodzących), którą przekazuje witryna odsyłająca w postaci (back)linków do strony docelowej. Określenie to opiera się na założeniu, że niektóre linki przekazują wartość i autorytet z jednej strony na drugą. Duża liczba linków przychodzących z witryn wysokiej jakości może przełożyć się na lepsze uplasowanie witryny w wynikach wyszukiwania.

a disappointment. This has a negative impact on user experience (UX).

It's only natural that by way of the evolution of the web, healthy URLs stop working a few years after they have appeared on the Internet. Link rot is a term used to refer to websites that contain many broken ("dead") links (Rhodes, 2010). *Link decay* progresses over time. It is associated with a growing number of broken links that refer to internal resources, but also to external resources, which, for example, ceased to exist or no longer exist at the given address (Król, 2018b). The link rot phenomenon has been known for a long time. The fact of it being a particular nuisance was noticed in the context of references to bibliographic sources, cited in academic works (Fichter, 1999; Taylor and Hudson, 2000).

The by-products of a decaying URL assume various forms. Two of them are the most typical and, at the same time, the most undesirable (Pinboard Blog, 2011): (1) the so-called *truly dead links*, also referred to as "blind links." They are relatively easy to observe and notice. Truly dead links cause one of the error messages (*an HTTP error code*) to appear, (2) the resource or page does not exist at the given location (or at the given address), but the website still exists (*page gone, site lives on*). This is most often reported by the 404 error message (The 404 "Not Found" response).

Lawrence (2001) pointed out two main reasons for the "decaying" of URLs: (1) personal websites (the so-called "home pages") or websites dedicated to singular, occasional events are sometimes abandoned by their owners who have not paid the subscription fee, (2) pages may be updated without maintaining the existing link structure (the resources might change their location in the directory tree). Rumsey (2002) showed that link damage occurs most often due to: (1) a change in the domain name, (2) a change in the directory, (3) a network failure, and (4) changes to the conditions of resource sharing.

Koehler (2004) listed the main messages associated with reporting broken or unavailable URLs: (1) reorganization or removal of files from the data server reported by the 404 HTTP code (Not Found) – page not found: a page or a resource does not exist on the server, (2) 400 Bad Request: the request cannot be handled by the server due to an irregularity perceived as the user's error (e.g. incorrect query syntax), (3) lack of access to the resource reported by the 401 HTTP error message that results from the limited access to the resource e.g. the resource exists in a password-protected directory (unauthorized access – requesting a resource that requires authentication), (3) internal server errors reported by the *504 Gateway Timeout* or *502 Bad Gateway* error messages.

zwykle rozczarowanie. Wpływa to negatywnie na wrażenia użytkownika (UX).

To naturalne, że w drodze ewolucji sieci zdrowe adresy URL przestają działać kilka lat po pojawieniu się w Internecie. Link rot, w wolnym tłumaczeniu – „szrotowisko linków” lub „zgnilizna linków”, to termin używany w odniesieniu do witryn internetowych, które zawierają wiele uszkodzonych („martwych”) linków (Rhodes, 2010). Rozpad linków (*link decay*) postępuje w czasie. Jest on związany z powiększającą się liczbą linków uszkodzonych, odsyłających do zasobów wewnętrznych, ale także do zasobów zewnętrznych, które np. przestały istnieć lub nie istnieją już pod danym adresem (Król, 2018b). Zjawisko link rot jest znane od dawna. Jego szczególną uciążliwość dostrzeżono na polu odwołań do źródeł bibliograficznych, przytaczanych w opracowaniach naukowych (Fichter, 1999; Taylor i Hudson, 2000).

Produkty rozpadu adresu URL przyjmują różną postać. Dwa z nich są najbardziej typowe i jednocześnie najbardziej niepożądane (Pinboard Blog, 2011): (1) tzw. prawdziwie niedziałające linki (*truly dead links*), określane też „martwymi linkami” lub „ślepyimi linkami”. Są one stosunkowo łatwe do zaobserwowania i odnotowania. Truly dead links powodują wywołanie jednego z komunikatów o błędzie (*an HTTP error code*), (2) dany zasób lub strona nie istnieje w danej lokalizacji (lub pod danym adresem), lecz witryna istnieje nadal (*page gone, site lives on*). Jest to najczęściej raportowane komunikatem o błędzie 404 (The 404 „Not Found” response).

Lawrence (2001) zwrócił uwagę na dwie główne przyczyny „rozpadu” adresów URL: (1) osobiste strony internetowe (tzw. strony „domowe”) lub strony internetowe pojedynczych, okazjonalnych wydarzeń bywają porzucane przez swoich właścicieli, którzy nie dokonali opłaty abonamentowej, (2) strony mogą zostać zmodernizowane bez zachowania dotychczasowej struktury linków (zasoby mogą zmienić lokalizację w strukturze katalogów). Rumsey (2002) wykazał, że uszkodzenie linku występuje najczęściej z powodu: (1) zmiany nazwy domeny, (2) zmiany struktury katalogów, (3) awarii sieci oraz (4) zmiany warunków udostępniania zasobu.

Koehler (2004) wymienił główne komunikaty związane z raportowaniem niedziałających lub niedostępnych adresów URL: (1) reorganizacja lub usunięcie plików z serwera danych raportowane kodem HTTP 404 (Not Found) – nie znaleziono strony: strona lub zasób nie istnieje na serwerze, (2) 400 Bad Request: żądanie nie może być obsłużone przez serwer z powodu nieprawidłowości postrzeganej jako błąd użytkownika (np. błędna składnia zapytania), (3) brak dostępu do zasobu raportowany komunikatem HTTP 401, który wynika z ograniczonego dostępu do zasobu np. zasób występuje w katalogu chronionym hasłem (nieautoryzowany dostęp – żądanie zasobu, który wymaga uwierzytelnienia), (3) wewnętrzne błędy serwera raportowane komunikatami – przekroczono limit czasu bramy HTTP (*504 Gateway Timeout*) lub błąd bramy HTTP (*502 Bad Gateway*).

Materials and methods

The research covered the websites of rural tourism facilities in Poland. The website addresses were obtained from the directory provided by the 'agroturystyka.pl' portal, the website of the Polish Federation of Rural Tourism "Gospodarstwa Gościnne" and through the Google search engine, from the so-called *organic search results*. Google's search engine was chosen because of its popularity around the world (Dergiades et al., 2018). The analyses involved websites in paid domains, including the national domain (the .pl, *country code top-level domain*, ccTLD), European, commercial, regional, and other subscription-based ones (see Król, 2017a).

In the examinations of the quality of internal and outgoing links (the scale of the link rot phenomenon), the Online Broken Link Checker application was used, which is available free of charge (brokenlinkcheck.com). The tests were carried out in the full mode, which made it possible to obtain information about the occurrence of each dead link (*Report all occurrences of each dead link*). Textual links leading to websites (hypertext documents) were analysed because the free version of the Broken Link Checker application does not allow for the verification of links referring to documents other than hypertext ones, e.g. pdf, doc, xls, ppt. In addition, the "damage" to graphical links (graphical files) was not verified.

In the next stage of the research, the websites were subjected to an SEO audit, which was carried out using three independent testing tools (tab. 1). Consequently, each site was described with the help of three SEO indicators expressed by a score on scales with different ranges. The final assessment of the degree of optimization of each site was obtained by applying the unitarization method (normalization of variables). It made it possible to describe each site using a synthetic, "accumulated" note – F-Score.

The normalization of variables was performed with the help of zero unitarization. With its help, the numerical ranges of variables have been unified and their absolute values have been transformed into relative ones (Kukuła and Bogocz, 2014). The degree of a website's optimization for search engines

Materiały i metody

Badaniami objęto witryny internetowe obiektów turystyki wiejskiej w Polsce. Adresy witryn pozyskano z katalogu udostępnionego przez portal 'agroturystyka.pl', serwis Polskiej Federacji Turystyki Wiejskiej „Gospodarstwa Gościnne” oraz z wyszukiwarki internetowej Google, z tzw. wyników organicznych (*organic search*). Wyszukiwarkę Google wybrano z uwagi na jej popularność w świecie (Dergiades i in., 2018). Badaniom poddano witryny w domenach płatnych, w tym w domenie krajowej (domena .pl, *country code top-level domain*, ccTLD), europejskiej, komercyjnych, regionalnych i innych abonamentowych (zob. Król, 2017a).

W badaniach jakości linków wewnętrznych i wychodzących (skali zjawiska link rot) wykorzystano aplikację Online Broken Link Checker, która jest udostępniana bezpłatnie (brokenlinkcheck.com). Testy wykonywano w trybie pełnym, który pozwala pozyskać informację o wystąpieniu każdego martwego łącza (*Report all occurrences of each dead link*). Analizie poddano linki tekstowe prowadzące do witryn internetowych (dokumentów hipertekstowych) ponieważ ogólnodostępna wersja aplikacji Broken Link Checker nie umożliwia weryfikacji linków odsyłających do dokumentów innych niż hipertekstowe np. pdf, doc, xls, ppt. Ponadto nie weryfikowano „uszkodzeń” linków przyjmujących postać obrazów (plików graficznych).

W kolejnym etapie badań witryny poddano audytowi SEO, który wykonano przy pomocy trzech niezależnych narzędzi testujących (tab. 1). W konsekwencji każda witryna została opisana przy pomocy trzech wskaźników SEO wyrażonych notą plasowaną w skali o różnej rozpiętości. Ostateczną ocenę stopnia optymalizacji każdej z witryn uzyskano stosując metodę unitaryzacyjną (normalizację zmiennych). Pozwoliła ona opisać każdą z witryn syntetyczną notą „skumulowaną” – F-Score.

Normalizację zmiennych wykonano przy pomocy unitaryzacji zerowanej. Przy jej pomocy ujednociono zakresy liczbowe zmiennych oraz przekształcono ich wartości bezwzględne na wartości względne (Kukuła i Bogocz, 2014). Stopień optymalizacji witryny dla

Table 1. Tools used in the examinations of websites

Tabela 1. Narzędzia wykorzystane w badaniach witryn

Tool/ Narzędzie	Tool type and object of measurement/ Typ narzędzia i przedmiot pomiaru	Scale of measurement/ Skala pomiaru
Pozycjonowanie.pl (online)	Measuring the degree of SEO optimisation for internet search engines. SEO indicator: SEO Score./ Pomiar stopnia optymalizacji witryny dla wyszukiwarek internetowych SEO. Wskaźnik SEO: SEO Score.	1-5
ZadroWeb (online)	Measuring the degree of SEO optimization for internet search engines./ Pomiar stopnia optymalizacji witryny dla wyszukiwarek internetowych SEO.	0-100
Sunspot SEO Audit (online)	Measuring the degree of SEO optimization for internet search engines. SEO indicator: general SEO score./ Pomiar stopnia optymalizacji witryny dla wyszukiwarek internetowych SEO. Wskaźnik SEO: ogólna ocena SEO.	0-170

Source: author's own study.
Źródło: opracowanie własne.

is an exogenous variable – it is desirable (expected) for its value to be as high as possible. Thus, in the unitarization process, the following formula (1) was used:

$$x'_{ij} = \frac{x_{ij} - \min_i \{x_{ij}\}}{\max_i \{x_{ij}\} - \min_i \{x_{ij}\}} \quad (1)$$

After normalization, the values of variables took a value from a closed interval with a range from 0 to 1. The F-Score indicator was the result of adding up the normalized values from individual measurements. It could, therefore, take a value ranging between 0 and 3 points.

Results and conclusions

A total of 919 websites have been examined. Broken links were recorded on 464 sites (about 50%), where about 65% of them only had a few broken links and, in many cases, it was a single *site-wide* link (copied onto each of the site's pages). Several dozen broken links were recorded on a hundred websites (tab. 2). These sites accounted for approximately 21.5% of all the sites with dead links on them. At the same time, websites on which several hundred or even several thousand broken links were identified had been recorded, and the highest value amounted to 3972 dead links on a single website.

wyszukiwarek jest stymulantą – pożądana (oczekiwana) jest jego jak największa wartość. Zatem w procesie unitaryzacji zastosowano formułę (1):

Wartości zmiennych po znormalizowaniu przyjmowały wartość z przedziału domkniętego od 0 do 1. Wskaźnik F-Score stanowił wynik zsumowania znormalizowanych wartości poszczególnych pomiarów. Mógł zatem przyjąć wartość od 0 do 3 punktów.

Wyniki i wnioski

Badaniom poddano łącznie 919 witryn internetowych. Niedziałające linki odnotowano na 464 witrynach (około 50%), przy czym na około 65% z nich odnotowano jedynie kilka nie działających linków, a w wielu przypadkach był to jeden link typu *site-wide* (powielany na każdej stronie serwisu). Na stu witrynach odnotowano kilkadziesiąt nie działających linków (tab. 2). Witryny te stanowiły około 21,5% witryn, na których zamieszczono martwe linki. Jednocześnie odnotowano witryny, na których zidentyfikowano kilkaset, a nawet kilka tysięcy nie działających linków, przy czym maksymalnie na jednej witrynie odnotowano 3972 martwych linków.

Table 2. Number of websites ordered by the number of broken links

Tabela 2. Liczba witryn internetowych według liczby nie działających linków

Number of broken links (from - to)/ Liczba uszkodzonych linków (od - do)	0	1-9	10-50	51-100	101-1000	>1000	In total/ Łącznie
Number of websites/ Liczba witryn	455	300	100	33	28	3	919
Percentage (%)/ Ujęcie procentowe (%)	49.51	32.64	10.88	3.59	3.05	0.33	100

Source: author's own study.

Źródło: badania własne.

In total, approximately 18 400 broken links have been recorded in the set of the analysed websites. About 60.4% of broken links referred users to other websites (outbound links). However, these were links that "lead to nowhere," reported by one of the HTTP error messages. The remaining links (around 39.6%) were internal links that led to non-existent pages of a given site (the so-called sub-pages). This was particularly important when the link referred to an offer's description. An example of such a link is given in Table 3. From the point of view of a website's search engine optimization, the location of the link alone was correct (a link within an article, surrounded by the text). However, it was a link that led to a non-existent page.

Łącznie w zbiorze badanych witryn odnotowano około 18,4 tys. nie działających linków. Około 60,4% nie działających linków odsyłało użytkowników do innych witryn internetowych (linki wychodzące). Były to jednak odesłania „prowadzące do nikąd”, raportowane jednym z komunikatów o błędzie HTTP. Pozostałe linki (około 39,6%) były linkami wewnętrznymi, które prowadziły do nieistniejących stron internetowych danej witryny (tzw. podstron). Miało to szczególne znaczenie, gdy link odsyłał do opisu oferty. Przykład takiego odesłania zamieszczono w tabeli 3. Z punktu widzenia optymalizacji witryny dla wyszukiwarek internetowych sama lokalizacji linku była właściwa (link w artykule otoczony tekstem). Jednak był to link, który prowadził do nieistniejącej strony.

Table 3. An example of a broken link referring to the description of a portfolio of services**Tabela 3.** Przykład niedziałającego linku odsyłającego do opisu oferty usługowej

The website's Internet address/ Adres internetowy strony http://www.ostoja-wilkow.pl/swieta-sylwester-gorach-swietokrzyskich/
Location of the broken link – the last paragraph of the text posted on the page – “ <i>In order to get acquainted with our comprehensive offer, we invite you to go to the offer tab and, should you have any questions, we would be very happy to provide you with information.</i> ” The link is the word “offer,” which refers to a non-existent page: http://www.ostoja-wilkow.pl/oferta/ Lokalizacja niedziałającego linku – ostatni akapit tekstu zamieszczonego na stronie – „ <i>W celu zapoznania się z naszą kompleksową ofertą zapraszamy do zakładki oferta, a w razie pytań bardzo chętnie służymy Państwu informacją</i> ”. Linkiem jest słowo „ oferta ”, które odsyła do nieistniejącej strony: http://www.ostoja-wilkow.pl/oferta/
Fragment of the HTML code (original record) <p>In order to get acquainted with our comprehensive offer, we invite you to go to the offer tab and, should you have any questions, we would be very happy to provide you with information.</p> Fragment kodu HTML (zapis oryginalny)<p>W celu zapoznania się z naszą kompleksową ofertą zapraszamy do zakładki oferta, a w razie pytań bardzo chętnie służymy Państwu informacją.</p>

Source: author's own study.

Źródło: opracowanie własne.

The qualitative analysis of broken outbound links showed that they often led to non-existent social media profiles. 18 dead links to Facebook and 39 dead links to Google+ have been recorded. In addition, references to non-existent foreign language versions of the site (on 8 websites) and non-existent sites promoting interactive agencies that created the site (recorded on 17 sites) have been recorded. In a large number of cases, these were broken *site-wide* links, placed in the footer of the site and copied onto each page. A large number of links to various non-existent websites, including weather forecast websites, as well as non-existent websites functioning as part of a larger website, e.g. Wikipedia, have also been recorded. Linking to external sites managed by a third party is therefore unfavourable.

Broken links accompanied by the “error 404” message were recorded most often (approximately 78.4%). In addition, in a large number of cases, the “bad host” message was noted (a server with this name does not exist or is unreachable). Broken links that led to websites which were unavailable for other reasons have also been recorded (tab. 4).

No linear relationship was noted between the number of broken links and the quality of the site expressed using the F-Score attribute ($r=0.034$, $n=919$, $\alpha=0.05$, $r^2=0.001$, $p>\alpha$). The number of broken links was not an attribute determining the degree of a website's search engine optimization. This is in line with the properties of SEO assessment because it consists of many attributes that are testimony

Analiza jakościowa niedziałających linków wychodzących pokazała, że często prowadziły one do nieistniejących profili w mediach społecznościowych. Odnotowano 18 martwych linków do serwisu Facebook oraz 39 martwych linków do serwisu Google+. Ponadto odnotowywano odwołania do nieistniejących obcojęzycznych wersji witryny (na 8 witrynach) i nieistniejących witryn firmujących agencje interaktywne, które utworzyły witrynę (odnotowane na 17 witrynach). W dużej liczbie przypadków były to niedziałające linki typu *site-wide*, zamieszczone w stopce witryny i powielane na każdej stronie. Odnotowano także dużą liczbę linków do różnorodnych, nieistniejących witryn, w tym serwisów pogodowych, a także nieistniejących stron internetowych funkcjonujących jako fragment większego serwisu np. Wikipedii. Linkowanie do witryn zewnętrznych, którymi zarządza podmiot trzeci jest zatem niekorzystne.

Najczęściej odnotowywano niedziałające linki, którym towarzyszył komunikat „błąd 404” (około 78,4%). Ponadto w dużej liczbie przypadków odnotowano komunikat „bad host” (serwer o tej nazwie nie istnieje lub jest nieosiągalny). Odnotowano także niedziałające linki, które prowadziły do witryn internetowych niedostępnych z innych powodów (tab. 4).

Nie odnotowano zależności liniowej między liczbą niedziałających linków a jakością witryny wyrażoną atrybutem F-Score ($r=0,034$, $n=919$, $\alpha=0,05$, $r^2=0,001$, $p>\alpha$). Liczba niedziałających linków nie była zatem atrybutem przesądzającym o ocenie stopnia optymalizacji witryny dla wyszukiwarek. Jest to zgodne

Table 4. Messages related to broken links recorded most often**Tabela 4.** Najczęściej odnotowywane komunikaty związane z niedziałającymi linkami

Error type (code)/ Rodzaj (kod) błędu	404	412	414	500	502	timeout	bad URL	bad host	other/ inne	In total/ Łącznie
Number of links/ Liczba linków	14414	6	6	408	121	199	66	3154	4	18378
Percentage (%)/ Ujęcie procentowe (%)	78.4	0.03	0.03	2.22	0.66	1.08	0.36	17.2	0.02	100

Source: author's own study.

Źródło: badania własne.

of a website's quality. A relatively strong linear relationship was noted between the size of the site expressed by the number of pages and the number of broken links ($r=0.61$, $n=919$, $\alpha=0.05$, $r^2=0.372$, $p<\alpha$). This is largely due to the frequent occurrence of *site-wide* links.

Discussion

The often dynamically progressing ephemerality and variability of resources make it so that the structure of the web changes. Internet addresses which, as recently as yesterday, led to useful resources might be *permanently unavailable* today. This is a natural process and it runs its course more so in an evolutionary model, rather than a revolutionary one. This is indicated by the test results. Hennessey et al. (2013) examined 14,489 different websites published in the years 1996-2010 in the Thomson Reuters Web of Science citation index and found that the average life cycle of these websites amounted to 9.3 years and only 62% of them have been archived. According to a study by Zittrain et al. (2014), over 70% of the URLs provided in articles published in the Harvard Law Review and other scientific journals, as well as approximately 50% of the URLs included in the rulings of the Supreme Court of the United States did not lead to (did not link to) originally cited information.

Spinellis (2003) attempted to determine how quickly archived information becomes obsolete. He showed that Internet resources were most often unavailable because they no longer existed on a given server (error code 404). In addition, numerous linking errors resulted from the hostname being incorrect (error code 901). Meanwhile, network/server problems (error code 504, Gateway Timeout) accounted for 8% of all errors. This was confirmed by research by Sellitto (2005). He demonstrated that missing resources were the main cause of broken links, reported by the code 404 (61.5%), as well as gateway errors (18%, code 502 – Bad Gateway) and gateway timeouts (17.6%, code 504).

Skowron (2018) conducted an analysis of the quality of outbound links posted on a thousand pages published on websites devoted to marketing. The research showed that among 1000 articles, authors of only 14 of them used links supplemented with a title (the title attribute). "Follow" links were used across most of the articles (70%). Links with the attribute "nofollow" were recorded in the case of only 6% of the articles, and in the case of 4%, simultaneous use of both types of links was observed. The articles analysed in the study most frequently contained several links (from 1 to 10) and as much as 20% of the articles did not contain any links.

Research on 201 websites of agritourism companies showed that their owners often had profiles on social media sites, but they did not demonstrate much activity in updating them, and they did not promote these profiles on their websites. The fact of possessing a social media account was tantamount to the presence of the so-called "social

z właściwością oceny SEO ponieważ składa się na nią wiele atrybutów świadczących o jakości witryny. Odnotowano relatywnie silną zależność liniową między wielkością witryny wyrażoną liczbą stron a liczbą niedziałających linków ($r=0,61$, $n=919$, $\alpha=0,05$, $r^2=0,372$, $p<\alpha$). W dużej mierze wynika to z częstego występowania linków typu *site-wide*.

Dyskusja

Efemeryczność i zmienność zasobów, często dynamicznie postępujące sprawiają, że zmienia się struktura sieci. Adresy internetowe, które jeszcze wczoraj prowadziły do przydatnych zasobów, dziś mogą być nieaktualne (*permanently unavailable*). Jest to proces naturalny i przebiega bardziej w modelu ewolucyjnym, niż rewolucyjnym. Wskazują na to wyniki badań. Hennessey i współautorzy (2013) przebadali 14 489 unikalnych stron internetowych zamieszczonych w latach 1996-2010 w indeksie cytowań Thomson Reuters Web of Science i ustalili, że średnia długości życia tych stron wyniosła 9.3 roku, a jedynie 62% z nich zostało zarchiwizowanych. Według badań Zittrain i współautorów (2014) ponad 70% adresów URL podanych w artykułach opublikowanych w Harvard Law Review i innych czasopismach naukowych oraz około 50% adresów URL zamieszczonych w opiniach Sądu Najwyższego Stanów Zjednoczonych Ameryki nie prowadziło (nie linkowało) do pierwotnie cytowanych informacji.

Spinellis (2003) podjął się próby określenia, jak szybko informacje archiwalne stają się nieaktualne. Wykazał, że zasoby internetowe były najczęściej niedostępne ponieważ nie istniały już na danym serwerze (kod błędu 404). Ponadto liczne błędy linkowania wynikały z nieprawidłowej nazwy hosta (kod błędu 901). Problemy sieciowe/serwerowe (kod błędu 504, przekroczony czas bramy, Gateway Timeout) stanowiły zaś 8% wszystkich błędów. Potwierdziły to badania Sellitto (2005). Wykazał on, że główną przyczyną niedziałających linków były brakujące zasoby, raportowane kodem 404 (61,5%), a także błędy bramy (18%, kod 502 – Bad Gateway) oraz przekroczenia czasu bramy (17,6%, kod 504).

Skowron (2018) przeprowadził analizę jakości linków wychodzących zamieszczonych na tysiącu stron internetowych opublikowanych w serwisach poświęconych marketingowi. Badania pokazały, że wśród 1000 artykułów, jedynie autorzy 14 z nich zastosowali linki uzupełnione tytułem (atrybut title). W większości artykułów stosowane były linki „follow” (70%). W przypadku jedynie 6% artykułów odnotowano linki opatrzone atrybutem „nofollow”, a w przypadku 4% odnotowano jednoczesne zastosowanie jednego i drugiego typu linków. Najczęściej badane artykuły zawierały kilka linków (od 1 do 10), przy czym aż 20% artykułów nie zawierało żadnego linku.

Badania 201 witryn internetowych gospodarstw agroturystycznych pokazały, że ich właściciele często posiadali profile w mediach społecznościowych, lecz nie wykazywali dużej aktywności w ich aktualizacji i nie promowali ich na swoich stronach inter-

media plugins” and links referring to social media profiles on a website (Król, 2018c). Therefore, studies conducted in such a manner were of both quantitative and qualitative nature. This is because they did not take into account the fact that many of these links might have been broken. As such, the presence of a link (a textual or a graphical one) to a social media site did not necessarily mean that there was indeed a profile on the social media site.

Dorocki et al. (2013) noted (n=64) in their studies that over 93% of the owners of rural tourism facilities had their own website (93.5%), but only 54% of them updated it “as needed.” At the same time, 24% of the respondents stated that the site was updated more often than once a year, and approximately 9% never updated the website. The owners of agritourism companies usually ran the website themselves (56.5%) or benefited from the help of a friend (19.6%). Simultaneously, approximately 22% of the respondents used the services of companies specializing in this field. The link rot phenomenon occurs more frequently on sites that regularly publish a lot of content, are not monitored or optimized (in the scope of SEO). The websites of rural tourism facilities, including agritourism companies, are usually not extensive, which limits the link rot phenomenon. At the same time, however, they are relatively seldom updated. They are optimized even more rarely. This can contribute to an increase in the number of broken links.

Practical implications

The link rot phenomenon is associated with unfavourable practical implications. Its severity may be a signal for search engine robots that a site is of inferior quality and should be placed in a worse position in the search results. Broken links to the details of an offer, a list of service fees or a form (a contact, booking, or payment one) may be particularly harmful. Links to foreign language versions of a website and social media sites which do not work are also unfavourable. Such links can reduce the user’s trust in a website, frustrate them, and cause them to stop browsing the site. This raises the question as to how to combat the phenomenon of link rot?

An obvious recommendation would be the regular audit and optimization of links. The link audit may be carried out with the help of one of the commonly available “broken link checker” internet applications. These applications scan websites in the search for broken links, find them and indicate their exact location, including in the HTML code. This makes it much easier to remove or repair broken links. In addition, taking preventive measures is recommended. Compliance with a few rules reduces the likelihood of the phenomenon of link rot spreading and may increase the long-term usability of the site (Leighton, 2015): (1) limiting the

netowych. Występowanie konta w mediach społecznościowych utożsamiane było z występowaniem na stronie internetowej tzw. „wtyczek społecznościowych” oraz linków odsyłających do portali społecznościowych (Król, 2018c). Tak przeprowadzone badania miały zatem charakter ilościowy i pogładowy. Nie uwzględnili bowiem faktu, że wiele z takich linków mogło być uszkodzonych. Obecność linku (tekstowego lub graficznego) do portalu społecznościowego nie musiał zatem oznaczać, że profil w portalu społecznościowym istniał.

Dorocki i współautorzy (2013) w swoich badaniach odnotowali (n=64), że ponad 93% właścicieli obiektów turystyki wiejskiej posiadało własną stronę internetową (93,5%), jednak już tylko 54% aktualizowało ją „w miarę potrzeb”. Jednocześnie 24% ankietowanych zadeklarowało, że aktualizacja witryny przeprowadzana jest częściej niż raz do roku, a około 9% nigdy nie aktualizowało witryny. Przeważnie właściciele gospodarstw samodzielnie administrowali stroną (56,5%) lub korzystali w tym zakresie z usługi grzeckościowej (19,6%). Jednocześnie około 22% badanych korzystało z usług firm specjalizujących się w tej dziedzinie. Zjawisko link rot występuje częściej na witrynach, które regularnie publikują dużo treści, nie są monitorowane (pod kątem SEO) i optymalizowane. Witryny obiektów turystyki wiejskiej, w tym gospodarstw agroturystycznych, przeważnie nie są rozbudowane, co ogranicza zjawisko link rot. Jednocześnie jednak są stosunkowo rzadko aktualizowane. Jeszcze rzadziej optymalizowane. Może to wpływać na wzrost liczby niedziałających linków.

Implikacje praktyczne

Zjawisko link rot ma niekorzystne implikacje praktyczne. Jego nasilenie może być sygnałem dla robotów wyszukiwarek, że witryna jest gorszej jakości i należy ją uplasować na gorszym miejscu w wynikach wyszukiwania. Szczególnie szkodliwe mogą być niedziałające linki do szczegółów oferty, cennika usług lub formularza (kontaktowego, rezerwacyjnego, płatniczego). Niekorzystne są także niedziałające linki do obcojęzycznych wersji witryny oraz serwisów społecznościowych. Takie linki mogą obniżyć zaufanie do serwisu, wzbudzić frustrację użytkownika i sprawić, że zaprzestanie on przeglądania witryny. Rodzi to zatem pytanie w jaki sposób przeciwstawić się zjawisku link rot?

Naturalnym zaleceniem jest regularny audyt i optymalizacja linków. Audyt linków można przeprowadzić przy pomocy jednej z ogólnodostępnych aplikacji internetowych typu „broken link checker”. Aplikacje te skanują witryny internetowe w poszukiwaniu uszkodzonych linków, wyszukują je i wskazują dokładne miejsce ich występowania, również w kodzie HTML. Znacząco ułatwia to usunięcie lub naprawę popsutych linków. Ponadto zalecana jest profilaktyka. Przestrzeganie kilku zasad zmniejsza prawdopodobieństwo rozszerzania się zjawiska link rot i może wpłynąć na wzrost długoterminowej użyteczności witryny (Leighton, 2015): (1) ograniczenie

number of outbound links to only include the most important ones. Linking to weather forecast websites or information from Wikipedia is not necessary, (2) avoiding linking to resources that may quickly disappear from the web, e.g. personal data, short-term content, pieces of news or press releases, (3) linking to original sources is advisable, (4) if possible, it is advisable to link to websites, not to files. Websites are usually more stable than files, e.g. pdf ones, (5) shortening links using external services, e.g. goo.gl, is not advisable (6) regular verification of the link profile is recommended.

Conclusions

The link rot phenomenon may negatively affect goal conversion and translate into the lower marketing (sales) efficiency of a site. Regardless of the scale of the phenomenon, link rot has a negative impact on the quality of a website. A site whose link profile has not been checked and updated for a long time may suffer from "link decay." Repairing or removing broken links may improve the ease of use and make the content of the site easier to discover by both users and search engines. Too many damaged links may indicate neglect and may translate into a website's score with regard to its quality being lower.

The research has shown that the scale of the link rot phenomenon is relatively small among the group of analysed sites. At this point, it should be mentioned that only references to hypertext documents were examined. Therefore, the actual number of broken links may be bigger. It seems, however, that the relatively small scale of the link rot phenomenon is not the result of conscious design measures or optimization. Perhaps it results from the fact that websites from the analysed group primarily play the role of internet business cards. These are not websites that are expanded, supplemented with new content. In most cases, they were not visibly updated. Singular links which referred to non-existent sites were recorded on many portals, however, these were often *site-wide* links. Most of the broken links were a consequence of the so-called phenomenon of ageing websites. At the same time, some cases of broken links referring to the description of an offer or a non-existent social media account have been recorded. Such links may be particularly unfavourable. They create negative impressions and make it so that the user has to become more involved in finding the content they are looking for. Removing broken links will improve the quality of a site and will have a positive impact on the algorithm-based assessment of the site carried out for the purposes of search engine rankings.

liczby linków wychodzących do najistotniejszych. Linkowanie do serwisów pogodowych lub informacji w Wikipedii nie jest konieczne, (2) unikanie linkowania do zasobów, które mogą szybko zniknąć z sieci np. treści osobistych, krótkoterminowych, aktualności lub informacji prasowych, (3) wskazane jest linkowanie do źródeł pierwotnych, (4) o ile to możliwe, wskazane jest linkowanie do stron internetowych, a nie do plików. Witryny są przeważnie bardziej stabilne niż pliki np. pdf, (5) nie jest wskazane skracanie linków, np. przy pomocy usług świadczonych przez podmioty zewnętrzne np. goo.gl (6) zalecana jest regularna weryfikacja profilu linków.

Podsumowanie

Zjawisko link rot może negatywnie wpływać na konwersję celu i przełożyć się na niższą skuteczność marketingową (sprzedażową) witryny. Bez względu na skalę zjawiska link rot wpływa ono negatywnie na jakość witryny internetowej. Witryna, której profil linków nie był sprawdzany i aktualizowany przez długi czas, może cierpieć na „zgniliznę linków”. Naprawa lub usunięcie niedziałających linków może poprawić wygodę użytkownika i sprawić, że zawartość witryny będzie łatwiejsza do odkrycia zarówno przez użytkowników, jak i wyszukiwarki internetowe. Zbyt duża liczba uszkodzonych łączy może świadczyć o zaniedbaniach i może przełożyć się na niższą ocenę jakości witryny.

Badania pokazały, że skala zjawiska link rot jest w zbiorze badanych witryn relatywnie nieduża. W tym miejscu należy wspomnieć, że badano jedynie odwołania do dokumentów hipertekstowych. W rzeczywistości liczba niedziałających linków może być zatem większa. Wydaje się jednak, że relatywnie nieduża skala zjawiska link rot nie wynika ze świadomych zabiegów projektowych lub optymalizacji. Być może wynika z faktu, że witryny z badanego zbioru to w większości internetowe wizytówki. Nie są to serwisy, które są rozbudowywane, uzupełniane o nowe treści. W większości przypadków nie były one w widoczny sposób aktualizowane. Na wielu witrynach odnotowano pojedyncze linki, które odsyłały do nieistniejących witryn, jednak często były to linki typu *site-wide*. Większość niedziałających linków była konsekwencją tzw. zjawiska starzenia się witryn internetowych. Jednocześnie odnotowano przypadki niedziałających linków, które odsyłały do opisu oferty lub nieistniejących kont w mediach społecznościowych. Takie linki mogą być szczególnie niekorzystne. Wywołują negatywne wrażenia i wymagają od użytkownika większego zaangażowania w wyszukanie treści. Usunięcie niedziałających linków podniesie jakość użytkową witryny, i będzie miało pozytywny wpływ na algorytmiczną ocenę witryny przeprowadzaną na potrzeby rankingów wyszukiwarek.

References/ Literatura:

1. Bielawa, A. (2011). Postrzeganie i rozumienie jakości – przegląd definicji jakości. *Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania*, 21, 144-152.
2. Dergiades, T., Mavragani, E., Pan, B. (2018). Google Trends and tourists' arrivals: Emerging biases and proposed corrections. *Tourism Management*, 66, 108-120. <https://doi.org/10.1016/j.tourman.2017.10.014>
3. Dorocki, S., Szymańska, A. I., Zdon-Korzeniowska, M. (2013). Przedsiębiorstwa agroturystyczne w gospodarce opartej na wiedzy. *Prace Komisji Geografii Przemysłu Polskiego Towarzystwa Geograficznego*, 24, 38-58.
4. Fichter, F. (1999). Do I look like a maid? Strategies for preventing link rot. *Online*, 23, 77-79.
5. Goh, D. H. L., Ng, P. K. (2007). Link decay in leading information science journals. *Journal of the American Society for Information Science and Technology*, 58(1), 15-24. <https://doi.org/10.1002/asi.20513>
6. Hennessey, J., Ge, S.X. (2013). A cross disciplinary study of link decay and the effectiveness of mitigation techniques. *BMC bioinformatics*, 14(14), 5. <https://doi.org/10.1186/1471-2105-14-S14-S5>
7. Kachniewska, M. (2014a). Wpływ digitalizacji kanałów dystrybucji na strukturę rynku usług pośrednictwa turystycznego. *E-mentor*, 1(53), 86-91.
8. Kachniewska, M. (2014b). Big Data analysis jako źródło przewagi konkurencyjnej przedsiębiorstw i regionów turystycznych. *Folia Turistica*, 32(32), 35-55.
9. Koehler, W. (2004). A longitudinal study of Web pages continued: a consideration of document persistence. *Information Research*, 9(2), 9-2.
10. Król, K. (2016). Globalne zmiany technologiczne i ich wpływ na promocję agroturystyki w internecie. *Roczniki Naukowe Ekonomii Rolnictwa i Rozwoju Obszarów Wiejskich*, 103(3), 84-100.
11. Król, K. (2017a). Internet domains promoting agrotourist farms. *Tourism*, 27(1), 23-29. <https://doi.org/10.1515/tour-2017-0003>
12. Król, K. (2017b). Wybrane formy internetowych rekomendacji w promowaniu agroturystyki. *Zagadnienia Doradztwa Rolniczego*, 4(90), 27-39.
13. Król, K. (2018a). Jakość witryn internetowych w zarządzaniu marketingowym na przykładzie obiektów turystyki wiejskiej w Polsce. Infrastruktura i Ekologia Terenów Wiejskich, III(2). Kraków: Komisja Technicznej Infrastruktury Wsi Oddziału Polskiej Akademii Nauk. <https://doi.org/10.14597/INFRAECO.2018.3.2.057>
14. Król, K. (2018b). Wstęp do encyklopedii linków. Audyt linków – wskazówki, ciekawostki, zalecenia. Narzędzia pomocne w charakterystyce witryny internetowej (część II). Kraków: HomeProject.pl. Pobrane z: <https://goo.gl/zzk3bF>
15. Król, K. (2018c). Synergia kanałów marketingowych w internetowej promocji agroturystyki. *Marketing i Rynek*, 3, 32-36.
16. Krzyżanowska, K. (2014). Skuteczność działań informacyjno-promocyjnych w turystyce wiejskiej. *Zeszyty Naukowe Uniwersytetu Szczecińskiego. Ekonomiczne Problemy Turystyki*, 3(27), 41-55.
17. Kukuła, K., Bogocz, D. (2014). Zero unitarization method and its application in ranking research in agriculture. *Economic and Regional Studies*, 7(3), 5-13.
18. Lawrence, S. (2001). Free online availability substantially increases a paper's impact. *Nature*, 411(6837), 521. <https://doi.org/10.1038/35079151>
19. Leighton, W.K. (2015). The growing problem of Internet „link rot” and best practices for media and online publishers. *Journalist's Resource*. Pobrane z: <https://goo.gl/n1T3aP>
20. Mary, P. Benbow, S. (1998). File not found: the problems of changing URLs for the World Wide Web. *Internet Research*, 8(3), 247-250. <https://doi.org/10.1108/10662249810217867>
21. Naramski, M., Herman, K., Szromek, A.R. (2015). Analiza porównawcza wybranych sposobów prezentacji lokalnej oferty turystycznej – studium przypadku. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 379, 278-287. <https://doi.org/10.15611/pn.2015.379.27>
22. Pinboard Blog (2011). Remembrance of Links Past. Pobrane z: <https://goo.gl/xbqbWB>
23. Prussak, W. (2006). *Zarządzanie jakością. Wybrane elementy*. Poznań: Wyd. Politechniki Poznańskiej.
24. Rhodes, S. (2010). Breaking down link rot: the Chesapeake project legal information archive's examination of URL stability. *Law Library Journal*, 102, 581-597.
25. Rumsey, M. (2002). Runaway train: Problems of permanence, accessibility, and stability in the use of Web sources in law review citations. *Law Library Journal*, 94(1), 27-39.
26. Sellitto, C. (2005). The impact of impermanent Web-located citations: A study of 123 scholarly conference publications. *Journal of the American Society for Information Science and Technology*, 56(7), 695-703. <https://doi.org/10.1002/asi.20159>
27. Sikorski, M. (2012). *Usługi on-line: jakość, interakcje, satysfakcja klienta*. Warszawa: Wyd. PJWSTK.
28. Skowron, R. (2018). Linki wychodzące – analiza na podstawie 1000 artykułów. Pobrano z: <https://goo.gl/GP4AzB>
29. Spinellis, D. (2003). The decay and failures of Web references. *Communications of the ACM*, 46(1), 71-77. <https://doi.org/10.1145/602421.602422>
30. Taylor, M. K., Hudson, D. (2000). Linkrot and the usefulness of Web site bibliographies. *Reference & User Services Quarterly*, 39(3), 273-277.
31. Tomczak-Woźniak, E. (2013). Polskie rolnictwo wobec wyzwań e-gospodarki. *Zeszyty Naukowe Uniwersytetu Szczecińskiego. Ekonomiczne Problemy Usług*, 105(2), 583-591.
32. Zittrain, J., Albert, K., Lessig, L. (2014). Perma: Scoping and addressing the problem of link and reference rot in legal citations. *Legal Information Management*, 14(2), 88-99. <https://doi.org/10.1017/S1472669614000255>