

The Problem of Reference Rot in Spatial Metadata Catalogues (Summary) ^{*}

Sergio Martin-Segura^{*}; Francisco Javier Lopez-Pellicer^{*}; Javier Noguera-Iso^{*};
Javier Lacasta^{*}; Francisco Javier Zarazaga-Soria^{*};

^{*}Department of Computer Science and System Engineering, University of Zaragoza, 50018
Zaragoza, Spain;

{segura,fjlopez,jnog,jlacasta,javy}@unizar.es

Palabras Clave: metadata, spatial data infrastructures, Reference Rot, Link Rot, Content Drift

Lugar de publicación: ISPRS International Journal of Geo-Information, Vol. 11, No. 1, p. 27, 2022

Impact factor: JCR 2.899 - Cuartil Q2 - Posición: 76 / 161 - Área: Computer Science/Information Systems

DOI: <https://doi.org/10.3390/ijgi11010027>

Resumen (Abstract). The content at the end of any hyperlink is subject to two phenomena: the link may break (Link Rot) or the content at the end of the link may no longer be the same as it was when it was created (Content Drift). Reference Rot denotes the combination of both effects. Spatial metadata records rely on hyperlinks for indicating the location of the resources they describe. Therefore, they are also subject to Reference Rot. This paper evaluates the presence of Reference Rot and its impact on the 22,738 distribution URIs of 18,054 metadata records from 26 European INSPIRE spatial data catalogues. Our Link Rot checking method detects broken links while considering the specific requirements of spatial data services. Our Content Drift checking method uses the data format as an indicator. It compares the data formats declared in the metadata with the actual data types returned by the hyperlinks. Findings show that 10.41% of the distribution URIs suffer from Link Rot and at least 6.21% of records suffer from Content Drift (do not declare its distribution types correctly). Additionally, 14.94% of metadata records only contain intermediate HTML web pages as distribution URIs and 31.37% contain at least one HTML web page; thus, they cannot be accessed or checked directly.

^{*} This paper is part of the RD project PID2020-113353RB-I00, supported by the Spanish MCIN/AEI/10.13039/501100011033/ and the project T59_20R supported by the Aragon regional Government.