

# مشکل جستجوی واژگان در مقالات پی‌دی‌اف

## فارسی: نامه به سردبیر

امراه شمسی<sup>۱</sup>، محمدجواد منصورزاده<sup>۲\*</sup>

۱. کارشناس ارشد کتابداری و اطلاع‌رسانی پزشکی، دانشگاه علوم پزشکی بوشهر، بوشهر، ایران.
۲. مرکز توسعه پژوهش‌های بالینی، بیمارستان شهدای خلیج فارس، دانشگاه علوم پزشکی بوشهر، بوشهر، ایران.
۳. دانشجوی دکترای کتابداری و اطلاع‌رسانی پزشکی، دانشگاه علوم پزشکی تهران، تهران، ایران.

### سردبیر محترم

پی‌دی‌اف قالب متنی- تصویری پرکاربرد است که توسط شرکت ادوبی<sup>۱</sup> در سال ۱۹۹۰ گسترش یافت. این قالب با هدف کاهش مشکلات جابه‌جایی شکلی و ساختاری فایل‌های الکترونیکی ایجاد شد که توأمان با گسترش وب، جایگاه خود را به‌عنوان قالب کتاب‌های الکترونیک تثبیت کرد (۱). پی‌دی‌اف به‌دلیل ثابت بودن محتوای فایل و نمایش یکسان آن در تمام سیستم عامل‌ها بین مردم پرکاربرد و محبوب به‌شمار می‌آید. این فرمت به‌حدی موفق بوده که امروزه اکثر مرورگرهای اینترنتی نیز این فرمت را به‌صورت پیشفرض پشتیبانی می‌کنند.

فایل‌های پی‌دی‌اف اکنون قالب رایج اکثر مقالات تمام متن در مجلات علمی است. قالب پی‌دی‌اف ابزارهای متنوعی در اختیار خوانندگان قرار می‌دهد که از مهم‌ترین آنها برای پژوهشگران می‌توان به امکان جستجوی واژگان در متن و همچنین قابلیت کپی کردن قسمتی از متن بدون لطمه زدن به فایل اصلی اشاره کرد. در واقع چنین قابلیت‌هایی در فایل مقاله می‌تواند منجر به تسریع در روند بازیابی اطلاعات و مرور متن، صرفه‌جویی در وقت و در نهایت بهبود کارایی پژوهشگر و استفاده بهینه از مقاله مورد نظر گردد. متأسفانه در برخی از مقالات تمام متن فارسی امکان جستجوی فارسی و انتخاب و کپی کردن قسمتی از متن فعال نبوده و بعضاً با محدودیت‌های در بازیابی واژگان روبه‌رو هستیم که این اتفاق می‌تواند موجب

گمراهی یا حتی عدم استفاده از مقاله شود. به‌عنوان مثال در فایل پی‌دی‌اف مقاله‌ای تحت عنوان «بررسی ارتباط لیومیوم رحمی با پارگی زودرس» کلمه «تومور» مورد جستجو قرار گرفت و هیچ نتیجه‌ای یافت نشد (۲). این در حالی است که این واژه مکرراً در متن مقاله استفاده شده و بازیابی نشدن آن می‌تواند خواننده را گمراه و وی را از بررسی کامل مقاله منصرف کند.

این مشکلات در فایل‌های پی‌دی‌اف می‌تواند به دلایل عمدی مانند سیاست مجله در محدودسازی دسترسی به محتوا و یا به‌صورت سهوی و به‌علت مشکلات نرم‌افزار تبدیل‌کننده فایل پی‌دی‌اف، عدم پشتیبانی نرم‌افزار پی‌دی‌اف خوان، عدم یکسان بودن صفحه کلید سازنده محتوا و پژوهشگر و استفاده از حروف عربی (مانند استفاده از ه، ی و ک عربی به جای ه، ی و ک در فارسی) در متن مقاله اشاره نمود.

پژوهش‌های جدید نشان می‌دهد بیش از نیمی از مقالات مجلات منتشر شده به‌صورت دسترسی آزاد در دسترس عموم هستند و با اندکی بررسی می‌توان ادعا کرد که تقریباً تمام مجلات ایرانی از جریان دسترسی آزاد<sup>۲</sup> تبعیت می‌کنند (۳، ۴). اغلب مجلات دسترسی آزاد در ایران از مجوزهای انتشار CC<sup>۳</sup> در سطوح مختلف استفاده می‌کنند. این مجوز به خوانندگان امکان خواندن، کپی‌برداری و بازنشر، ترکیب و یا تغییر محتوا را در سطوح مختلف می‌دهد (۵). از مجلاتی که از مجوزهای دسترسی آزاد استفاده می‌کنند انتظار می‌رود تا بر اساس مجوزهای CC حقوق خوانندگان را

<sup>2</sup> Open access

<sup>3</sup> Creative Commons

<sup>1</sup> Adobe Inc

خوانندگان فارسی زبان از ابزار جستجو، انتخاب و کپی متن در مقالات به نحوی نقض حقوق خوانندگان و پژوهشگران است. از این رو به سردبیران و مدیران اجرایی مجلات فارسی پیشنهاد می شود به این مسئله بیشتر توجه کنند تا این امکان برای خوانندگان فراهم شود.

جهت سهولت استفاده رعایت کرده و به آنها احترام بگذارند. چنین مشکلاتی در مجلات فارسی در حالی دیده می شود که اکثر مجلات معتبر دنیا چه با دسترسی آزاد مثل مجموعه PLOS و چه مجلات با مجوزهای تجاری نظیر Lancet و Nature قابلیت های جستجو و کپی متون مقالات را برای کاربران خود فعال کرده اند. بنابراین محروم کردن

## منابع

1. Kuchta T, Lutellier T, Wong E, Tan L, Cadar C. On the correctness of electronic documents: studying, finding, and localizing inconsistency bugs in PDF readers and files. *Empir Softw Eng* 2018; 23(6):3187-3220.
2. Razavi M, Rashidi Fakari F, Rigi A, Farzaneh F. The Relationship between Uterine Leiomyoma with Premature Rupture of Membranes in Pregnant Women. *Iran J Obstet Gynecol Infertil* 2019; 22(1):15-19.
3. Piwowar H, Priem J, Larivière V, Alperin JP, Matthias L, Norlander B, et al. The State of OA: A large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ* 2018; 6:e4375.
4. Laakso M, Polonioli A. Open access in ethics research: an analysis of open access availability and author self-archiving behaviour in light of journal copyright restrictions. *Scientometrics* 2018; 116(1):291-317.
5. About The Licenses - Creative Commons. Available at: <https://creativecommons.org/licenses/?lang=en> Accessed: 23 Feb 2019.