



راهنمای خواندن نتایج شاخص جهانی داده باز



|  |  |
| --- | --- |
| شناسه سند | |
| موضوع | داده باز |
| طبقه‌بندی | عمومی |
| وضعیت انتشار | منتشر شده |
| نسخه | نهایی |
| تاریخ تصویب | ‏18 / 02 / 1396 |
| نویسنده / مشارکت کننده | محمدحسین خانی |

شاید برای شما هم سوال شده باشد 0% یا 100% باز یعنی چه؟ پاسخ این سوالات را از زبان دست‌اندرکاران شاخص داده باز بخوانید.

قصد داریم شیوه‌ی خواندن نتایج شاخص جهانی داده باز را مرور کنیم. در ادامه به سوالاتی شبیه سه مورد زیر پاسخ خواهیم داد:

* امتیاز نهایی به چه معناست؟
* چگونه امتیازات متغیر بین %0، %40 یا %70 را تفسیر کنم؟
* امتیاز 0% به چه معناست؟

# امتیاز یعنی چه؟

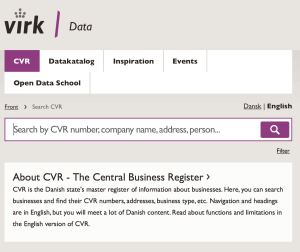
امتیازدهی ما (اعم از 0٪ باز تا 100٪ باز) لزوماً به معنای بهبود وضعیت داده باز نیست. به همین دلیل است که هر امتیاز کمتر از 100 درصد تنها نشان می‌دهد که یک مجموعه داده تا حدی، و نه کاملا، باز است. سطوحِ باز بودن شامل اطلاعات عمومی، داده‌هایی که دسترسی به آن‌ها کنترل‌شده است و همچنین شکاف‌های داده می‌شود.

برای مثال، امتیاز 70٪ می‌تواند حاکی از آن باشد که ما داده‌ی ماشین‌خوانی یافته‌ایم که دسترسی به آن کنترل‌شده است و نمی‌توان به‌دفعات آن را دانلود نمود.

توجه داشته باشید امتیاز زیر 100٪ به معنی عدم دسترسی، دسترسی محدود یا دسترسی عمومی است. در اینجا ما به توضیح هر یک از این مفاهیم در نظر و عمل می‌پردازیم.

# داده در دسترس عموم

زمانی داده به‌صورت عمومی در دسترس است که مردم بتوانند آن را آنلاین و بدون هیچ محدودیت دسترسی مشاهده نمایند. این بدان معنا نیست که میتوان داده‌ها را دانلود نمود و یا آن را آزادانه مورداستفاده مجدد قرار داد. اغلب بدان معناست که داده‌ها در قالب HTML در یک وب‌سایت ارائه‌شده‌اند.



تصویر بالا یک رابط جستجوی ثبت شرکت را نشان می‌دهد. این رابط امکان جستجوی هدفمند را برای شرکت‌های مستقل فراهم می‌سازد اما امکان بازیابی تمام داده‌ها به‌صورت همزمان وجود ندارد. نتایج جستجوهای شخصی (نه انبوه) در فرمت HTML نمایش داده‌شده که می‌توان آن را به‌صورت PDF دانلود نمود (دقت داشته باشید که PDF فرمتی ماشین‌خوان نیست.)؛ بنابراین امتیاز 70٪ به آن تعلق میگیرد و باز بودن آن به‌صورت زیر در نوار آیکون رتبه‌بندی ما ترسیم میشود:



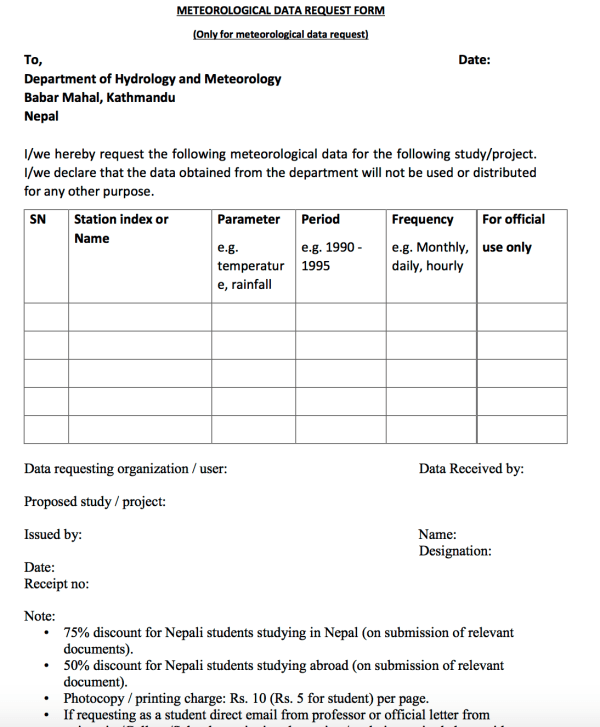
# داده‌ای که دسترسی به آن کنترل‌شده است

زمانی دسترسی به یک داده کنترل‌شده است که ارائه‌دهنده تعیین کند که داده توسط چه کسی، چه زمانی و چگونه قابل دسترسی است. کنترل دسترسی شامل این موارد است:

* ثبت‌نام
* شناسایی و احراز هویت
* فرمهای درخواست داده، توافقنامه اشتراک‌گذاری داده (قید نمودن موارد استفاده)
* سفارش یا خرید داده.

دلایل بسیاری برای ایجاد کنترل دسترسی در داده وجود دارد که مدیریت ترافیک وب‌سایت یا نظارت بر نحوه استفاده از جمله‌ی آن‌هاست. از جمله موضوعات چالش برانگیز داده باز، پذیرش یا عدم پذیرش داده‌های نیازمند ثبت نام و احراز هویت به عنوان باز است. اما در مورد داده‌هایی که دسترسی به آن‌ها نیازمند تکمیل فرم‌های درخواست داده است، اتفاق نظر بر این است که جزء داده‌ باز به حساب نمی‌آیند.

تصویر زیر، یک فرم درخواست داده است. دسترسی به داده‌ها نیازمند پرداخت هزینه‌ی مشخصی است. این مورد، مانع از ارزیابی داده توسط تیم تحقیقاتی ما می‌شود و قادر به بررسی فرمتی که داده در آن قالب ارائه خواهد شد نیستیم و از سوی دیگر نمیتوانیم بررسی کنیم که آیا داده‌ها واقعاً داده‌های پیش‌بینی آب‌وهوا هستند یا خیر (موردی که مدنظر ما بود)؛ بنابراین این داده‌ها که دسترسی به آن‌ها کنترل‌شده امتیاز 0 کسب کردند و 0٪ باز بودند. در مقابل، داده‌های کنترل دسترسی شده اغلب امتیاز بالایی، تا 85٪ ، کسب کردند (زیرا ما برای کنترل‌های دسترسی نظیر الزامات ثبت‌نام 15 امتیاز از 100 کم می‌کنیم).



# چگونه امتیاز 0% را بخوانیم؟

دلایل بسیاری مبنی بر اینکه چرا مجموعه داده‌ها امتیاز 0٪ دریافت می‌کنند وجود دارد. دلایل اصلی را در اینجا مشاهده میفرمایید:

## **شکاف داده (Data Gaps)**

شکاف داده میتواند به این معنی باشد که دولت هیچ‌گونه داده‌ای در دسته‌بندی داده‌شده ارائه نکرده است. در برخی موارد، اگر GODI امتیاز صفر درصد را نشان دهد شاهد شکاف داده خواهیم بود. به‌عنوان‌مثال، در کشورهای آفریقای غربی که سیستم کنترل کیفیت هوا در آن‌ها وجود ندارد یا کشورهایی که هیچ سیستم کد پستی ندارند. شکاف داده نشان می‌دهد که سیستم‌های اطلاعاتی دولت آماده تولید داده باز نیستند، گاهی اوقات به این دلیل که منابع از بین رفته‌اند، یا اساسا از اولویت‌های دولت مدنظر نبوده‌اند.

## موجود است، اما تنها برای استفاده دولتی

گاهی اوقات دولت داده را در اختیار دارد اما به دلایل بسیاری را آن را در اختیار عموم مردم قرار نمی‌دهد.

## **فاقد جزئیات**

ازآنجاکه معیار ما به دنبال جزئیات است تمام مجموعه داده‌هایی که به این سطح از جزئیات نرسیده باشند را غیرقابل‌دسترس قلمداد می‌کنیم.

به‌عنوان‌مثال، بریتانیا نتایج انتخابات را منتشر کرده است، اما نه در سطح حوزه رأی‌گیری که یک سطح مهم برای تشخیص تقلب محسوب می‌شود؛ بنابراین، بااینکه داده‌هایی در مورد انتخابات بریتانیا وجود دارد، اما سطح آن‌ها مناسب نیست و به‌عنوان ناموجود در نظر گرفته می‌شود.

## با معیارهای ما سازگار نیست

ما به دنبال مجموعه داده‌هایی خاص در GODI هستیم. هنگامی‌که این مجموعه ها دارای تمام ویژگی مدنظرمان نیستند آن‌ها را به‌عنوان غیر قابل دسترس قلمداد می‌کنیم.

برای مطالعه بیشتر پیرامون نحوه خواندن نتایج GODI به index.okfn.org/interpretation مراجعه فرمایید.

<https://blog.okfn.org/2017/05/02/how-to-read-the-global-open-data-results/>