

بيانات استخدام الكتب ذات الوصول المفتوح: ما مدى تقارب إحصاءات (COUNTER) وغيرها؟

رونالد سنايدر



الملخص:

يقارن هذا المقال بين خدمتين واسعتي الاستخدام لقياس استخدام الكتب ذات الوصول المفتوح، هما (Google Analytics) والإصدار الخامس من مدونة ممارسات (COUNTER) كما تطبقها (IRUS-UK)، مستفيدًا من انتقال مكتبة (OAPEN) إلى منصة جديدة مبنية على (DSpace 6) في أبريل 2020. يحلل المقال بيانات تنزيل نحو 11000 كتاب وفصل في المدة من 15 أبريل إلى 31 يوليو 2020، بعد مراجعة لأدبيات قياس أثر هذه الكتب، ويبيّن أن (Google Analytics) سجلت أكثر من 3.6 ملايين تنزيل مقابل 1.5 مليون في (COUNTER)، وأن الفروق تتفاوت تفاوتًا كبيرًا من بلد إلى آخر، وأن 6% فقط من العناوين الأعلى ترتيبًا جاءت متقاربة الترتيب في الخدمتين، مع أمثلة تفصيلية لخمس عناوين. ويخلص إلى أنه لا توجد قاعدة بسيطة للتحويل بين المقياسين، وأن اختيار خدمة القياس يؤثر تأثيرًا كبيرًا في النتائج، فينبغي الحذر في الاستنتاج والشفافية في الخيارات.

الكلمات المفتاحية: الكتب ذات الوصول المفتوح، بيانات الاستخدام، (COUNTER)، (Google Analytics)، قياس الأثر

مقدمة:

منذ إطلاقها في عام 2010، أتاحت مكتبة (OAPEN) كتبًا وفصولًا خضعت للتحكيم العلمي بالوصول المفتوح (open access, OA). [1] وبحلول فبراير 2021، كانت المجموعة قد نمت إلى 14500 كتاب وأكثر من 700 فصل. وابتداءً من يونيو 2013، زودتنا (IRUS-UK) ببيانات استخدام متوافقة مع الإصدار الرابع من (COUNTER) لمكتبة (OAPEN). [2] وتجاوزت المكتبة عتبة عشرة ملايين تنزيل في الربع الأول من عام 2020.

وفي أبريل 2020، انتقلت مكتبة (OAPEN) إلى منصة جديدة مبنية على (DSpace 6)، وهو نظام المستودعات مفتوح المصدر. ومن بين ما أتاحه ذلك أننا صرنا قادرين على رصد جميع الأحداث التي تجري على المنصة باستخدام (Google Analytics, GA). وفي المدة نفسها، بدأت (IRUS-UK) العمل على تطبيق الإصدار الخامس (R5) من مدونة ممارسات (COUNTER).

ولذلك فهذه لحظة مناسبة للمقارنة بين هذين المقياسين الواسعي الاستخدام لبيانات الاستخدام. ونهدف، بوصف بيانات استخدام مكتبة (OAPEN) المستمدة من (Google Analytics) ومن الإصدار الخامس من (COUNTER)، إلى فهم الفروق فهماً أفضل. ولا نقصد إصدار حكم، إذ لا نرى أن أحدهما أفضل من الآخر. فهذان النظامان طوّرا من منظورين مختلفين: فبينما صُمم (GA) على النحو الأمثل لوصف ما يجري على موقع ويب معين، ولا سيما من منظور التسويق والمبيعات، يهدف (COUNTER) إلى توفير بيانات موحدة يمكن استخدامها للتجميع والمقارنة عبر بيانات متعددة.

مراجعة الأدبيات:

ليس هذا المقال الأول بأي حال في دراسة بيانات استخدام الكتب ذات الوصول المفتوح. إذ يمكن النظر إلى بيانات الاستخدام بوصفها مؤشراً على أثر هذه الكتب: فكثيراً ما يُستخدم الانتشار الجغرافي وعدد التنزيلات مؤشرين. وإلى جانب التنزيلات، تحظى بيانات الاستشهادات والمقاييس البديلة (altmetrics) أيضاً باهتمام الباحثين، وقد درس عدد من الناشرين أثر الوصول المفتوح في كتبهم.

ففي دراسة حالة لمطبعة كلية لندن الجامعية (UCL Press)، قارن مونتغمري وآخرون (Montgomery et al.) [3] عدة مصادر لأرقام التنزيل لفهم كيفية تأثرها بالأحداث المهمة المتعلقة بالترويج للعناوين المنشورة. ولم يكن (GA) مهياً لتسجيل أرقام التنزيل، بل استُخدم هنا لتوفير معلومات عن زوار موقع الكلية. وشددت الدراسة كثيراً على أن كل منصة توفر بيانات استخدام مبنية على مبادئ مختلفة، ولذلك لم تُجمع أرقام التنزيل من المستودعات الثلاثة.

وحللت مطبعة جامعة ستوكهولم (Stockholm University Press) إحصاءات الاستخدام وبيانات الاستشهادات والمقاييس البديلة، مقرونة باستبانة لمواقف وسلوك المؤلفين والمحررين الذين نشروا كتباً بالوصول المفتوح. [4] وخلص المؤلفون إلى أن هناك فروقاً داخل تخصصات أكاديمية بعينها، لكنهم ذكروا أيضاً أن تفسير المقاييس ما يزال معقداً.

وأجرت شركة (Springer Nature) دراسة حالة، استناداً إلى 3934 كتاباً منها 281 كتاباً بالوصول المفتوح، تبحث في الفروق في الأثر بين الكتب المنشورة بالوصول المفتوح والكتب المنشورة بطريقة مغلقة. [5] وخلص المؤلفون إلى أن إتاحة الكتب بالوصول المفتوح زادت عدد التنزيلات، وأن الانتشار الجغرافي، ولا سيما التنزيلات من البلدان المنخفضة والمتوسطة الدخل، اتسع كذلك. وفضلاً عن ذلك، استشهد بالكتب ذات الوصول المفتوح ضعف ما استشهد بنظيراتها «المغلقة».

ومؤخراً، تناولت دراسة تايلور (Taylor) [6] عدد المرات التي تُذكر فيها الكتب ذات الوصول المفتوح في الشبكات الاجتماعية ووسائل الإعلام الجماهيري والمدونات ووثائق السياسات. وبحسب الدراسة، هناك «ميزة للوصول المفتوح»، لكن الآليات الكامنة وراءها ليست واضحة في الوقت الحاضر. ومرة أخرى، تظهر الفروق بين التخصصات الأكاديمية.

ومن المحاولات الأخرى لفهم أثر الكتب ذات الوصول المفتوح مقال سنايدر (Snijder). [7] فبتصنيف المستخدمين، سعى المؤلف إلى جمع بيانات كمية عن الأثر العلمي والصلة المجتمعية للعناوين المنزلة. ومن البيانات المقيسة، ارتبط أكثر من 27% مباشرة بمستخدمين أكاديميين، في حين أن أكثر من 45% من التنزيلات يُرجح بدرجة عالية أنها صادرة عن عامة الجمهور أو غيرهم من غير الأكاديميين، وهو مؤشر محتمل على الأثر المجتمعي.

وقدّم أوزايغن (Ozaygen) [8] تحليلًا تقنيًا موسعًا لبيانات استخدام مجموعة من 28 كتابًا حديث النشر بالوصول المفتوح في عدة تخصصات أكاديمية، قدمها 13 ناشراً. وكانت هذه المجموعة التجريبية لمبادرة Knowledge (Unlatched) [9] التي أُتيحت في عام 2013. ويجمع هذا التحليل بين عدة تقنيات لتقديم صورة شاملة عن كيفية استخدام الكتب أو إتاحتها على الويب، وعن أماكن ذلك.

وللعثور على أثر كتاب بعينه من الكتب ذات الوصول المفتوح وتحليله، يحتاج المرء إلى قدر كبير من الوقت والجهد. وللمساعدة على حل هذه المشكلة، يجري إنشاء صندوق استثماني لبيانات استخدام الكتب الإلكترونية ذات الوصول المفتوح (Open Access eBook Usage Data Trust, OAEBU). [10] وهو مشروع تجريبي مدته عامان لتطوير واختبار نماذج للبنية التحتية والسياسات والحوكمة من أجل إنشاء صندوق استثماني لبيانات عالمي لبيانات استخدام الكتب ذات الوصول المفتوح. [11] وإلى جانب جمع بيانات الاستخدام، يهدف الصندوق الاستثماني للبيانات إلى التوافق مع أولويات المؤلفين والمؤسسات مع احترام المعايير الأخلاقية في استخدام المقاييس.

وتهدف مدونة ممارسات (COUNTER) إلى تزويد المكتبات بإحصاءات متسقة وقابلة للمقارنة عن الموارد الإلكترونية التي تقيتها. [12] فالمكتبات لا تحتاج إلى قياس كيفية استخدام مواردها الخارجية وتقييمها فحسب، إذ قد تكون هناك موارد تتوقف أسعارها على الاستخدام، بل تحتاج أيضًا إلى تقدير دور المكتبة نفسها تقديرًا كميًا. [13] وتستخدم المكتبات أيضًا (GA) أداة للمساعدة على توضيح كيفية استخدام موقعها على الويب، بما في ذلك فهرس المكتبة. [14]

ومن مراجعة الأدبيات هذه، يمكننا أن نستنتج أن بيانات استخدام الكتب والفصول ذات الوصول المفتوح تؤدي دورًا مهمًا للناسخين والمكتبات على حد سواء. ومن الواضح أيضًا أن الحصول على البيانات أبعد ما يكون عن السهولة، وفوق ذلك ما تزال هناك أوجه كثيرة من عدم اليقين: فروق في أنواع البيانات (التنزيلات، والاستشهادات، والمقاييس البديلة) الآتية من منصات متعددة قد تولّد بيانات غير قابلة للمقارنة. ويضاف إلى ذلك ضرورة تفسير النتائج. وسيوضح القسم التالي الفروق بين مقياسين واسعي الاستخدام يبلغان عن الحدث نفسه: تنزيل الكتب والفصول من مكتبة (OAPEN).

البيانات:

يناقش هذا المقال بيانات تنزيل ما يقرب من 11000 كتاب وفصل من مكتبة (OAPEN)، في المدة من 15 أبريل 2020 إلى 31 يوليو 2020. فحين يُنزل كتاب أو فصل، يسجّل (GA) ذلك، وفي الوقت نفسه تُرسل إشارة إلى (IRUS-UK). وقد استُخدمت النتائج المبلغ عنها للمقارنة في هذا المقال.

ويسجّل (GA) أشياء أكثر بكثير من التنزيلات: فهو يرصد جميع الزيارات إلى موقع الويب ويجمع معلومات عن الزوار. والتحدي هو العثور على بيانات الاستخدام ذات الصلة بهذه المقارنة دون غيرها. وقد أنشأنا تقريرًا مخصصًا يرصد التنزيلات، لا زيارات صفحات الويب. ويُسمّى ذلك في (GA) «حدثًا» (event) في الفئة «(Bitstream)». ولا تحتوي

بيئة (DSpace) الخاصة بـ (OAPEN) على «ملفات الكتب» فحسب، بل يرافق كل عنوان أيضًا ملف صورة للفلاف. وقد استبعدنا تنزيلات ملفات الأغلفة من التقارير. فضلًا عن ذلك، ولضمان استخدام بيانات قابلة للمقارنة، تُستبعد «الروبوتات» (bots) المعروفة.

أما جمع البيانات في (IRUS-UK) فيركز على استخدام المنشورات فحسب. إذ تُقيّم التنزيلات وفقًا لإرشادات (R5) ويبلغ عنها في «تقرير تصفية المواد» (Item Filter Report). واستخدمنا هنا المقياس «(Total_Item_Request)»، الذي يُعرّف بأنه العدد الإجمالي للمرات التي تُزل فيها النص الكامل لمادة محتوى أو عُرض. ومن الأمور الجوهرية في إعداد تقارير (COUNTER) إزالة أي بيانات استخدام يُعدّ أنها لم تصدر عن قصد من مستخدم بشري. [15] ومن ثم تُستبعد التنزيلات الآلية التي تجربها «الروبوتات».

وتتيح كلتا المنصتين، (GA) و(R5)، إمكانية إزالة التكرار من بيانات الاستخدام، وتُسمّى «الأحداث الفريدة» (Unique Events) في (GA) و«(Unique_Item_Requests)» في (R5). ولأننا لم نكن متأكدين من أن المنصتين تستخدمان التعريفات نفسها للحدث الفريد، قررنا عدم استخدام هذا المقياس. وترد معايير اختيار البيانات في الجدول 1.

الجدول 1: اختيار بيانات (GA) و(R5):

(COUNTER R5)	(Google Analytics)	
(IRUS-UK), (Jisc)	(Google)	الجهة المزودة
تقرير المواد (COUNTER R5 Item Report) / عوامل التصفية المستخدمة: المستودع: (OAPEN DSpace)	فئة الحدث: (Bitstream) / عوامل التصفية المستخدمة: استبعاد (jpg) من (Page) (جزء عنوان (URL) الذي يلي اسم النطاق) / تصفية الروبوتات: استبعاد جميع الزيارات من الروبوتات وبرامج الزحف (spiders) المعروفة	التقرير/عامل التصفية
(Total_Item_Requests)	(Total Events)	البيانات المستخدمة
شهري	يومي	التواتر
15 أبريل 2020 إلى 31 يوليو 2020	15 أبريل 2020 إلى 31 يوليو 2020	المدة
17	16	الوصف

ويُنتج عن ذلك مجموعتا بيانات: التنزيلات الشهرية التي يقيسها (GA)، والاستخدام الذي يبلغ عنه (R5)، مجتمعةً كذلك حسب الشهر. وتتألف كلتا المجموعتين من العدد الإجمالي للتنزيلات لكل عنوان، موزعًا حسب البلد. وهكذا، في يوليو 2020، وبحسب بيانات (R5)، نُزل كتاب «الإثنية والعرق وعدم المساواة في المملكة المتحدة» (Ethnicity, Race and Inequality in the UK) [18] 1433 مرة، وكان القراء يقيمون في 54 بلدًا مختلفًا. وحين ننظر في بيانات (GA)، تختلف الصورة قليلًا: 1360 تنزيلًا من 21 بلدًا.

وفي هذا المثال، يكون الفرق بين (R5) و(GA) صغيرًا نسبيًا، لكن يظهر عادةً تفاوت كبير بين مجموعتي البيانات. وعمومًا، لكي تكون بيانات الاستخدام متوافقة مع (COUNTER)، يجب أن تخضع لقواعد أكثر صرامة كي يُبلغ عنها، مقارنةً بقياسات (GA). وعند مقارنة العدد الإجمالي للتنزيلات، تبلغ بيانات (R5) نسبة 58% من إجمالي (GA).

وفي الأقسام التالية، سنقارن بيانات (GA) و (R5) على عدة مستويات: بدءًا بالمجاميع، مرورًا ببيانات البلدان، وصولًا إلى المقارنة على مستوى الكتاب. وجميع البيانات متاحة، ويمكن الاطلاع على التفاصيل في بيان إتاحة البيانات في نهاية هذا المقال.

الاستخدام الإجمالي:

كما ذكر من قبل، فإن عدد التنزيلات الذي يبلغ عنه (GA) أكبر بكثير من (R5). فالعدد الإجمالي للتنزيلات في (GA) يتجاوز 3.6 ملايين: أي أكثر من مليون تنزيل شهريًا. وفي المقابل، يبلغ العدد الذي يبلغ عنه (R5) مليونًا ونصف المليون (1.5 مليون) تنزيل: أي نحو 400000 تنزيل شهريًا.

وعند النظر في البيانات الشهرية، يتضح أن العلاقة بين (GA) و (R5) ليست مباشرة تمامًا. فالفرق بالنسبة المئوية يتفاوت من شهر إلى آخر: ففي مايو كان الفرق 54%، وارتفع إلى 64% في يوليو. وبطبيعة الحال، لا تكفي ثلاثة أشهر لإثبات وجود اتجاه، لكن سيكون من المثير للاهتمام إجراء تحليل آخر بعد عام.

المقارنة حسب البلد:

تُظهر بيانات الاستخدام الإجمالي موزعةً حسب البلد انتشار الاستخدام في معظم بلدان العالم؛ فكل من (GA) و (R5) يرصد تنزيلات من أفغانستان إلى زيمبابوي. وتتبع البيانات النمط نفسه: عدد محدود من البلدان تُنزل فيه الكتب والفصول بكثرة نسبيًا، وعدد كبير من البلدان يقل فيه الاستخدام، وهو ما يُعرف بنمط «الذيل الطويل» (long tail).

والأكثر إثارة للاهتمام هو النظر في الفروق بين البلدان «الخمس عشرة الأولى»: الولايات المتحدة، وألمانيا، والمملكة المتحدة، وفرنسا، والهند، وأستراليا، والصين، وهولندا، وروسيا، وإندونيسيا، وكندا، وإيطاليا، وجنوب أفريقيا، والنمسا، وإسبانيا. فالوصول المفتوح ظاهرة عالمية بوضوح، لا تقتصر على البلدان الأكثر ثراءً. وتفضي مقارنة العدد الإجمالي للتنزيلات لهذه البلدان إلى استنتاج مألوف: فإجمالي (R5) يبلغ 58% من إجمالي (GA). وهذا يتسق مع نمط الاستخدام الإجمالي.

غير أن مقارنة بيانات (R5) و (GA) على مستوى كل بلد على حدة، تُظهر فروقًا كبيرة. فـ (GA) يورد أكثر من خمسة أضعاف عدد التنزيلات للولايات المتحدة وفرنسا والصين وروسيا. وفي المقابل، فإن الأرقام الخاصة بأستراليا وكندا والنمسا متقاربة.

وتكشف المقارنة بطريقة أخرى عن فروق بين (GA) و (R5). فبحسب بيانات (GA)، تستحوذ عناوين الشبكة الموجودة في الولايات المتحدة على النصيب الأكبر من الاستخدام؛ إذ تبلغ التنزيلات الأمريكية قرابة ثلث الإجمالي، أي ثلاثة أضعاف ألمانيا، البلد الثاني. أما بيانات (R5) فتظهر توزيعًا أكثر توازنًا، بفروق أصغر بكثير بين البلدان الأكثر تنزيلًا.

الفروق على مستوى العنوان:

المستوى الأخير الذي سناقشه هو الفروق بين (GA) و(R5) عند النظر في العناوين الفردية. ونظرًا إلى أن مجموعتي بياناتنا تحتويان على ما يقرب من 11000 عنوان، فإن مناقشة كل عنوان مناقشة مستفيضة ستؤدي إلى التكرار دون فائدة كبيرة. غير أن مقارنة ترتيب العناوين تساعد على تكوين صورة. وببساطة: يُرتَّب كل كتاب وفقًا لعدد التنزيلات، بحيث يحتل الكتاب صاحب أعلى عدد من التنزيلات المرتبة الأولى، وهكذا. والخطوة التالية هي مقارنة ترتيب العناوين: فالفروق بين ترتيب (GA) وترتيب (R5) توضح اختلاف قياس استخدام الكتاب أو الفصل.

ويوضح الجدول 2 مؤشرًا أوليًا على التفاوت الكبير بين (GA) و(R5). فعند النظر في العناوين الـ 500 الأعلى ترتيبًا في (GA) التي هي أيضًا ضمن العناوين الـ 1000 الأعلى ترتيبًا في (R5)، لم يتقارب ترتيب سوى 6% من العناوين تقاربًا نسبيًا. ومن أمثلة ذلك كتاب «فرانكنشتاين» (Frankenstein)، [19] الذي جاء في المرتبة الخامسة في (GA) والمرتبة الثالثة في (R5). كما أن جزءًا صغيرًا نسبيًا، هو 20%، يتألف من كتب وفصول يقع ترتيبها في حدود 50 مرتبة بعضها من بعض. ويمكن هنا الاستشهاد بكتاب «الأساطير التي صنعت أمريكا» (The Myths That Made America) [20] مثالًا: إذ جاء في المرتبة السادسة في (GA) والمرتبة الثالثة والعشرين في (R5). أما المجموعة الكبرى من العناوين فتتباعد مراتبها أكثر، مثل كتاب «صحة الناس وصحة الكوكب ومسؤوليتنا» (Health of People, Health of Planet and Our Responsibility) [21] الذي جاء في المرتبة الثانية في (GA) لكنه جاء في المرتبة الثالثة والثمانين في (R5).

الجدول 2: مقارنة ترتيب (GA) و(R5):

الترتيب	عدد العناوين	النسبة المئوية
الفرق > 10	30	6%
الفرق > 50	98	20%
فرق أكبر	372	74%
المجموع	500	100%

وتبرز الفروق في ترتيب الكتب بوضوح: فكثير من العناوين التي تحتل مراتب متقدمة في (GA) تأتي في مراتب متأخرة في (R5)، والعكس صحيح، دون نمط واضح يفسر هذا التفاوت.

وتصف الأقسام الفرعية التالية الكتب الخمسة الأعلى ترتيبًا في (GA) مع «بياناتها المقابلة» في (R5). وسنرى أن استخدام كل عنوان يُمثل تمثيلًا مختلفًا تمامًا في (GA) و(R5)، وأن العنوان صاحب المرتبة الأولى في (GA)، وهو «(m-Learning - die neue Welle? Mobiles Lernen für Deutsch als Fremdsprache)» («التعلم المتنقل: الموجة الجديدة؟ التعلم المتنقل للألمانية لغةً أجنبية»)، [22] لم يُعثر عليه ضمن العناوين الـ 1000 الأولى ترتيبًا في (R5). ولذلك فإن العنوان الأول في القسم التالي هو ثاني أكثر العناوين تنزيلًا في بيانات (GA).

صحة الكوكب ومسؤوليتنا): (Health of People, Health of Planet and Our Responsibility) صحة الناس

لهذا العنوان، يبلغ (GA) عن قرابة 28000 تنزيل، وفي المقابل، لا يبلغ (R5) إلا عما يزيد قليلاً على 1200 تنزيل، مما يفضي إلى تفاوت كبير في الترتيب. وترد بيانات الترتيب والتنزيل في الجدول 3.

الجدول 3: ترتيب كتاب (Health of People, Health of Planet and Our Responsibility) وتنزيلاته:

(R5)	(GA)	
83	2	المرتبة
1,274	27,962	عدد التنزيلات

(Access Controlled) (الوصول المتحكم فيه):

في حالة هذا الكتاب، جرى أكثر من 7000 تنزيل في يوم واحد. وتستبعد (IRUS-UK) المستخدمين الذين ينزلون 40 منشورًا أو أكثر في يوم واحد، أو أولئك الذين ينزلون المنشور نفسه أكثر من 10 مرات في يوم واحد. وترد بيانات الترتيب والتنزيل في الجدول 4.

الجدول 4: ترتيب كتاب (Access Controlled) وتنزيلاته:

(R5)	(GA)	
119	4	المرتبة
1,083	12,457	عدد التنزيلات

(Frankenstein) (فرانكنشتاين):

يبلغ كل من (GA) و (R5) عن تنزيلات كثيرة، مما يفضي إلى ترتيب عالٍ في المقياسين كليهما. ويتناقض نمط التنزيل تناقضًا صارخًا مع «تنزيل اليوم الواحد» لكتاب (Access Controlled). وترد البيانات في الجدول 5.

الجدول 5: ترتيب كتاب (Frankenstein) وتنزيلاته:

(R5)	(GA)	
3	5	المرتبة
5,675	8,583	عدد التنزيلات

(The Myths That Made America) (الأساطير التي صنعت أمريكا):

للهولة الأولى، يبدو نمط الاستخدام شبيهًا جدًا بنمط كتاب (Access Controlled): ذروة استخدام في مدة قصيرة. غير أن الذرى ليست بالحجم نفسه: فأعلى عدد من التنزيلات في يوم واحد لم يتجاوز 1,300، وهو أقل بكثير من التنزيلات الـ 7000 لكتاب (Access Controlled). وترد البيانات في الجدول 6.

الجدول 6: ترتيب كتاب (The Myths That Made America) وتنزيلاته:

	(GA)	(R5)
المرتبة	6	23
عدد التنزيلات	8,299	2,098

(Ethnicity, Race and Inequality in the UK) (الإثنية والعرق وعدم المساواة في المملكة المتحدة):

لهذا العنوان، فإن عدد التنزيلات الذي سجله (GA) أقل في الواقع من الذي سجله (R5). وهذا يخالف النمط المعتاد في المقارنة. وترد البيانات في الجدول 7.

الجدول 7: ترتيب كتاب (Ethnicity, Race and Inequality in the UK) وتنزيلاته:

	(GA)	(R5)
المرتبة	7	1
عدد التنزيلات	7,828	10,511

الاستنتاجات:

بيانات استخدام الكتب ذات الوصول المفتوح مهمة لكثير من أصحاب المصلحة. غير أنه لا يوجد معيار مقبول عالميًا يستخدمه جميع مزودي مجموعات الكتب ذات الوصول المفتوح. وبصرف النظر عن الفروق في البيانات المقدمة، فإن جمع البيانات ليس مهمة سهلة.

ويهدف هذا المقال إلى بيان كيف تبلغ خدمتان واسعتا الاستخدام للقياس، هما (Google Analytics) والإصدار الخامس من (COUNTER)، عن الأحداث نفسها. فقد اتخذت كل من الخدمتين خياراتها الخاصة بشأن النتائج المعروضة وما لا يُبلغ عنه. وظهرت فروق كبيرة خلال المدة المدروسة: إذ أبلغ (GA) عن 3.6 ملايين تنزيل مقابل 1.6 مليون تنزيل ذكرها (R5). وفضلًا عن ذلك، لا توجد قاعدة تقريبية بسيطة «لتحويل» مقاييس (GA) إلى (R5): فعلى مستوى مجاميع البلدان وعلى مستوى العناوين الفردية، نرى أرقامًا متباينة تباينًا شديدًا. فعلى سبيل المثال، بيانات الاستخدام كما يبلغ عنها (GA) أعلى بكثير مقارنةً بـ (R5) بالنسبة إلى الولايات المتحدة، في حين أن البيانات الخاصة بأستراليا متطابقة تقريبًا. ويصدق ذلك أيضًا على كتاب (Access Controlled) مقارنةً بكتاب (Frankenstein).

وقد يكون من المفري استنتاج أن الاستخدام كما يبلغ عنه (GA) «أصدق»، لأنه يبدو أقل تقييداً فيما يُقاس. وليس الأمر كذلك. فأولاً، خضعت بيانات (GA) في هذه المقارنة بالفعل للتصفية لإزالة الاستخدام الصادر عن «الروبوتات» المعروفة. وثانياً، كما أظهر مثال كتاب (Ethnicity, Race and Inequality in the UK)، قد يكون الاستخدام الذي يبلغ عنه (GA) أحياناً أكثر تقييداً من (R5).

وما اتضح جلياً هو أن لاختيار خدمة القياس عواقب كبيرة على النتائج المعروضة. ومن ثم ينبغي استخلاص الاستنتاجات من النتائج بحذر. فعلى سبيل المثال، ما الذي ينبغي أن نستنتجه من أن العنوان الأكثر تنزيلاً بحسب (GA) لم يُعثر عليه حتى ضمن العناوين الـ 1000 الأكثر تنزيلاً بحسب (R5)؟

فليس أحد المقياسين أفضل من الآخر، لكن ينبغي أن نوضح الخيارات التي اتخذناها. ففي نهاية المطاف، مقياس الكتب ذات الوصول المفتوح معقدة، ولن نجني من الوضوح إلا الفائدة.

بيان إتاحة البيانات:

البيانات كاملة متاحة في مستودع البيانات المشار إليه في المراجع.

المراجع:

1. "OAPEN Library", OAPEN Foundation, 2010, <https://www.oapen.org> (accessed 15 February 2021).
2. "International Collaboration and Value: Working with OAPEN", IRUS-UK Case Studies, 2018, https://irus.jisc.ac.uk/documents/IRUS-UK_working_with_OAPEN.pdf (accessed 15 February 2021).
3. Lucy Montgomery et al., "Getting the Best out of Data for Open Access Monograph Presses: A Case Study of UCL Press," Learned Publishing 31, no. 4 (1 October 2018): 335-44, DOI: <https://doi.org/10.1002/leap.1168> (accessed 17 February 2021).
4. Sofie Wennström et al., "The Significant Difference in Impact: An Exploratory Study about the Meaning and Value of Metrics for Open Access Monographs," ELPUB 2019 23d International Conference on Electronic Publishing (June 2019), DOI: <https://doi.org/10.4000/proceedings.elpub.2019.9> (accessed 15 February 2021).
5. Ros Pyne et al., "The Future of Open Access Books: Findings from a Global Survey of Academic Book Authors," Springer Nature Open Access Books, (June 2019), DOI: <https://doi.org/10.6084/m9.figshare.8166599> (accessed 15 February 2021).
6. Megan Taylor, "Mapping the Publishing Challenges for an Open Access University Press," Publications 7, no. 4 (December 2019): 63, DOI: <https://doi.org/10.3390/publications7040063> (accessed 17 February 2021).

7. Ronald Snijder, "Measuring Monographs: A Quantitative Method to Assess Scientific Impact and Societal Relevance," *First Monday* 18, no. 5 (May 2013), DOI: <https://doi.org/10.5210/fm.v18i5.4250> (accessed 17 February 2021).
8. Alkim Ozaygen, "Analysing the Usage Data of Open Access Scholarly Books: What Can Data Tell Us?," (Thesis, Curtin University, 2019), <https://espace.curtin.edu.au/handle/20.500.11937/79585> (accessed 17 February 2021).
9. Lucy, Montgomery and Celeste Feather, "Knowledge Unlatched Pilot Collection to Become Open Access - Nearly 300 Libraries Globally Pledge Their Support," *Knowledge Unlatched* (blog), 10 March 2014, <https://knowledgeunlatched.org/2014/03/knowledge-unlatched-pilot-collection-to-become-open-access-nearly-300-libraries-globally-pledge-their-support/> (accessed 19 February 2021).
10. "Developing a Pilot Data Trust for Open Access Ebook Usage," *Educopia Institute*, https://educopia.org/data_trust/ (accessed 17 February 2021).
11. Christina Drummond, "Engaging Stakeholder Networks to Support Global OA Monograph Usage Analytics," *Collaborative Librarianship* 12, no. 2 (21 October 2020), <https://digitalcommons.du.edu/collaborativelibrarianship/vol12/iss2/9> (accessed 15 February 2021); "Developing a Pilot Data Trust for Open Access Ebook Usage," *Educopia Institute*, https://educopia.org/data_trust/ (accessed 15 February 2021).
12. "COUNTER | About Us", COUNTER, 2014, <http://www.projectcounter.org/about.html> (accessed 15 February 2021).
13. Gayle Baker and Eleanor J. Read, "Vendor-Supplied Usage Data for Electronic Resources: A Survey of Academic Libraries," *Learned Publishing* 21, no. 1 (1 January 2008): 48-57, DOI: <https://doi.org/10.1087/095315108X247276> (accessed 15 February 2021); Virginia Kinman, "E-Metrics and Library Assessment in Action," *Journal of Electronic Resources Librarianship* 21, no. 1 (16 April 2009): 15-36, DOI: <https://doi.org/10.1080/19411260902858318> (accessed 15 February 2021).
14. Tabatha Farney and Nina McHale, "Introducing Google Analytics for Libraries," *Library Technology Reports* 49, no. 4 (2013): 5-8, <https://www.journals.ala.org/index.php/ltr/article/view/4269> (accessed 15 February 2021); Steven J. Turner, "Website Statistics 2.0: Using Google Analytics to Measure Library Website Effectiveness," *Technical Services Quarterly* 27, no. 3 (May 2010): 261-78, DOI: <https://doi.org/10.1080/07317131003765910> (accessed 15 February 2021).

15. "7.0 Processing Rules for Underlying COUNTER Reporting Data," COUNTER, <https://www.projectcounter.org/code-of-practice-five-sections/7-processing-rules-underlying-counter-reporting-data/> (accessed 15 February 2021).
16. 'About Events - Analytics Help', Google, 2020, <https://support.google.com/analytics/answer/1033068?hl=en#Anatomy> (accessed 15 February 2021).
17. "Release 5: Understanding Investigations and Requests," COUNTER, 26 May 2017, <https://www.projectcounter.org/release-5-understanding-investigations-and-requests/> (accessed 15 February 2021).
18. Bridget Byrne et al., *Ethnicity, Race and Inequality in the UK: State of the Nation* (Bristol: Policy Press, 2020), <https://library.oapen.org/handle/20.500.12657/22310> (accessed 15 February 2021).
19. Ed, Finn, David Guston, and Jason Scott Robert, *Frankenstein : A New Edition for Scientists and Engineers* (Cambridge, Massachusetts: The MIT Press, 2017), <https://library.oapen.org/handle/20.500.12657/31387> (accessed 15 February 2021).
20. Heike Paul, *The Myths That Made America : An Introduction to American Studies* (transcript Verlag, 2014), <https://library.oapen.org/handle/20.500.12657/31456> (accessed 15 February 2021).
21. Wael Al-Delaimy, Veerabhadran Ramanathan, and Marcelo Sánchez Sorondo, *Health of People, Health of Planet and Our Responsibility: Climate Change, Air Pollution and Health* (Cham: Springer Nature, 2020), <https://library.oapen.org/handle/20.500.12657/37742>, DOI: <https://doi.org/10.1007/978-3-030-31125-4> (accessed 15 February 2021).
22. Haymo Mitschian, *m-Learning - die neue Welle? Mobiles Lernen für Deutsch als Fremdsprache* (Kassel: Kassel University Press, 2010), DOI: https://doi.org/10.26530/OAPEN_355293 (accessed 15 February 2021).
23. OAPEN usage data, DOI: <https://doi.org/10.17026/dans-x9z-4q2w>.