

# التقييم العالمي للمواقع الإلكترونية لقبول عناوين البريد الإلكتروني في عام 2019

بين  
المجموعة التوجيهية لقبول العالمي (UASG)  
رابطة شركات البرمجيات البرازيلية (ABES)  
والحوكمة التمهيدية

9 آب (أغسطس) 2019



## جدول المحتويات

3	المقدمة
4	الملخص التنفيذي
5	فريق البحث
6	المنهجية
6	الإجراءات
9	مخطط البيانات الوصفية
11	النتائج
13	التقييم والترابطات
13	عينة من حالات فشل إجراءات التحقق
13	رفض جميع عناوين البريد الإلكتروني
14	اختبارات الارتباط الإقليمية
15	HTML5: معوقات القبول الشامل
17	هل يمكن أتمتة هذه التقارير؟
18	الخلاصة



## المقدمة

إن الهدف من القبول الشامل (UA) هو ضمان استخدام جميع أسماء النطاقات وعناوين البريد الإلكتروني من قبل جميع التطبيقات والأجهزة والأنظمة المدعومة بالإنترنت. ويتجلى هذا في كل من نطاقات المستوى الأعلى العامة (gTLD) الجديدة وتلك التي لا تحتوي على الحروف اللاتينية. وعلى الرغم من أن البعض قد يفترض أن هذه النطاقات تعمل بنفس الطريقة التي تعمل بها النطاقات القديمة، إلا أن الأمر ليس كذلك، ولا تزال مشاكل التوافق هي الأكثر شيوعاً مما ينبغي.

وهذا البريد الإلكتروني:  
دون@رسيل.السعودية

ويتمثل هدفنا لهذا البريد الإلكتروني:  
测试1@server.technology

للحصول على نفس معدل القبول كما في هذا النطاق:  
user@test.org

كلفت المجموعة التوجيهية للقبول الشامل (UASG) بهذه الدراسة الاستقصائية باعتبارها متابعة لاختبار مماثل تم في عام 2017<sup>1</sup>، وذلك كجزء من مبادرة أوسع لتعزيز فهم المجتمع للاختناقات والمشكلات الرئيسية المحيطة بالتوافق الواسع لجميع أسماء النطاق المتاحة حالياً.

وكان الهدف يتمثل في تقييم امتثال القبول الشامل UA لأفضل 1000 موقع إلكتروني في العالم (وفقاً لـ Alexa<sup>2</sup>) وذلك عن طريق أخذ عينات من الممارسات وأساليب التطوير المختلفة في نطاق البريد الإلكتروني في نماذج على الويب واختبارها بالممارسة العملية. وتعد نماذج البريد الإلكتروني هذه مناسبة تماماً لاختبار الجوانب المتنوعة من UA، حيث يمكن التحقق من نقاط الفشل المختلفة بما في ذلك معايير HTML والتنفيذ، وكذلك استخدام JavaScript ولغات الويب الأخرى على شبكة الإنترنت.

من الواضح أن عدداً من المطورين لا يضعون في اعتبارهم أحدث حالات الاستخدام. ونتيجة لذلك، يحتاج الأشخاص الذين يرغبون في استخدام عناوين مبتكرة أو لديهم القدرة على الكتابة بلغاتهم الخاصة، إلى أن يكون لديهم بريد إلكتروني احتياطي في حالة فشل معالجة عنوانهم المفضل. وهناك ثمة حاجة إلى فهم أين توجد هذه المشكلات حتى يمكن التوصل لحل لها في نهاية المطاف.

وسيقدم باقي هذا التقرير معلومات حول منهجية الاستبيان وتحديد النتائج والمشكلات والحلول المحتملة.

<sup>1</sup><https://uasg.tech/wp-content/uploads/2017/09/UASG-Report-UASG017.pdf>  
<sup>2</sup><https://www.alexa.com/topsites>



## المخلص التنفيذي

بعد تصفية النطاقات الزائفة، أثبت عدد 527 من أفضل 1000 موقع تم تجميعها من قائمة Alexa أنها قابلة للاختبار. استمر الفريق في البحث عن مواقع ويب قابلة للاختبار حتى الترتيب رقم 1922 من القائمة، وتم التوصل إلى ألف نطاق قابل للاختبار عند هذه النقطة. وتم بعد ذلك إجراء التقييمات اليدوية للتحقق من قبول صيغ العناوين التالية والمتزايدة التعقيد بأشكالها التالية:

ascii@idn.ascii test@普遍接受-测试.org	ascii@ascii.newlong test@test.example	ascii@ascii.newshort test@test.exp
arabic.arabic@arabic (RTL) دون@رسيل.السعودية	unicode@idn.idn 测试5@普遍接受-测试.世界	unicode@ascii.ascii 测试1@test.org

تم حفظ النتائج بعد ذلك بقاعدة بيانات MongoDB وذلك لتسهيل الرجوع إليها في مخطط البيانات الوصفية الذي تم إنشاؤه أثناء وضع المنهجية، والنتائج الواردة في التقرير مستمدة من هذا العمل. ومع الأخذ بنظر الاعتبار الفرق في القبول بين اختبارات 2019 و2017، يمكننا ملاحظة نمط مشجع من القبول المتزايد لجزء معين من مهمة القبول الشامل UA، كما هو ملاحظ في أدناه:

حالة الاختبار	2017	2019
ascii@ascii.newshort	%91	%97
ascii@ascii.newlong	%78	%84
ascii@idn.ascii	%45	%50
unicode@ascii.ascii	%14	%13
unicode@idn.idn	%8	%8
arabic.arabic@arabic (RTL)	%8	%7

يمكن تفسير بعض التباين بسبب استخدام مجموعات مختلفة من البيانات، حيث يستخدم كل اختبار قائمة من أفضل مواقع الويب كما كانت موجودة أثناء مرحلة جمع البيانات في كل عملية تقييم. هذا يعني أن معدلات القبول لبعض حالات الاختبار قد لم تنخفض ولكنها لم تتغير. ومع ذلك، فإن الزيادة في قبول النطاقات الجديدة وأسماء النطاقات المدوّلة (IDNs) في المستوى الثاني تُظهر تقدماً لا يمكن إنكاره وهي مسألة ذات أهمية كبيرة.

إن المشكلة الأكثر أهمية التي وجدناها هي `<input type="email">`، وهو حقل مخزون HTML5 لمعالجة هذه البيانات، التي يعتمد عليها عدد كبير من مواقع الويب. والحقيقة هي: أنها ليست متوافقة مع القبول الشامل UA. إننا نؤمن أن الأولوية الأولى لأصحاب المصلحة المعنيين يجب أن تكون تحسين هذا المعيار. وهذا من شأنه زيادة القبول لدى مجلس الإدارة، وخاصة إذا كان يمكن لمواقع الجاهزة للقبول الشامل اختيار استخدام `<input type="eaiemail">` بدلاً عن ذلك، مشيرةً إلى قدرتها على قبول هذه العناوين.



## فريق البحث

**[المشرف]** باولو ميلبيت روك هو خبير مخضرم في مجال التكنولوجيا يحظى بخبرة في التجارة الدولية، والذي تفاوض مع أكثر من 100 شركة في العديد من البلدان المختلفة. وهو المؤسس المشارك ونائب رئيس جمعية البرمجيات البرازيلية (ABES). تأسست ABES في عام 1986، وهي أكثر الكيانات تمثيلاً في هذا القطاع حيث تضم حوالي ألفي شركة منتسبة وتابعة.

**[المنسق]** مارك دبليو داتيسجلد هو سفير مجموعة UASG حاصل على درجة البكالوريوس والماجستير في العلاقات الدولية، ويركز على حوكمة الإنترنت وأثار التكنولوجيا على صنع السياسات العامة والخاصة. وهو يتشاور مع الشركات والأفراد في مشاركتهم في المؤسسات والمحافل الدولية المتعلقة بالتكنولوجيا وذلك تحت العلامة التجارية التمهيدية للإدارة.

**[المختبر الرئيسي]** سافيو فينيشيوس دو موراييس هو أحد أعضاء برنامج ICANN NextGen وسفيره، وهو باحث في مجال الأمن مع التركيز على إنترنت الأشياء في بيئات المكاتب الصغيرة والمنازل (SOHO) بهدف تخفيف تأثير الهجمات الكبيرة على أجهزة إنترنت الأشياء. وهو يحظى بخبرة في العمل ويشارك بنشاط في تطوير الويب وإدارة النظم.

**[المختبر]** إدسون سيليو فيريرا أراوجو هو طالب شاب في هندسة الكمبيوتر يعمل كمطور للأنظمة في شركة Grendene S/A. وهو يساهم في وقت فراغه في مشاريع مفتوحة المصدر ويعمل في مساعي حوكمة الإنترنت. وهو خريج برنامج Youth@IGF.

**[المختبر]** جوناك مينديز فيوريني هو شاب فني في مجال الحوسبة، يدرس حالياً هندسة الكمبيوتر في جامعة فيدرال دو إسبيريتو سانتو (UFES). ويشارك في المشاريع المتعلقة بالشبكات الرقمية وحلول الشبكات، وكان من الشغوفين بالبرمجيات الحرة منذ عام 2012. يعتبر فيوريني خريج برنامج Youth @ IGF.

**الداعمون الكرام لنا وهم:** دون هولاندر، وأجاي داتا، وسرمد حسين، وجنيفر تشونغ، ونيفالدو كليتو، ورودرغو دي لا بارا، ودانييل فينك، وسيدا أكيولوت، وفاندا سكارتيزيني، وروبنز كوهل، وشركة NIC.br.



## المنهجية

### الإجراءات

يأتي أساس منهجيتنا من "UASG017: تقييم مواقع الويب لقبول مجموعة متنوعة من عناوين البريد الإلكتروني"، وهي دراسة رئيسية أجريت في أيلول (سبتمبر) 2017 والتي وضعت الامتثال للقبول الشامل UA على نطاق واسع لأول مرة. وتوضح دراسة 2019 هذه مزيداً من التفاصيل حول المكونات الأساسية للتحقيق السابق بهدف النهوض بتلك الأساليب.

وكان القرار المنهجي المهم الذي تم اتخاذه يتمثل في أنه لم يتم استخدام مجموعة المواقع التي خضعت الى اختبار عام 2017. ففي الوقت الذي تكون هناك قيمة كبيرة في المقارنة بما لو كانت مواقع الويب قد أحرزت تقدماً في معدلات امتثالها أم لا، إلا إنه في الحقيقة إن هذه العملية سوف لن تكون بمثابة تقييم لحالة مواقع الويب الحالية ولكن بدلاً عن ذلك ستكون بشكل إجمالي كدراسة جديدة مختلفة، أي دراسة نافذة.

كانت هناك أيضاً بعض الاختلافات في الإجراءات التي يسترشد بها هذا القرار، منها حقيقة أن مجموعة البيانات النهائية التي خضعت للاختبار السابق تحتوي على حوالي 750 إدخالاً، بينما بلغ العدد 1000 إدخال في اختبار 2019. وسيكون من غير الواقعي إدراج المزيد من الإدخالات بأثر رجعي في مجموعة بيانات 2017، لذلك كان استخدام قاعدة بيانات جديدة هو الخيار الأكثر منطقية في هذه الحالة. وللمضي قدماً، يمكن الآن اتباع أي الخيارين.

كانت الخطوة الأولى المتخذة في هذه الدراسة هي إدراج عناوين URL القابلة للاختبار لأفضل 1000 موقع في العالم وفقاً لأداة التحليل التنافسي Alexa، والتي تم تنفيذها في أوائل عام 2019. كان يجب أن تفي مواقع الويب المختارة بمعايير عدم كونها برمجيات ضارة ولوجود حقل إدخال بريد إلكتروني في مكان ما في صفحاتها. لم يتم وضع أسئلة ذات صلة بالمحتوى وبذلك تتم معاملة جميع مواقع الويب على قدم المساواة على أنها عاملة بغض النظر عن مواضيع محتوى كل منها.

ومن القائمة الأولية لأفضل 1000 موقع ويب، استوفى 527 منها فقط معايير الاختبار، مما دفع الفريق إلى المضي قدماً في القائمة لإيجاد المزيد من المواقع المرشحة، وفي نهاية المطاف تم الوصول إلى العدد 1922 موقع، والذي أصبح الإدخال الأخير في مجموعة البيانات. لقد تعين علينا أن نكون أكثر دقة في تسمية مجموعة البيانات، حيث أنها كانت لتكون قائمة "أفضل 1000 موقع ويب قابل للاختبار وفقاً لـ Alexa".

قام الفريق أيضاً بالبحث عن جنسية مواقع الويب باستخدام سجل WHOIS الذي كان لا يزال محدثاً، وذلك للحصول على انطباع عن مدى الامتداد الجغرافي للاختبار. ويشير هذا إلى حقيقة أن الجزء الكبير من القائمة لا يزال يتركز حول نصف الكرة الغربي، وخاصة الذي تهيمن عليه الولايات المتحدة. وفي نصف الكرة الشرقي، تبرز الصين وروسيا واليابان والهند وكوريا الجنوبية وتايوان وهونغ كونغ باعتبارهم يحظون بالحضور الأقوى.



التصنيف الكامل حسب رمز البلد يظهر في أدناه.

يرجى ملاحظة أن العدد الكلي لا يصل إلى 1000 بسبب عدم القدرة على اكتشاف أصل عدد صغير من مواقع الويب في القائمة.

1	VC	2	AS	1	LV	8	HK	47	CN	3	AE
1	VE	1	SC	2	MA	7	ID	1	CR	1	AM
4	VG	3	SE	1	MU	3	IE	8	CY	5	AR
		2	SG	2	MX	6	IL	11	CZ	1	AT
		1	SI	1	NG	1	IM	29	DE	13	AU
		1	SK	4	NL	16	IN	3	DK	1	BA
		2	TH	2	NO	7	IR	1	DO	1	BD
		1	TN	35	PA	14	IT	1	EC	1	BE
		3	TO	1	PE	24	JP	2	EG	1	BG
		9	TR	3	PH	1	KE	18	ES	13	BR
		12	TW	10	PL	11	KR	26	FR	10	BS
		2	UA	4	PT	2	KZ	9	GB	1	BY
		23	UK	2	RO	1	LA	1	GI	30	CA
		381	US	33	RU	13	LU	3	GR	2	CH

ثم تم إنشاء مجموعة من ستة صناديق بريد من أجل التفاعل مع حقول النموذج لعناوين URL التي تم اختيارها، مع وجود مستويات متزايدة من التعقيد لكل منها فيما يتعلق بالنطاقات القديمة المعتمدة على ASCII. وتم الاحتفاظ بنفس العناوين من بداية الاستبيان وحتى نهايته. وأن قائمة صناديق البريد هي كما يلي:

حالة الاختبار	مثال
ascii@ascii.newshort	test@test.exp
ascii@ascii.newlong	test@test.example
ascii@idn.ascii	test@普遍接受-测试.org
unicode@ascii.ascii	测试1@test.org
unicode@idn.idn	测试5@普遍接受-测试.世界
arabic.arabic@arabic (RTL)	دون@رسيل.السعودية

خلافاً لما في اختبار 2017 ، لم يتم استخدام صندوق بريد "ascii@ascii.idn". في البداية، لم تكن هذه نيّة الفريق، حيث أنه وفقاً لطلبات تقديم التعليقات RFC ذات الصلة، من المفترض أن تكون هذه حالة استخدام مدعومة. ومع ذلك، بعد محاولة تسجيل بريد إلكتروني لحالة الاختبار باستخدام هذه الصيغة على خمس سجلات مختلفة والتي انتهت جميعها إلى حدوث إخفاق بسبب أن الصيغة لم تكن مدعومة، توصلنا مع قيادة مجموعة UASG إلى أنه سيتم استبعاد هذا للاختبار الحالي لأنه لم يتم اعتماده بطريقة جادة في الوقت الراهن.

أحد العناصر المأخوذة من الاختبار السابق هو حالة اختبار "arabic.arabic@arabic"، والتي إذا تم تمثيلها باستخدام تسمياتها الصحيحة، ستكون حالة الاختبار من اليمين إلى اليسار (RTL). لقد عدلنا هذا الاسم الخاطئ إلى حد ما عن



طريق إضافة RTL إلى الرمز، ولكن منعنا عدم وجود اختبار للغات العبرية والفارسية والأردية والسندية والبيديشية ولغات أخرى ذات صلة من نقل الاسم بالكامل. ويتعين على اختبارات أخرى أن تأخذ هذا بنظر الاعتبار.

وكانت الخطوة التالية في العملية هي إجراء اختبار عملي لنماذج البريد الإلكتروني عن طريق إدراج عنوان حالات الاختبار واحداً تلو الآخر وتقديم النماذج المراد معالجتها وفق مواقع الويب. وتم تخطي بعض النماذج في هذه المرحلة بسبب متطلبات التحقق من رسائل SMS أو رموز التحقق captcha غير القابلة للحل على لوحة مفاتيح قياسية غربية. في هذه الحالة الثانية، فإنه على الرغم من أن العدد لم يكن عالياً بما فيه الكفاية لعرض النتائج، إلا أنه من المحتمل قد عمل على إمالة الاستبيان إلى حد ما.

ففي كل مرة تم فيها تنفيذ هذه العملية بنجاح، تم تخصيص ست تصنيفات مقبولة أو مرفوضة للموقع، وفقاً للمعايير التالية:

يؤشر بوضع علامة مقبول في الحالات التالية:

- ✓ اشترك نتج عن رسالة نجاح.
- ✓ قبول الاشتراك ولم يتم الإبلاغ عن أي خطأ.
- ✓ عرض رسالة بريد إلكتروني تقول "أنت مسجل بالفعل"3.

يؤشر بوضع علامة مرفوض في الحالات التالية:

- ✗ يعرض موقع الويب خطأ بمجرد إدخال العنوان.
- ✗ تقديم البيانات يرسل رسالة خطأ.
- ✗ لم يُسمح لإتمام الاشتراك.

كان الجدول الناتج بالصيغة التالية:

الموقع الإلكتروني	البريد الإلكتروني 1	البريد الإلكتروني 2	البريد الإلكتروني 3	البريد الإلكتروني 4	البريد الإلكتروني 5	البريد الإلكتروني 6
test.org	مقبول	مقبول	مقبول	مرفوض	مرفوض	مرفوض
ページ.日本	مقبول	مقبول	مقبول	مقبول	مقبول	مرفوض

أثناء اختبار أول 100 موقع ويب، استغرق الأمر 10 دقائق كمعدل لإكمال اختبار كل موقع. ومع تقدم الاختبار، انخفض معدل الوقت المستغرق إلى خمس دقائق تقريباً لكل موقع ويب، وهو ما نعتبره توقعاً حقيقياً للمحاولات المستقبلية عند تكرار هذه المنهجية من قبل أي أحد.

وأثناء إجراء الاختبارات، تم حفظ رمز HTML الخاص بكل صفحة محلياً حتى يمكن استخراج رموز التحقق من حقل البريد الإلكتروني لتحليلها. استغرقت هذه العملية فترات متباينة، اعتماداً على نوع الحل والتقنية المستخدمة. وإذا تم إجراء التحقق من جانب الخادم أو HTML5، فإنه يمكن تحديد ذلك في غضون دقيقة واحدة أو أقل. ومع ذلك، قد تستغرق تقنيات أخرى مثل JavaScript ما يصل إلى 15 دقيقة لكل موقع ويب، وذلك وعلى الأرجح لأن الرمز كان غالباً مصغراً<sup>4</sup>.

وبدلاً من مجرد نقل هذه النتائج إلى جدول بيانات رقمي، اختار الفريق تنظيمها في قاعدة بيانات MongoDB، وهي عبارة عن منصة موجهة نحو JSON والتي لا تحتاج إلى نمذجة. أثبت هذا الحل أنه قابل للتطوير ومربح، وبإمكان كل جهة تقوم بالاختبار الوصول إلى أحدث إصدار من الاختبارات في أي وقت كان. كان من السهل تصور البيانات وبإمكان

<sup>3</sup> في حالة يكون فيها موقع الويب لديه قاعدة بيانات مشتركة مع قاعدة مسجلة سابقاً.

<sup>4</sup>التعريف: إزالة البيانات غير الضرورية أو المكررة بدون التأثير على كيفية معالجة المورد بواسطة المتصفح، مثل تعليقات الرمز والصياغة وإزالة الرمز غير المستخدم واستخدام أقصر أسماء للمتباينات والوظائف وما إلى ذلك. [Wikipedia]





حتى المستخدم الأقل خبرة تشغيله. وبشكل عام، فقد اعتبرناه خياراً ممتازاً لقاعدة بيانات SQL أو جدول بيانات مشترك عبر الإنترنت.

بمجرد إنشاء قاعدة بيانات رئيسية في MongoDB، فإنها تتيح لإنشاء "مجموعات" غير محدودة، تكون شبيهة بالجدول. تحتوي المجموعات على جميع النتائج ذات الصلة بأي اختبار ختبار معين، وبمجرد تحديد اسمها بشكل صحيح، فإنها ستكون متواجدة كمجموعات بيانات قائمة بذاتها قابلة للتشغيل المتبادل. باتباع التسمية وفق الصيغة "ua\_<scope>\_<year>"، يمكن جمعها جميعاً في نفس قاعدة البيانات المحفوظة بأسماء يمكن للإنسان قراءتها.

على سبيل المثال، تسمى مجموعة هذه الدراسة "ua\_global\_2019"، وهذا يعني أنه إذا تم تسمية المجموعة للسنة التالية باسم "ua\_global\_2020"، فيمكن للقائمين التواجد في نفس المكان دون مشاكل، وسيكون بمقدور الباحثين الرجوع إلى قائمتي البيانات. ومن المهم الإشارة إلى أن أسماء المجموعات تعتبر مسألة حساسة جداً. يمكن تسمية استبيان يجري مستقبلاً في المكسيك مثلاً باسم "ua-regional\_Mexico\_2019"، وما إلى ذلك.

وكانت الخطوة النهائية بعد ذلك تتمثل في تجميع كافة البيانات، وتحويل قاعدة البيانات الناتجة إلى ملف من نوع CSV، وأخذها للمعالجة والتقييم وبذلك يمكن إستخلاص مثل هذا التقرير. يمكن نقل أي ملف من نوع CSV يمثل لسمات مخطط البيانات الوصفية المفصلة أدناه إلى قاعدة البيانات بسهولة باستخدام الأمر التالي، والذي يفترض وجود منظومة برمجية Ubuntu:

```
mongoimport --db ua_database --collection <scope>_<year> --
host=<hostname> --username=<username> --password=<password> --
drop --type csv --file <file_address>/ua_global_2019.csv --
headerline
```

### مخطط البيانات الوصفية

مع اختيار برنامج قاعدة بيانات MongoDB لتخزين البيانات، توصل الفريق إلى مخطط بيانات وصفية نأمل أن يتم استخدامه في الاختبارات المستقبلية، بحيث يمكن لجميع النتائج من عينات مختلفة أن تشترك جميعها في نفس المعيار ومقارنتها حسب الرغبة تحت أي عدسة يختارها الباحثون.

سنقدم الآن هذا المعيار ونشره.

```
{
  "_id": {"$oid": "00001"}
  "domain": "test.org",
  "url":
  "https://www.test.org/signup",
  "rank": "1000",
  "testable": "Yes",
  "code": "<input
type='email'>",
  "comments": "Field triggers
a captcha",
  "mailboxes":
  {
    "mail1": "Accepted",
    "mail2": "Accepted",
    "mail3": "Accepted",
    "mail4": "Rejected",
    "mail5": "Rejected",
    "mail6": "Rejected"
  }
}
```



- فيما يلي وظائف كل سلسلة:
- `id_`: معرف فريد من نوعه تم إنشاؤه تلقائياً داخل قاعدة البيانات بأكملها.
  - `النطاق`: النطاق، غير المسبوق بـ "www".
  - `url`: عنوان URL يحتوي على المسار الكامل للصفحة التي تحتوي على النموذج.
  - `المرتبة`: ترتيب موقع الويب في هذه المجموعة بعينها.
  - `قابل للاختبار`: يشير إلى ما إذا كان الفريق قد تمكن في النهاية من اختبار موقع الويب أو لا.
  - `الرمز`: يحتوي على السلسلة التي تتحقق من صحة النموذج، إن وجد.
  - `التعليقات`: أي تعليق ذي صلة.
  - `صناديق البريد`: قائمة التصنيفات لكل حالة اختبار.

بالنسبة للاختبارات المستقبلية، نحن ننظر في تضمين حقل "eai"، والذي يشير إلى ما إذا كان خادم البريد الخاص بنطاق معين يدعم تدويل عنوان البريد الإلكتروني، مما يعني في الأساس أنه سيكون بإمكانهم تبادل البريد في UTF-8 وبالتالي في الكود الموحد (اليونيكود). ز هذا من شأنه أن يوفر لنا نظرة أكثر عمقاً حول ما هو المستوى العام لدعم هذه التكنولوجيا، مع الأخذ بنظر الاعتبار عدد نقاط الفشل الموجودة في الطريق إلى التبادل الناجح للرسائل من عناوين أكثر تعقيداً.

بالنسبة لتقرير 2019 هذا بشكل خاص، فقد أدرجت سلسلة "البلد" بحيث يمكن تقييم فكرة أفضل حول جغرافية القائمة. ومع ذلك، لا يُمكن أن يعتبر هذا جزءاً من المعيار بسبب التغييرات التي تم إجراؤها على قاعدة بيانات WHOIS والتي تعيق عمليات البحث الضخمة مثل تلك التي أجراها الفريق، مما يجعله غير عملي من الآن فصاعداً. إذا سمح بروتوكول RDAP في النهاية بإجراء مثل هذه الاستعلامات، فقد تصبح السلسلة صالحة مرة أخرى.

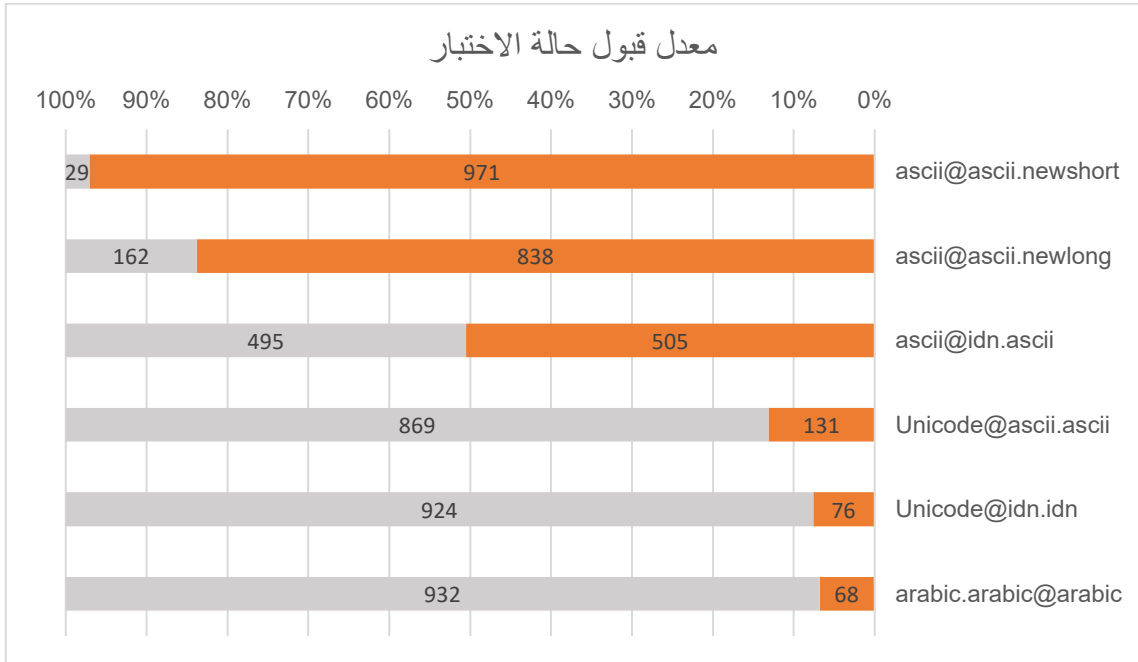


## النتائج

### النتائج الإجمالية للاختبار بالأرقام

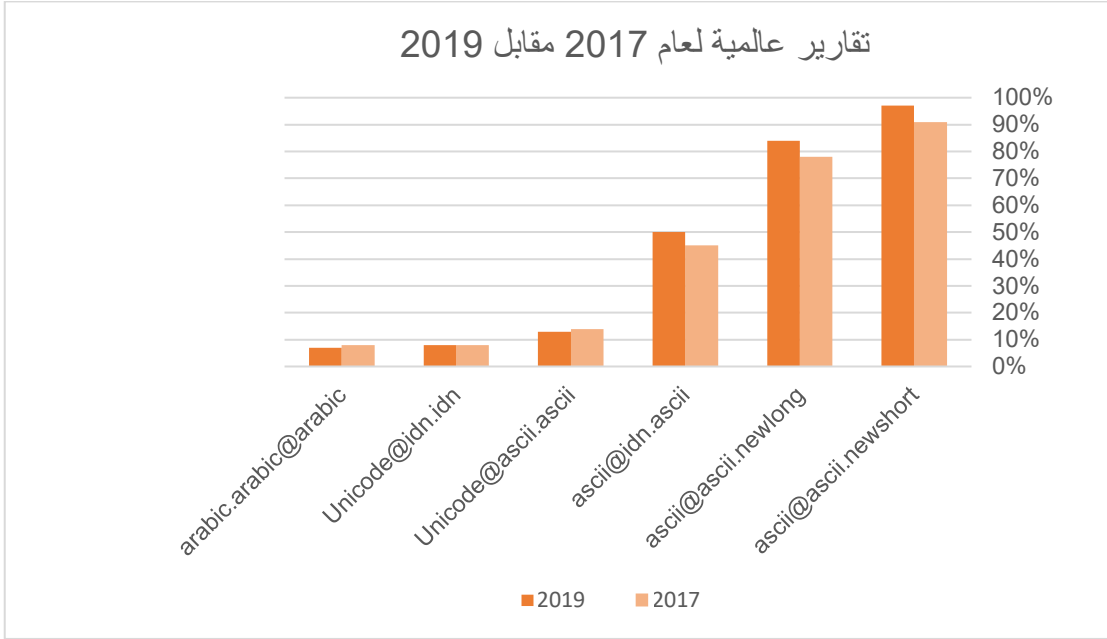
حالة الاختبار	مقبول	مرفوض
ascii@ascii.newshort	971	29
ascii@ascii.newlong	838	162
ascii@idn.ascii	505	495
Unicode@ascii.ascii	131	869
Unicode@idn.idn	76	924
arabic.arabic@arabic (RTL)	68	932

### النتائج الإجمالية للاختبار في الرسم البياني





## مقارنة النتائج الاجمالية للاختبار لعام 2017 مقابل 2019





## التقييم والترابطات

سنستعرض في هذا القسم بعض النتائج التي حصل عليها الفريق أثناء تقييم الرموز، وذلك لإلقاء الضوء على الممارسات السيئة التي تحيط باستخدام نموذج البريد الإلكتروني على الويب. وانهزنا الفرصة أيضاً لإجراء تجربة بسيطة بالاستفادة من جمع البيانات الإقليمية.

وتجدر الإشارة إلى ما يشبه وإلى حد كبير ما لوحظ في استطلاع عام 2017، وهي أنه لا يوجد نهج موحد لترميز وظائف التحقق من الصحة بخلاف ما هو الحال في لغة HTML5 وبعض النصوص القياسية، مثل "jquery-validation.js". يعتمد عدد كبير من مواقع الويب على التعبيرات العادية (RegEx) في لغة JavaScript، مما يحقق درجات متفاوتة من النجاح في استخدامها. ولكن كقاعدة عامة، نادراً ما تعتبر هذه التعبيرات حلاً جيداً.

### عينة من حالات فشل إجراءات التحقق

رفضت مواقع الويب التي تحتوي على كود أفضل بشكل عام ولكنها ضعيفة في الامتثال للقبول الشامل UA البريد الإلكتروني تماماً حينما تم إدخاله في النموذج. وقبلت مواقع الويب ذات الرمز الضعيف العنوان للوهلة الأولى ولكنها أظهرت رسالة "حدث خطأ ما". في حالات قليلة، حدد موقع الويب حالات الاختبار الأكثر تعقيداً على أنها ضارة، وصنفتها على أنها هجمات أو محاولات قرصنة.

فيما يلي بعض الرسائل التي ظهرت بشكل عام، والمتباينات الصغيرة منها، عندما تم رفض حالة اختبار:

يرجى إدخال عنوان بريد إلكتروني نافذ.
صيغة بريد إلكتروني غير صالحة.
عزراً، هذا البريد الإلكتروني لا يبدو صحيحاً. يرجى التحقق من أنه البريد الإلكتروني الصحيح.
يجب ألا يحتوي النص قبل الرمز @ على رموز.
حدث خطأ.

يظهر نوع واحد من الرسائل على وجه الخصوص - وهو أن هناك مشكلة إذا كان الحرف @ مسبوقة بالرموز. كان هذا مطابقاً جداً عبر ملاحظتنا ويبدو أن هذا هو المفهوم الخاطئ الرئيسي الذي يسبب مشاكل فيما يخص التوثيق.

### رفض جميع عناوين البريد الإلكتروني

واجه الفريق بعض الأمثلة السيئة بشكل خاص لممارسات الترميز ضمن أفضل 1000 موقع إلكتروني في العالم. لقد فصلنا فيما يلي مثالاً واحداً في لغة JavaScript والذي فشل في كل حالة اختبار، وهو يأتي من موقع ويب فيتنامي على الأقل. وهي ليست عينة منعزلة أيضاً، ومواقع الويب الأخرى في القائمة ترتكب نفس الخطأ بطرق متنوعة.

RegEx:

```
/^[-a-z0-9~!$%&*_{\'}?]+\.([-a-z0-9~!$%&*_{\'}?)+*@[a-z0-9_][-a-z0-9_]*\.([-a-z09_+)*\.(aero|arpa|biz|com|coop|edu|gov|info|int|mil|museum|name|net|org|pro|travel|mobi|[a-z][a-z])|([0-9]{1,3}\.[0-9]{1,3}\.[0-9]{1,3}\.[0-9]{1,3}))(:[0-9]{1,5})?$/i
```

ما يقوم به هذا الجزء من الرمز عبارة عن اختبار صعب لمعلومات محددة جداً تحدد هيكل عنوان البريد الإلكتروني الرائد، مع الحرص الشديد على إدراج بعض نطاقات المستوى الأعلى (TLD) المدعومة في القائمة البيضاء بالإضافة إلى نطاقات المستوى الأعلى TLD الرائدة المكررة مثل ".org" و ".edu". ليس لهذا معنى كبير، ولكن تم التوصل إلى تباينات من هذا أيضاً في حالات أخرى.



## اختبارات الارتباط الإقليمي

على الرغم من أن هذا ليس هدفاً رئيسياً من الدراسة، إلا أن الفريق أبدى اهتماماً بمحاولة ربط امتثال UA مع أفضل مواقع الويب من البلدان التي تستخدم نصوصاً كتابية مختلفة عن اللاتينية، مع توقع أن مستويات الامتثال قد تكون أعلى. بعد دراسة كل من النصوص السيريلية والهان، استنتجنا أن هذه العلاقة تبدو أنها غير موجودة على هذا المستوى. ربما يمكن للاختبارات المحلية المستقبلية أن تقيم بشكل أفضل ما إذا كان أفضل 100 موقع ويب على المستوى الإقليمي فإن المواقع الإلكترونية تكون أفضل بما يكفي، وهو سؤال مهم يطرح نفسه.

يمكن ملاحظة النتائج أدناه:

- **حالة لغة الهان:** كانت المجموعة تضم 47 موقعاً إلكترونياً من الصين، و8 من هونغ كونغ، و2 من سنغافورة، و12 من تايوان؛ وإجمالاً، لم تتمكن سوى 4 مواقع من الصين من التعامل مع حالة الاختبار المتوسطة "unicode@ascii.ascii".
- **حالة اللغة السيريلية:** كانت المجموعة تضم 37 موقعاً إلكترونياً من روسيا، وموقعين من أوكرانيا؛ وإجمالاً، تمكنت 5 مواقع روسية فقط من التعامل مع حالة الاختبار "unicode@ascii.ascii" المعتدلة.



## HTML5: معوقات القبول الشامل

بالرجوع إلى "ua\_regional\_Brazil\_2018"<sup>5</sup>، أشارت النتائج إلى أن 30 بالمائة من أفضل 50 موقعاً في البلاد قد استخدمت `<input type="email">` كحل لها للتحقق من صحة عناوين البريد الإلكتروني على جانب المتصفح من المعادلة. توقع الفريق أن يجد نمطاً مماثلاً في الاستطلاع العالمي، وقد ثبت صحة ذلك. يتراوح معدل الاستخدام العالمي لهذه السلسلة بين 20 إلى 30 في المائة وفقاً للأنماط التي لاحظناها.

وتتمثل المشكلة في أنه في منتصف عام 2019، لم يكن قد ارتبط المعيار HTML5 بالقبول الشامل بعد. وهذا يجعلنا نشعر بالقلق مع استمرار نمو انتشار المعيار واعتماد المطورين بشكل متزايد على وظائفه لجعل الرمز الخاص بهم يعمل في إطار العديد من السيناريوهات التي تتطلبها الحالة الحالية للويب.

فيما يلي النمط الحالي للقبول لـ `<input type = "email">`:

حالة الاختبار	النتيجة
ascii@ascii.newshort	مقبول
ascii@ascii.newlong	مقبول
ascii@idn.ascii	مقبول
unicode@ascii.ascii	مرفوض
unicode@idn.idn	مرفوض
arabic.arabic@arabic (RTL)	مرفوض

تعزز هذه النتيجة فقط تأكيدنا السابق للمفهوم الخاطئ الرئيسي الذي يبدو أنه يجب ألا توجد رموز قبل الحرف @.

إذا كان ما بين 20 إلى 30 بالمائة من مواقع الويب تستخدم هذه السلسلة، فعلى العكس، فهذا يعني أن نفس عدد مواقع الويب يفشل في هذه المرحلة من التحقق من الصحة بسبب عدم وجود معايير مناسبة في HTML5. بمعنى آخر، فيما يتعلق بجانب المتصفح من المعادلة، فإن هذا هو الإجراء الأكثر أهمية الذي يمكن اتخاذه لتعزيز UA في الوقت الحالي.

بعد التحقق من حالة هذه المشكلة، وجدنا أن المسودة 5.3 من المواصفات، في القسم المتعلق بالتحقق من صحة البريد الإلكتروني<sup>6</sup>، تتوقع الامتثال لـ RFC 6531 و RFC 5890، مما يجعل المعيار متوافقاً مع UA. ومع ذلك، فإن هذه المراجعة كانت في حالة "المسودة الأولى" منذ عام 2017 مع عدم وجود تاريخ متوقع لتصبح "توصية مرشحة" أو "توصية" نهائية.

نحن نجد أن هذه المسألة مهمة جداً نظراً لأهمية الأمر حتى يتمكن المجتمع من الوصول إلى القبول العالمي، خاصة فيما يتعلق بالأسماء المدوّلة. فهذه مشكلة من النوع الذي لا يمكن للمجتمع تركها بدون معالجة، ولا بد من اتخاذ خطوات واسعة في علاقتنا مع بائعي W3C و WHATWG وبائعي المتصفح، وذلك من أجل تسريع عملية التنفيذ لهذا الجانب من المواصفات.

ويعتبر هذا الوقت المناسب خاصة أن W3C و WHATWG وافقا مؤخراً على اتخاذ نهج مشترك نحو وضع المعايير<sup>7</sup>. من المهم الوصول إلى الأطراف ذات الصلة والمطالبة بإضافة خيار "eaimail" لمواقع الويب الجاهزة للقبول الشامل UA، تماماً مثل الضغط على بائعي المتصفحات الرئيسيين للإشارة إلى أنهم مهتمون بجعل هذا الخيار قابلاً للتطبيق.

<sup>5</sup> <http://central.abessoftware.com.br/Content/UploadedFiles/Arquivos/Estudos,%20Pesquisas%20e%20Pareces/Universal-Acceptance-in-Brazil-2018.pdf>

<sup>6</sup> <https://www.w3.org/TR/html53/sec-forms.html#email-state-typeemail>

<sup>7</sup> <https://www.w3.org/blog/2019/05/w3c-and-whatwg-to-work-together-to-advance-the-open-web-platform/>



يوصي فريق البحث بجعل حل هذه المشكلة في مقدمة الأولويات.





## هل يمكن أتمتة هذه التقارير؟

أثناء إعداد هذه الدراسة، طرحت المجموعة التوجيهية للقبول العالمي UASG أسئلة في اجتماعات ICANN وخلال جلسات التخطيط حول جدوى البرمجيات المؤتمتة في إنتاج تقارير مثل هذا التقرير. حاول الفريق إيجاد إجابة عملية على هذا السؤال لكنه توصل إلى استنتاج مفاده "كلا"، لا يمكن أتمتة هذه الاختبارات في الوقت الحاضر.

وهذه هي مخاوفنا الرئيسية:

- هناك الكثير من التنوع فيما يتعلق بكيفية ترميز مواقع الويب، وما هي البرامج النصية التي تستخدمها، وكيف يتم تنظيم النماذج. عمومًا، سيكون من المستحيل عمليًا حساب جميع المتغيرات.
- يُستخدم التصغير على نطاق واسع من أجل تقليل أوقات تحميل المواقع والبرامج النصية. ويتطلب هذا في بعض الأحيان أن ينظر شخص ما لديه خبرة في الترميز في النتائج.
- هناك عدد كبير من مواقع الويب التي تتطلب نوعاً من رموز التحقق Captcha أو لديها بعض الفروق الدقيقة في نماذجها، مما يجعل الأتمتة عديمة الفائدة.

يمكن تطبيق نفس الأسباب المنطقية في مسألة الاستعانة بمصادر خارجية للدراسة لغير الخبراء كما في حالة استخدام موقع Amazon's Mechanical Turk. على الرغم من أن هذا قد ينجح من حيث إجراء الاختبار الفعلي، إلا أن التنسيق سيظل ضرورياً لإعداد قائمة النطاقات، واستخراج الرمز، ومن ثم إصدار تقرير بناءً أكثر من مجرد كونه إحصائياً. ومع ذلك، فهذا هو نهج أكثر كفاءة بكثير من الأتمتة.



## الخلاصة

إذا أخذنا في الاعتبار أفضل 100 موقع في العالم كمرجع، فقد تقبل خمسة منهم فقط جميع حالات الاختبار: "godaddy.com"، و"txxx.com" [NSFW]، و"spotify.com" و"espn.com" و"quora.com" كان يتعين علينا تخطي حالات الاختبار الصعبة مثل "unicode@idn.idn" واختبار العربية من اليمين الى اليسار (RTL)، فسيبقى لدينا تسعة مواقع ويب فقط تتقبل أول أربع حالات للاختبار. ونظراً لأننا نتحدث عن أكثر الصفحات التي يتم الوصول إليها في العالم، فهذا أمر بالغ الأهمية.

فبينما نحن متفائلون بشأن احتمال وصول الشركات الجديدة إلى الامتثال الكامل في المستقبل القريب، ونجد أنه من الواقعي أن تتجه الأجيال الجديدة في نفس الاتجاه، إلا أنه من الواضح أن النطاقات المدوّلة تحتاج إلى دفعة وأن هناك حاجة إلى اتخاذ إجراء استباقي لكي تزدهر. لا يقتضي الأمر وضع سياسات أقوى فحسب، بل يجب جعل الإنترنت أكثر درايةً بوجود مثل هذا النوع من رسائل البريد الإلكتروني والنطاقات - وهو أمر لن يتحقق إلا من خلال التواصل الاستباقي.

في الختام، إننا نؤمن أن الخطوة التالية المهمة ستكون تطوير لوحة معلومات تتيح تقييم جميع التقارير الخاصة بالقبول الشامل UA التي يتم إصدارها باستخدام مخطط البيانات الوصفية المقترح في هذه الوثيقة. وسيؤدي هذا إلى إنشاء مستودع طويل الأمد يسمح بإجراء قياسات واقعية من قبل أي طرف مهتم وسيكون بمثابة أداة مرجعية للمجتمع ككل.