اجتماع ICANN78 | الجمعية العمومية السنوية – التخطيط لمراجعة تشابه السلاسل في الجولة التالية من نطاقات gTLD
الخميس الموافق 26 أكتوبر/تشرين الأول 2023 - من الساعة 9:00 إلى 10:00 بتوقيت هامبورغ

بيتينان كوارمورنباتانا: شكرًا. أهلاً ومرحبًا بكم في جلسة التخطيط لمراجعة تشابه السلاسل في جلسة الجولة التالية من نطاقات gTLD. أنا محدثتكم بيتينان وأنا مديرة المشاركة لهذه الجلسة. يُرجى العلم بأن هذه الجلسة يجري تسجيلها، وتحكُمُها معايير السلوك المتوقعة في ICANN. وتتضمن هذه الجلسة ترجمة فورية وتدوينًا آنيًا للحوار في الوقت الفعلي، وهو ما يمكن عرضه بالنقر على الأيقونات الموجودة بشريط أدوات برنامج زووم. وإذا أردت التحدث، فيرجى رفع يدك وسوف ينادي عليك منسق الجلسة. أود يمكنك كتابة أسئلتك في مربع الدردشة. وجمع المشاركين مطالبين بالتكرم بذكر أسمائهم قبل التسجيل وأن يتحدثوا بوتيرة معقولة. وبذلك، أحيل الكلمة إليك، سرمد. شكرًا.

سرمد حسين: شكرًا لك، بيتينان، وطاب صباحكم جميعًا هنا وطابت أوقات كل من يشاركوننا عبر الإنترنت. ربما تكون هذه الجلسة الخاصة هي الأولى في سلسلة من الجلسات، والتي سوف تركز على كيفية تخطيطنا لمراجعة تشابه السلاسل في الجولة التالية من نطاقات gTLD. ويشاركني هنا أسموس فرايتاغ، والذي سيتحدث أكثر حول بعض الإرشادات التوجيهية التي كنا نطورها مع ميشيل ديسميتر، فهما يتشاوران معنا في هذا المشروع. وقبل أن نصل إلى إلقاء نظرة على العرض التوضيحي نفسه، فقد رأيت أنه قد يكون من المفيد أن أعرض عليكم أولاً قدرًا من السياق ومعلومات الخلفية حول ما نقصده بلفظ مثل ومماثل ومميز لأن هذه من الأشياء التي سوف نتحدث حولها خلال حوارنا اليوم.

ومن ثم عندما نتحدث حول السلاسل أو مقاطع نطاق المستوى الأعلى، المقدم طلبات للحصول عليها أو التي تكون قيد الاستخدام، فهي مؤلفة بالفعل من حروف ونقاط أكواد بنظام يونيكود ويمكن بشكل افتراضي أن تكون مماثلة أو شبيهة أو متميزة عن بعضها بعضًا. وعندما تكون نقاط الأكواد مجمعة معًا أو عندما تكون هذه الحروف مجمعة معًا من أجل تشكيل سلاسل أكبر، فمن الممكن اعتبار السلاسل المقابلة لها أيضًا مشابهة أو مماثلة أو متميزة افتراضيًا. إذن من الممكن أن يكون التماثل في طبقتين، الحرف في حد ذاته، والذي تتألف منه السلسلة وفي مستوى

ملاحظة: ما يلي هو ما تم الحصول عليه من تدوين ما ورد في ملف صوتي وتحويله إلى ملف كتابي/نصّي. ورغم أن تدوين النصوص يتمتع بدقة عالية، إلا إنه في بعض الحالات قد تكون غير مكتملة أو غير دقيقة بسبب المقاطع غير المسموعة والتصحيحات النحوية. تنشر هذه الملفات لتكون بمثابة مصادر مساعدة للملفات الصوتية الأصلية، ولكن لا ينبغي أن تُعامل معاملة السجلات الرسمية.

السلسلة وفي بعض النصوص، ومن الملفت، عندما نقوم بتجميع الحروف معًا في سلاسل فإنها في حقيقة الأمر تغير من شكلها.

ومن ثم فإن النظر فقط في الحروف الفردية غير كافٍ، على الرغم من توفيره لبعض المعلومات، ولكن في نهاية المطاف بالنسبة لمراجعة تشابه السلاسل، فقد يحتاج المرء أيضًا أن ينظر في كامل السلسلة. ولكن فقط لكي أقدم لكم بعض معلومات الخلفية، إليكم بعض الأمثلة على الحروف ونقط أكواد يونيكود الخاصة به من نصوص مختلفة لكي أوضح لكم بعض الفئات. والسطر الأول هنا، وهو المتشابه من الناحية البصرية، يمثل بشكل أساسي تلك الحروف التي يستخدمها هؤلاء المستخدمين لذلك النص أو المجتمع الذي يستخدم ذلك النص ويرون أنها مشابهة أو لا يمكن تفريقها عن بعضها الآخر. وفي بعض الحالات، قد تكون مختلفة قليلاً، لكن ليس من السهل التمييز بينها. وفي بعض الحالات، تكون متطابقة تمامًا، بمعنى أنها في حقيقة الأمر بنفس نوع الخط المستخدم وتستخدم نفس رموز الحروف أو الصورة في تشفير ذلك الحرف.

وإذا ما نظرنا في المثال اللاتيني أو باللغة السيريلية، نجد أن الحرف C والحرف السيريلي الذي يضم نقطة رمز رقم 0441، فإنما متطابقين تمامًا. وبالمثل بالنسبة للغة العربية، ومن ثم فهذا التشابه يمكن أن يكون عبر النصوص، ولكنه قد يكون كذلك ضمن النص الواحد. فإذا ما نظرتم في المثال العربي، سوف ترون أن الحرف الممثل أو المشفر باستخدام يونيكود في 0649 و06CC، أنهما متطابقين أيضًا. على الأقل في هذا السياق. ومن ثم فإن تلكم أمثلة على ما نطلق عليه اسم متشابه بصريًا.

إذن عملية تشابه السلسلة، فإننا في حقيقة الأمر ننظر فقط في التفسير البصري وليس أي تفسير آخر. فالحروف المتشابهة بصريًا، فإنها محددة على ذلك النحو من خلال مجتمعات النصوص التي كنت أعمل فيها وهي قواعد إنشاء مسميات منطقة الجذر. إذن من الناحية الأساسية، فإن هذا يعطينا مثالاً على حرفين من اللاتينية واليونانية، وهما ما يمكن اعتبارهما متماثلين.

وبعد ذلك مثال آخر من الديفنغارية والبنغالية، وهما ما يمكن اعتبارهما متشابهين في حقيقة الأمر. وبعد ذلك حروف أخرى، يوضح السطر الأخير أشياء، أو حروف مختلفة تمامًا عن بعضها بعضًا. على سبيل المثال، الحرف A والحرف F في الإنجليزية أو مجرد حروف اعتباطية أخرى في نصوص ولغات أخرى. إذن فإن السطر الأول الذي يمثل الحروف المتشابهة أو التي لا يمكن تمييزها تم تحديدها بالفعل من خلال مجتمعات النصوص وتم ترميزها بالفعل بشكل

واضح بأنها متباينات في قواعد إنشاء مسميات منطقة الجذر. ومن ثم أعتقد أن هذا الجزء تم تشفيره بالفعل وتمت معالجته.

ماذا عن هذا الجزء من العمل، ما يتعلق بمراجعة تشابه السلاسل، يرتبط على الأكثر بهذا السطر الأوسط حيث توجد حروف متشابهة من الناحية البصرية وبعد ذلك تكون السلاسل المكونة بتلك الحروف المتشابهة بصريًا. وبعد ذلك يكون الأمر عائدًا إلى هيئة المراجعة المستقلة والتي سيتم إعدادها من أجل الجولة التالية، والتي ستقرر بعد ذلك ما إن كانت سلستان تقعان في هذه الفئة ويكونان متشابهين بصريًا، وهل هما متشابهين لحد اللبس أم لا.

فإذا كانا متشابهين لحد اللبس، فيجب ألا تتم المتابعة فيهما لأن ذلك يؤدي إلى ليس في نطاقات المستوى الأعلى التي يتم تفويضها وهذا ليس أمرًا جيدًا بالنسبة للمستخدمين النهائيين. لكن إذا كانا متشابهين ولكن يمكن التفريق بينهما رغم ذلك، فيمكن المضي قدمًا فيهما. وذلك إلى حد ما تفريق أدق يجب القيام به من خلال الهيئة المعنية بمراجعة بتشابه السلاسل وهذا ما يخص عملية مراجعة تشابه السلسلة. إذن فإن ما نعمل عليه في نهاية المطاف هو هذا السطر الأوسط ومحاولة كشف التفاصيل لمعرفة ما هي الآلية الجيدة لجعل هذه العملية أكثر قوة.

فلندخل إلى العرض التوضيحي نفسه. السلسلة، إن ما سنقوم به في هذه الجلسة هو قبل أن ندخل في حقيقة الأمر في المزيد من التفاصيل، سوف نقوم فقط بمراجعة بعض توصيات الخلفية التي تأتينا من السياسة التي تحفز على هذا العمل. وتأتي مجموعة التوصيات الأولى من التقرير النهائي حول عملية وضع سياسات الإجراءات اللاحقة لنطاقات gTLD الجديدة أو الإجراءات اللاحقة اختصارًا. ثم بالإضافة إلى ذلك، هناك المزيد من التفاصيل التي تضاف إلى تشابه السلسلة. في توصيات المرحلة الأولى من عملية وضع السياسات المعجّلة لأسماء النطاقات المدوَّلة أو IDN EPDP اختصارًا. برجاء ملاحظة أن عملية وضع السياسات المعجّلة لأسماء النطاقات المدوَّلة منشورة في الوقت الحالي. وهي توصيات المرحلة الثانية للتعليق العام. وما نشاركه هنا هو التوصيات التي تم نشرها خلال التعليقات العامة.

ومن ثم هذه ليست هي التوصيات النهائية من عملية وضع السياسات المعجَّلة. فقد حصلوا على التعليقات والآراء من خلال عملية التعليق العام وهم الآن بصدد مراجعة بعض من تلك التوصيات. لكن تلك التوصيات النهائية لم يتم نشرها إلى الآن. إذن فبعض التفاصيل التي نشاركها حول عملية وضع السياسات المعجّلة لأسماء النطاقات المدوَّلة قد تتغير بمجرد نشر التوصيات النهائية.

ولغرض هذا العرض التوضيحي، فإننا نستخدم الإصدار الذي تم نشره بالفعل من أجل نسخة التعليقات العامة إلى أن يصبح الإصدار التالي متاحًا. وبذلك، سوف نراجع بعد ذلك تلك التوصيات ونقترح أو نشارك استنادًا إلى تلك التوصيات المنهجية المستندة إلى النطاق ونتائج عملية مراجعة تشابه السلاسل للجولة التالية من نطاقات gTLD.

ومن بين الأشياء -- كما سترون هذه التوصيات فإنها تنادي بإعداد إرشادات توجيهية لمراجعة تشابه السلاسل من أجل تنفيذها عبر لجنة مراجعة تشابه السلسلة والتي سوف تقوم بهذا العمل. ومن ثم فقد دعونا أسموس الذي سيساعدنا على وضع هذه الإرشادات التوجيهية التي ستعود بالنفع في نهاية المطاف على هذه العملية. وهذه الإرشادات التوجيهية ليست نهائية في الوقت الحالي فنحن ما زلنا في بداية وضع وتطوير هذه الإرشادات. إذن فإن ما سنقوم به هو أن أسموس سوف يطلعنا على عملية التفكير الحالية حول المكان الذي سننطلق إليه في هذه المسألة وسوف نطلب بشكل واضح تعقيباتكم وآرائكم هنا لمعرفة ما إن كان هذا هو الاتجاه الصحيح الذي يجب أن نسلكه وما هي الاعتبارات الأخرى التي قد نرغب في اتخاذها في عملية تطوير ووضع هذه الإرشادات التوجيهية. إذن فهذه عبارة عن مراجعة بشكل ما للجلسة واسمحوا لي بالدخول في الموضوع مباشرة.

إذن بداية من توصيات السياسة بالأساس القسم 24 في الإجراءات اللاحقة فإنها تتحدث أو تركز على مراجعة تشابه السلسلة ومجموعة العمل بالطبع تؤكد على التوصية 2 من سياسة 2007 والتي تنص على أن السلاسل يجب ألا تكون مشابهة إلى حد اللبس مع نطاق مستوى أعلى حالي أو اسم محجوز. إذن فهذا إلى حد ما أساس للمكان الذي يمكننا البدء منه. التأكيد رقم 24.2 -- إذن هذا أيضًا ينبني على جولة 2012 لنطاقات gTLD والتي توفر بعض التفاصيل الإضافية حول طبيعة نطاق تلك المقارنات لمراجعة السلاسل تلك ولكن الأهم من ذلك أنه تقدم لن تعريفًا بشكل ما لما تعنيه كلمة متشابه إلى حد اللبس بالنسبة للسلاسل وإذا ما قرأتم تلك العبارة فإنها تقوم بأن كلمة مشابه تعني السلاسل ولذلك فإن كلمة مشابه تعني أن السلاسل متشابهة إلى حد التسبب في إمكانية إرباك المستخدم في حالة تفويض أكثر من سلسلة واحدة في منطقة الجذر.

إذن من الناحية الأساسية، فقد عُهد إلى الهيئة المعنية بتشابه السلسلة في نهاية المطاف بتقرير وتحديد ما يميز بين مجرد التشابه البسيط في مقابل ذلك التشابه الذي سوف يركز على المستخدمين النهائيين بقدر من الإمكانية وذلك النوع من التفاصيل التي سوف يتوجب استنتاجها في نهاية

المطاف. وهناك بعض التوصيات الإضافية التي تتناول أداة تلقائية قد لا تكون كافية للقيام بذلك والمعالجة اليدوية بحاجة إلى بعض التفاصيل فيما يخص عملية الاعتراض.

وبعد ذلك في توصيات المرحلة الأولى من عملية وضع السياسات المعجَّلة لأسماء النطاقات المدوَّلة، فإن هناك المزيد من التفاصيل التي تضاف من أجل مراجعة تشابه السلسلة. تُبيّن التوصية رقم 4.1 بوضوح أن نطاق المقارنات للسلاسل المتقدم بطلبات للحصول عليها يجب أن تقارن مع نطاقات gTLD الحالية، ونطاقات ccTLD الحالية، والسلاسل المطلوبة لتكون نطاقات المستوى الأعلى لرموز البلدان ذات أسماء النطاقات المدوَّلة وغيرها من نطاقات gTLD المقدم طلبات لها، وجميع السلاسل في قائمة الأسماء المحجوزة وأي حرفين من نوع ASCII تم حجزهما من أجل أكواد الدول.

والأمر لا يتعلق في حقيقة الأمر بالقوائم وحسب ولكنه يرجح بأنه ليست السلسلة الأصلية وحسب ولكن هذه السلاسل سوف تتم مقارنتها إلا أن المتباينات المخصصة لها سوف تتم مقارنتها مع المتباينات المخصصة ومتباينات الكتل. وبعد ذلك متباينات الكتل أيضًا سوف تتم مقارنتها مع متباينات الكتل الأخرى. وسوف نتطرق إلى تلك التفاصيل لاحقًا أيضًا، لكن هذا في حقيقة الأمر يوفر لنا الكثير من تفاصيل ما قد تبدو عليه تلك المقارنات. ومن ثم فإن هذه المحادثات حول نطاق المقارنات في تلك التوصية.

وبعد ذلك فإنه يتطرق في حقيقة الأمر إلى القول بأنها تدرك أنه نطاق المقارنة قد يصبح كبيرًا للغاية في بعض الأحيان، لأن لديك سلاسل مقارنة إضافة إلى المتباينات الخاصة بها مع السلاسل الأخرى ومتباياناتها، وفي بعض الأحيان تتحول تلك المجموعات -- قد يكون لأي سلسلة في حقيقة الأمر مئات المتغيرات المختلفة. ومن ثم فإنكم تقارنون مئات المتباينات بمئات المتباينات للسلسلة ولسلاسل حالية.

ومن ثم في حال أصبح البعض من هذه المقارنات كبيرًا وقد تكون من الواضح جدًا أن البعض من هذه المقارنات من الممكن ألا تؤدي في حقيقة الأمر إلى أي مشكلات تخص تشابه السلاسل، وبعد ذلك يمكن أن تقوم الهيئة المعنية بتشابه السلسلة بحذف بعض من تلك المقارنات. لكن هذه المقارنات، ومن ثم فإن هذه المقارنات تتيح قدرًا من الكفاءة في العملية والإجراءات، ولكن إلى الحد الذي -- إذا كنت تقوم على سبيل المثال، بمقارنة الصينية بالعربية، فإن احتمالية التشابه منخفضة للغاية بين هذين الزوجين من النصوص.

لكن إذا كنت تقارن اللغة السيريلية باللاتينية، فإن احتمالية التشابه قد تكون مرتفعة جدًا. ولكن لا يجب تخفيضها اعتباطيًا وحسب وبعض الإرشادات التوجيهية التفصيلية للغاية يجب وضعها حول ما يمكن أن يكون، حسنًا، فهي تقول بالفعل ما يجب مقارنته، لكن تلك الإرشادات التوجيهية والمعايير يمكن أن تحدد بعد ذلك ما يمكن تجاوزه ربما حيث تتوافر فرص منخفضة للغاية لحدوث تشابه عبر مختلف أزواج النصوص. ومن ثم فإن هذه الإرشادات الخاصة تتيح إمكانية حدوث تلك الاستثناءات، ولكن من خلال الإرشادات التوجيهية، والتي يجب أن توضح بطبيعة الحال مسبقًا.

وفي نهاية المطاف، فإن هذه التوصية في حقيقة الأمر تقول بأنه إذا كانت دليك مشكلة تتعلق بالتشابه في حقيقة الأمر مع أي واحد من السلاسل في مجموعة متباينات، فسوف تتأثر مجموعة المتباينات بالكامل، والأمر لا يتعلق بتلك السلسلة وحسب. ومن ثم إذا ما تبين أن سلسلة واحدة متشابهة مع سلسلة أخرى في أي من نطاقات gTLD المقدم لها طلب في الجولة، على سبيل المثال، فإن المتباين الكامل لتلك السلسلة ومجموعة المتباينات الخاصة بها سوف تدخل في تنافس مع مجموعة المتباينات الكلية في السلسلة الأخرى، وليس في تلك السلسلتين، على سبيل المثال. وبالإضافة إلى ذلك، في حالة تطابق المتباينات بشكل مباشر، حتى وإن لم تكن متماثلة، فإنها تدخل أيضًا في مجموعة التنافس. إذن هذا هو سياق السياسة إلى حد ما، أعني السياسة الإرشادية، والتي تطالبنا بأن نحدد منهجية للنطاق بالإضافة إلى النتائج. إذن نعود مرة أخرى إلى النطاق، كما يمكن أن تروا بعد ذلك، ما الذي يجري اقتراحه في المقدمة هنا هو سلسلة gTLD المقدم طلب للحصول عليها.

وهي السلسلة الأساسية في حد ذاتها، أعني السلسلة المقدم طلب للحصول عليها، وبعد ذلك من الممكن الحصول على بعض مقاطع الأسماء المتباينة والقابلة للتخصيص وسوف يكون بها مقاطع تسمية متباينة تسلسلية. ويجري احتساب هذه الأشياء من خلال قواعد إنشاء مسميات منطقة الجذر. وبالمثل، وعلى الجانب الآخر، وعلى النقيض من ذلك، سوف تجدون فئات المقارنة التي تم تحديدها في عملية وضع السياسات المعجّلة لأسماء النطاقات المدوَّلة. وهي تشمل نطاقات gTLD الموجودة أو نطاقات gTLD التي تم التقدم بطلبات للحصول عليها على سبيل المثال، لكنها ما تزال قيد التناول من الجولة السابقة.

ومن ثم ما تزال هناك سلاسل ربما لم تتم معالجتها على سبيل المثال، من جولة نطاقات gTLD لعام 2012. ومن ثم فإنها أيضًا جزء من هذه العملية. ونطاقات ccTLD الموجودة حاليًا أو

نطاقات المستوى الأعلى ذات رموز البلدان أو غيرها من نطاقات المستوى الأعلى لرموز البلدان الموجودة في عملية تقديم الطلبات. ومن ثم فإنها مطلوبة، لكنها غير مفوضة تكرمًا أو تم تقييمها بنجاح. أما سلاسل نطاقات gTLD الأخرى المقدم لها طلبات في نفس الجولة، والأسماء المحجوزة وأي نطاق ASCII مكون من حرفين لأن نطاقات ASCII ذات الحرفين محجوزة لنطاقات المستوى الأعلى ذات رموز البلدان ومدارة بمعرفة وكالة إدارة بموجب المعيار ISO 3166.

إذن وبشكل أساسي، من الممكن أن تكون ذات سلسلة أساسية، ومتغيرات قابلة للتخصيص ومتباينات سلسلة وفقًا لما تم حسابه عن طريق قواعد إنشاء مسميات منطقة الجذر. وجميع تلك السلاسل بحاجة إلى مقارنة مع جميع السلاسل المقدم طلبات للحصول عليها ومقاطع الأسماء المتباينة سوف تكون بحاجة إلى مقارنة معها جميعًا. إن كلمة "نعم" أمامكم تشير إلى المكان الواجب إجراء مقارنات فيها. أما "نعم" مع علامة النجمة أو تلك المناطق الرمادية غير النشطة قليلاً، فإنها تسرد المقارنات التي قد تقرر الهيئة حذفها في حالة -على سبيل المثال- وفقًا لم أعلنته سابقًا، إذا كنتم تقارنون الصينية بالعربية، فقد تقرر اللجنة بأن فرص حدوث التشابه منخفضة للغاية ولذلك من الممكن أن تقرر بالنسبة لبعض السلاسل الخاصة أن تزيل تلك المقارنات. وتلكم هي المقارنات المميزة بعلامة "نعم" مع نجمة. إذن كما ترون، لا يمكن للهيئات إلى الآن إزالة جميع المقارنات. فهناك بعض المقارنات التي يمكن حذفها فقط.

وبعد ذلك علامة "لا" وهي ترمز إلى المنطقة التي لا تقع في نطاق مقارنة مراجعة تشابه السلسلة. ومن ثم متباينات حجب في مقابل متباينات حجب لن تخضع للمقارنة في العملية. هذا ملخص لما تم النص عليها في حقيقة الأمر في السياسة 4.1، أعني التوصية 4.1 من عملية وضع السياسات المعجّلة لأسماء النطاقات المدوَّلة. والمنهجية تتمثل بالطبع في أننا في نقطة البداية، أنه يتوجب علينا في حقيقة الأمر تجميع قائمة بجميع هذه السلاسل التي يجب أن تخضع للمقارنة. وهي بحاجة لأن تأتي من جميع الفئات المختلفة. ونحن أيضًا بحاجة إلى تجميع قائمة تضم المتباينات القابلة للتخصيص وأيضًا ربما متباينات الحجب.

وفي بعض الحالات، قد تكون قائمة متباينات الحجب كبيرة لسلسلة خاصة، وليس بشكل عام، ولكن لسلسلة خاصة، فقد تكون كبيرة في حقيقة الأمر، وفي هذه الحالة سوف يتوجب علينا التوصل إلى آلية مجدية أكثر من أجل القيام بتلك الأنواع من المقارنات. وسوف يتم ترك هذه التفاصيل بالطبع إلى حد ما في الإرشادات التوجيهية والبعض منها إلى اللجنة نفسها. وبعد ذلك

بمجرد تحديد مجموعة السلاسل، فسوف تقرر اللجنة ما هي سلاسل الحجب التي يجب حذفها أو ما هي المقارنات التي يجب حذفها، إن وجدت، مع توثيق مسوغات ذلك القرار. وبعد ذلك سوف يحددون الطلبات المختلفة ذات نفس مجموعة السلاسل المتباينة من أجل التنافس. فإذا ما كانت هناك سلاسل متباينة عن بعضها الآخر، أي في طلبين مختلفين، فيجب أن يدخلا بشكل واضح في التنافس.

ومن ثم سوف يدخلان في التنافس بسبب ذلك. أو أن هناك سلستان متشابهتان، ومن الواضح أنهما يدخلان في التنافس أيضًا. وبعد ذلك بالنسبة للبقية، سوف يجريان مقارنة السلاسل من أجل التأكد من تشابه السلسلة إلى حد تشابههما إلى حد اللبس. وهذا التعريف الخاص بماهية التشابه إلى حد الإرباك سوف يأتي بشكل واضح من إرشادات مراجعة تشابه السلسلة التي يجري وضعها وتطويرها في الوقت الحالي. وبعد ذلك استنادًا إلى ذلك، سوف تضع تعريفًا لمجموعات التنافس، وتوثق مجموعات التنافس والمسوغات الخاصة بمراجعة تشابه السلاسل. وسوف يعود ذلك مرة أخرى من الهيئة إلى ICANN وسوف يتم نشر النتائج. ومن ثم فإن هذا إلى حد ما وجهة نظر رفيعة المستوى للغاية للعملية التي سوف يتم اتباعها واستنادًا إلى الإرشادات التوجيهية المقدمة إلينا، إضافة إلى التوصيات والإرشادات.

وبالطبع فإن هناك العديد من النتائج المحتملة. وسوف تقوم هيئة مراجعة السلاسل بتوثيق النتائج التالية. فقد تكون هناك سلسلة مقدم لها طلب، وتكون مشابهة لأحد نطاقات gTLD الموجودة حاليًا وبحاجة إلى توثيق. وفي تلك الحالة بطبيعة الحال، فقد لا يكون من الممكن المتابعة فيها على سبيل المثال. وسوف نتحدث حول ذلك في الشريحة التالية. حيث يمكنهم توثيق وجود سلسلة مشابهة لسلسلة نطاق gTLD مقدم طلب للحصول عليها في الجولة السابقة أو مشابه لنطاق ccTLD موجود حاليًا أو مشابه لنطاق مدول مطلوب أو مشابه لاسم محجوز أو مشابه لسلاسل مكونة من حرفين. أو ما هو غير مدرج هنا عندما يكون ذلك مشابهًا لواحد من هذه السلاسل، فهذا يعني في حقيقة الأمر أنه مشابه أو متباين قابل للتخصيص مشابه أو أنه متباين قابل للتخصيص أو متباين حجب. ومن ثم فإن جميع الخيارات التي ناقشناها في الشريحة السابقة، فإذا ثبت أي من تلك، فسوف تتأتى هذه النتيجة.

وأخيرًا، فإنهم يجدون أن السلسلة غير مشابهة لأي من الفئات التي تم سردها وفي تلك الحال من الممكن المضي قدمًا في الموافقة على السلسلة. ويمكنني نشر نتائج المراجعات المتعددة للسلاسل أو موقعها على الويب لأي أحد من أجل المراجعة في إطار العملية التي يجري تطويرها. ومن

ثم سوف يتطرق هذا الأمر إلى المزيد من التفاصيل. وأنا لن أخوض في التفاصيل وسوف أقرأ عليكم كل شيء، لكن كان لتلك النتائج في نهاية المطاف بعض العواقب.

إذن لكي نتناول مثال بسيط فقط الموجود في أعلى جهة اليسار، في الصف الأول، في حال تبين أن نطاق gTLD المتباين المقدم طلب للحصول عليه أو أي من السلاسل المتباينة له هو نفس نطاق gTLD موجود، فلا يمكن المتابعة فيه. فإذا ما تبين أنه متباين لسلسلة نطاق gTLD حالي، فلا يمكن المتابعة فيه إلا عندما يكون بالطبع متغيرًا قابلاً للتخصيص وأن يكون هناك طلب مقدم له لنفس مشغل السجل وفي تلك الحالة يكون طلب متباين ومن ثم يمكن في تلك الحال المتابعة فيه.

ومن ثم إذا كان متشابهًا من الناحية البصرية، لكنه ليس متباينًا لأي نطاق gTLD موجود، فلا يمكن المتابعة فيه. ويمكنكم استعراض الشرائح ومعرفة مجموعة العواقب الأخرى أو نتيجة هيئة مراجعة تشابه السلسلة. وهذا بالطبع سوف يرشدنا بعد ذلك إلى وضع العملية نفسها. وبذلك، سوف أحيل الكلمة إلى أسموس والذي سوف يستعرض معكم الإرشادات التوجيهية في حد ذاتها والتي يجري تطويرها في الوقت الحالي، والكلمة إليك أسموس.

أسموس فرايتاغ: حسنًا، شكرًا لك، سرمد. لقد قمت بتغطية كل ما كنت أنوي قوله تقريبًا. فلننتقل إلى الشريحة التالية. واسمحوا لنا أن ننتقل إلى الشريحة التالية. ونريد فقط أن نؤكد على أن إمكانية حدوث اللبس هو الدرجة التي يمكن فيها أن تحل تسميتان محل بعضهما بعضًا. وكلمة تحل محل هي المفهوم الأساسي هنا. وكلمة قابل للبس، نعتقد أن هناك عتبة ما موجودة وهي في حقيقة الأمر غير محددة التعريف بموضوعية وبشكل متكرر، لكن يتعين علينا اختيار أحدها بحيث نقول بأن إمكانية حدوث اللبس قد وصلت إلى مستوى يكون مقلقًا بالنسبة لمنطقة الجذر. وهي مهمة بالنسبة لمنطقة الجذر لأن إمكانية حدوث اللبس موجودة بشكل واضح في طيف دقيق بداية من إمكانية لبس قليلة جدًا إلى كبيرة جدًا.

لننتقل إلى الشريحة التالية، رجاءً. تشابه السلاسل هو مجموعة فرعية من إمكانية اللبس التي تستند إلى التشابه في الشكل. ومن ثم، قد يكون لديكم مفهوم التشابه المسبب للبس، وهو درجة التشابه التي يجب عليكم فيها جعل المسميات فيها مسببة للبس، بمعني تجاوز ذلك الحد بصرف

النظر عن العتبة المحددة والتي تفترضون تجاوزها وهو الحد الذي يكون عنده حد اللبس غير محتمل بعد ذلك.

الشريحة التالية من فضلك. إن الإدراك الأساسي يتمثل في أن من الصعب للغاية أن نضع معيارًا ثابتًا للغاية لمدى اللبس لأن ما يبدو مسببًا للبس يختلف بالنسبة لكل مستخدم، لأنه أيضًا مختلف بالنسبة للسياق الذي يتم فيه عرض المسميات. إن من هم ملمون أو على دراية بنص أو لغة محدد في نص محدد اجتازوا سنوات وسنوات من قراءة سلاسل بالنص الخاص بهم ولديهم طريقة محددة في النظر إليها قد لا تكون لمن هم ملمون وعلى دراية بلغة مختلفة. إذن فإن أحد الأشياء في إنشاء ووضع إرشادات توجيهية هو ضرورة تحديد طبيعة المستخدم المستهدف.

لننتقل إلى الشريحة التالية، رجاءً. إذن أول الطرق المختصرة التي سوف نقترحها هي أن المستخدم الذي نستهدفه هو ما نطلق عليه مهتم بشكل معقول باحتياجات العملاء. وهذا، إذا كنتم لا تنظرون في حقيقة الأمر فيما تنقرون فوقه، فسوف تخلطون أي شيء بأي شيء. فإذا كنتم تعملون في تنضيد طباعة الحروف، فسوف تحددون فروقًا لا يمكن للمستخدمين العاديين تحديدها. ومن ثم سوف يتعين علينا التوصل إلى أرضية مشتركة ومعقولة بشكل ما هنا.

أم المطلب الآخر الذي يمكن أن يكون معقولاً فهو أن نقول، عندما يتطرق الأمر إلى عناوين الإنترنت والسلاسل، فإننا نستهدف في حقيقة الأمر أشخاصًا على دراية بالنص أو اللغة التي يتم عرض المسمى بها. وهذا يعني على الرغم من ذلك أنه لا يمكننا في حقيقة الأمر استبعاد المواقف التي يبحث فيها المستخدم عن مسمى أو عنوان بالنص المستخدم أو بلغته الأم، ولكن عوضًا عن ذلك تُعرض عليه نتيجة أخرى تستخدم في حقيقة الأمر نصًا أجنبيًا بالنسبة للمستخدم. ويتصادف أن يكون لهذا الأمر سلسلة متشابهة في الظاهر. وعلى الرغم من ذلك، فإن ما يمكننا استبعاده من تعريف المستخدم المستهدف لدينا أو تعريف سياق مستهدف هو المواقف التي يحاول فيها المستخدم أن يقرر بين تسميتين أو عنوانين بنص لا يمكن للمستخدم أن يقرأها في حقيقة الأمر.

وإذا كنا نحاول أن نفهم ما هو التشابه المؤدي للبس بين السلاسل الصينية وفقًا لما يراه المستخدم المتحدث الأصلي للغة الأثيوبية، فإن هذا السؤال لا إجابة عنه بأن كل مسمى أو عنوان واحد باللغة الصينية بالنسبة له قد يبدو من حيث الشكل مثل المسمى الصيني الآخر. ونحن نفترض أيضًا أن قسمًا كبيرًا من المستخدمين بحكم الأمر الواقع على دراية بمجموعة نظام الترميز المعياري الأمريكي لتبادل المعلومات ASCII الفرعية. ويمكننا أيضًا أن نفترض لأسباب تاريخية

أن العديد منهم في حقيقة الأمر يتطلعون لبحث عن مسمى ASCII الصحيح من أجل الوصول إلى مورد من الموارد. ومن ثم فإن تعريف المستخدم المستهدف يجب أن يغطي السياق الذي يبحث فيه شخص من أصحاب اللغة بنص مختلف عنه عن مسمى بنظام الترميز ASCII.

وفي نهاية المطاف، يجب أن تكون هناك نقطة أخرى. فنحن نستبعد السياق الذي يحاول فيه المستخدم التفريق بين عنوانين أو مسميين حاضرين بصريًا في الوقت ذاته. وهذا من المواقف المستخدمة بشكل شائع جدًا بالنسبة للاختبارات المصطنعة ودراسات الأبحاث الخاص بالتشابه إلى حد اللبس، ولكن المواقف الحياتية الطبيعية تتمثل في أن المستخدم لديه فكرة عن اسم وبعد ذلك ينظر في تجسيد مادي لشيء ما ويجب عليه أن يقرر ما إن كان التجسيد المادي يطابق ما يبحث عنه أم لا. ونادرًا ما تكون لدينا الحالة التي يعرض علينا فيها عنوانان أو مسميان جنبًا إلى جنب ويجب علينا الاختيار من بينهما.

لننتقل إلى الشريحة التالية، رجاءً. والمفهوم الأساسي هنا هو أننا عندما نحاول القيام بعملية مراجعة تشابه سلاسل، فإنه بالتأكيد من الأشياء المستندة إلى السلاسل لأن الأمر مهم في نهاية المطاف، ليس إن كانت هناك نقاط أكواد فردية متشابهة، ولكن هل السلسلة الناجمة عن ذلك متشابهة أو من الممكن أن تسبب اللبس. ومن بين الأشياء التي تؤدي، حتى وإن كانت لديكم نقاط أكواد مسببة للبس، إذا كان هناك عنوان أو تسمية تحتوي على سلسلة فرعية فريدة أي أنها سلسلة فرعية مؤلفة من نقاط أكواد ليس لها أي مكافآت مشابهة في أي مكان آخر، فمن المفترض أن يكون ذلك كافيًا لجعل التسمية الكاملة فريدة لأنه ليس هناك أي طريقة لإنشاء شيء يمكن خلطه مع تلك التسمية، حتى وإن كانت بعض الأجزاء الأخرى في التسمية أو العنوان بها مكافئ مسبب للبس.

أما الجانب الآخر الملفت في الأمر هو أنه عندما ينظر الناس في العناوين أو التسميات، فإنهم يعاملونها معاملة مهمة القراءة العادية لهم أو التعرف على الكلمات. ونحن نعلم أن التعرف على الكلمات مسألة شمولية، بمعنى أن الناس لا يقرأون الكلمات حرفًا بحرف، ولكنهم يقرأون الكلمة في مجملها. وهذا يعني أنه كان ما يكفي من الكلمة بأكملها يتطابق مع توقع مع، فمن الممكن أن يؤدي إلى سياق يؤدي في الوعي الباطن إلى إدراك من جانب المستخدم على الرغم من أن بعض التفاصيل خاطئة. ومن ثم فإن القراءة الشمولية يمكن أن يؤدي إلى إلغاء بعض الفروقات البسيطة فيما هو معروض.

وهذا يعني أنه على الرغم من أن شيئا ما يجب أن يكون بارزًا، فليست هذه هي الحالة النموذجية لذلك وهو أن العديد من قراء اللغة الإنجليزية لن يتعرفوا ببساطة على أنكم قد أضفتم علامة تشكيل أو علامة صغيرة فوق أو أسفل أي حرف من أجل إنشاء عنوان أو مسمى زائف. وعند القراءة، فإن تلك العلامات غير موجودة. ولا تتم ملاحظتها. حيث تعتبر في أفضل الأحوال أنها غبار على الشاشة وكون المستخدمون سعداء إلى حد ما بالنقر فوق العنوان الخطأ الذي يحتوي على علامة تشكيل فوق حرف في مكان ما ليس هو مكانه الطبيعي.

إذن أعتقد أنني قد غطيت النقطة الثالثة بالفعل. فلننتقل إلى الشريحة التالية. من الجدير التمييز بين التشابه عن مفهوم المتباينات. الشريحة التالية. إذن فيما يخص المتباينات، فإننا نعرف المسمى الذي يكون بمعنى من معاني الكلمات مشابهًا ويعالجه المستخدم بنفس معالجة التسمية أو العنوان الأصلي. وذلك يعني أنه لا يحظى بنفس المظهر أو أنه قريب للغاية من نفس الشكل بحيث يعتبره المستخدم أنه متطابق. ولكن المتباينات غير مقتصرة على التشابه البصري وهو من المفاهيم الهامة للغاية. وحالات الاستخدام الشائع على سبيل المثال، هي الشكل المبسط في مقابل الشكل التقليدي للحروف الصينية. ويتم اعتبار هذه الحروف في نهاية الأمر نفس الحروف وقد يواصل المستخدم معاملتها بشكل مختلف حتى وإن لم تكن متشابهة من الناحية البصرية. وفي الأثيوبية، فإن لدينا حالة ملفتة تقدم فيها العديد من الحروف نفس الصوت ومن ثم يبادل الناس بينها بكل أريحية عن كتابة اللغة.

مرة أخرى، فهي لا تبدو متشابهة لكنها ذات تسمية وتهجئة بمجموعة واحدة من الحروف وتعامل نفس معاملة وتهجئة التسمية بمجموعة مختلفة من الحروف. وعلاقة المتباينات تتمثل في علاقة تكافؤ يشير إليها نفس الشيء كما أن جميع علاقات التكافؤ متعدية ومنتقلة. وهذا يعني أن أي متباين لأي تسمية هو أيضًا متباين لجميع المتباينات الأخرى أيضًا. وعلى الرغم من ذلك، فإن التشابه بالمعنى العام يعني أن هناك بعض الاختلاف رغم أنه دقيق.

وبسبب الاختلاف القابل للاكتشاف على الرغم من احتمالية تجاهل القراء له هو أن عنوانين أو تسميتين يمكن أن يكونا مشابهين لعنوان أو تسمية ثالثة لكنهما قد لا يكون قريبين بما يكفي من حيث الشكل لبعضهما لكي يعتبرا متشابهين. ومن ثم فإن التشابه وإمكانية اللبس ليسا متعديين أو متنقلين وهو ما يجعلنا نحذف بعض الأدوات الفعالة التي لدينا في التعامل مع مجموعات المتباينات.

الشريحة التالية من فضلك. إذن هناك بعض المتباينات التي تستند إلى تشابه بصري وإذا كان الأمر كذلك، فإما أن لديها مظهر متطابق أو شبه متطابق تمامًا. ومن بين الأشياء التي لا تراعيها المتغيرات هو أن المتغيرات الأخرى غير المتشابهة جدًا لدرجة عدم النظر فيها هو التشابه المستند إلى الحروف الكبيرة. إذن فإن عناوين ومسميات ASCII يمكن عرضها بالحروف الصغيرة أو الكبيرة وحتى أسماء النطاقات المدوَّلة يتم قبولها بالحروف الكبيرة من خلال غالبية وكلاء المستخدمين ومن ثم يمكن الحصول على مسميات وعناوين تكون معروضة بالحروف الكبيرة ويتعين على المستخدم أن يحدد ويفهم ما إن كان هذا هو المسمى والعنوان الذي يريده أم لا. ونحن لم نقم في منطقة الجذر بتضمين المتباينات بسبب إمكانية النقل والتعدي. ولدي مثال لتوضيح كيف أن ذلك الأمر يمكن أن يوقعك في ورطة. وعلى الرغم من ذلك، بالنسبة لتشابه السلاسل حيث لا يكون لدينا قيد إمكانية النقل والتعدي فإن إمكانية حدوث اللبس استنادًا إلى المكافئ الأساسي بالحروف الكبيرة يجب أن يكون في نطاق الدراسة.

الشريحة التالية رجاءً. إذن إليكم مثال على طريقة حدوث ذلك. حيث تبدأ بالحرف H في أعلى اليسار ويكون معك الحرف H في وضعه الكبير وغالبًا ما نضعه في مجموعة ASCII ولن يتعقب غالبية المستخدمين بشكل واع ما إن كانوا قد رأوا عنوانًا أو تسمية بحرف كبير أو حرف صغير وفي حقيقة الأمر فإن مسميات ASCII لكلا الشكلين صحيحين بالتساوي. وحرف H في الحالة الكبيرة يطابق بدقة من حيث الشكل حرف eta الكبير باليونانية. وسوف يكون ذلك في المعتاد عبارة عن متباين باستثناء أننا نتجاهل الحالة الكبيرة للحروف بالنسبة لتعريفات المتغيرات ولكن لديكم تشابه شبه تام. فإذا ما كانت لديكم تسمية أو عنوان يحتوي على حرف eta في الحالة الكبيرة، فلا يمكن للمستخدم التفريق عندئذ بين اليونانية واللاتينية. أما الشكل المصغر من الحرف eta فهو مختلف إلى حد ما لكنه ما يزال مشابهًا إلى كبير مع الحرف الصغير N بحيث يتسبب في إمكانية حدوث لبس. ليس لأن كل شخص من المحتمل، إذا كانت الأشياء معروضة في السياق الصحيح فربما يلاحظون غياب أو حضور ذلك الحرف الممتد قليلاً في ثنايا الحرف الثاني إلى اليمين.

أما الحرف الصغير n فله الحالة الكبيرة N في نظام الترميز ASCII وهو ما يتصادف أنه بنفس المظهر على وجه الدقة مثل الحرف الكبير nu باليونانية. وأنا أؤكد لكم مرة أخرى أنه ليس متباينًا لأننا نستبعد ذلك ولكن إذا كان هناك عنونا أو مسمى معروضًا بالحالة الكبيرة للحروف ومن المفترض أن تنقر فوقه، فلا يمكنك أن تحدد ما إن كنت تنظر في عنوان باللغة اليونانية أو

اللاتينية. أما الحرف nu باليونانية هو الشكل المصغر الذي تراه هناك مع نقطة الترميز 03BD وهو بدوره مشابه إلى حد ما للحرف الصغير v وبما يكفي لأن تعتقد أنه في سياق ما نفس المسميات والعناوين، فلديك هنا بالتأكيد مسببات للبس. والآن فيما يخص الحالة الصغيرة v فإن لها مكافئ الحرف الكبير للحالة الكبيرة وهي V وإذا ما حاولت إكمال الدائرة فسوف تدرك أن الحرف الكبير V والحرف الصغير h مختلفان تمامًا وبشكل مطلق من الناحية البصرية. وليس ثمة فرصة للبس بين الحرفين.

وهذا من بين الأسباب وراء عدم محاولة التعامل مع هذه الأشياء في مستوى المتغيرات حيث يمكن اعتبار كل واحد من هذه الأشياء متباينًا لجميع الحروف الأخرى. وسوف يكون لديك جميع حروف N وحروف H وحروف V باللغة اللاتينية متباينات لبعضها الآخر وهذا أمر غير عملي تمامًا. ولكن بالنسبة لتشابه السلاسل، فسوف يتوجب علينا تضمين هذه التأثيرات لأنه في غير ذلك يمكن أن تؤدي إلى استحداث مسميات وعناوين مربكة للغاية. وفي حقيقة الأمر، فإن المسميات التي لا يمكن للمستخدمين تحديد في أي النصوص هي. ولتناول الحالة، فإنك تعد تسمية HN فإن كانت في الحالة الكبيرة للأخرى فيمكنك معرفة إن كانت باللغة اليونانية أو اللاتينية.

لننتقل إلى الشريحة التالية، رجاءً. إذن هذا تكرار إلى حد ما للمصفوفة الكبيرة للأشياء التي تعرض لها سرمد بالشرح. إذن يتوجب علينا بسبب جميع المتباينات تعتبر إلى حد ما متشابهة للتسمية والعنوان، فإنها تصبح مسببة للبس والإرباك مع أي تمسية مفوضة أو متغيراتها. وقد استثنى سرمد حالة مقارنة المتغيرات المحجوبة مع متغيرات أخرى محجوبة. وأعتقد أن هذا من منظور فني، وفي حقيقة الأمر خطأ لأنه وحتى إن كانت المتغيرات المحجوبة يتم معاملتها من جانب المستخدم بأنه شيء يمكن قبوله بأنه نفس الشيء، وأن يكون لديك شيئين بديلين معقولين مستبعدان من المقارنة فذلك يبدو إغفالاً في رأيي.

لنتناول الشريحة التالية. ثمة تفاعلات شيقة مع النصوص المختلفة. واسمحوا لنا بأن نستكشفها في الشرائح التالية. وإليكم حالة أخرى، هذه هي الحالة الصغيرة، كيفية التفاعل بين النصوص المختلفة. إذن في الأعلى، لدينا جهة اليسار حرف لاتيني، Aang، وهو الحرف الذي لا يحتاجه الكثير منا لكنه مستخدم في بعض اللغات. وعلى الجهة اليمنى، لدينا حرف أرميني والذي يمكنكم الجدل بشأن تشابهه من عدمه. وقد عرضت عليكم هذين الاثنين بتداخل بسيط للغاية لكي أوضح لكم أن العديد من الناس قد يعتبرانهما مختلفين بسبب انحناء الساق في الاتجاه المقابل. وعلى

الرغم من ذلك، إذا ما أضفنا إلى المزيج، حرف Ada اليوناني، فإن له ساق غير معقوفة بأي من الاتجاهين. وفجأة يكون لديك تداخل كبيرة للغاية في التشابه.

ومن ثم من الممكن واعتمادًا على ما إن كنت قد نزلت، وما إن كان الاثنين في الأعلى متشابهين بما يكفي إلى حد اللبس، فإذا ما قررت بأنهما غير متشابهين بما يكفي، فربما تسمحون بأن يكونا حاضرين، ولكن بعد ذل كلا يمكن أن يكونا حاضرين في نفس الوقت مثل حرف Ada. لننتقل إلى الشريحة التالية، رجاءً. إن مفهوم ذلك التشابه يعود إلى الناظر إليه في حد ذاته، فالأمر مهم عندما يتطرق إلى النصوص ذات الطبيعة المختلفة تمامًا، لأن ما يقوم به القراء بالنسبة لكل نص، فإنهم ينظرون في المزايا الخاصة لكي تكون لهم القدرة على تحديد الكلمات بعيدًا عن بعضها بعضًا في حين أنهم سوف يتجاهلون المزايا الأخرى التي يكون الناس فيها غير مدركين بأن النص الذي قد يفكرون فيه مهم.

إذن ما هو نوع المزايا الرسومية التي تهمنا فهي مختلفة من نص إلى آخر. وعلى الرغم من ذلك بالنسبة لبعض النصوص، فإن لدينا بعض البيانات المتاحة والتي فكر فيها البعض بالفعل حول ما يمكن أن يكون مسببًا للإرباك واللبس، في حين أن هناك نصوص أخرى لا تسبب ذلك. وبعد ذلك هناك نصوص المشكلة فيها صعبة للغاية لأنه نصوص تشكل فيها الحروف المتجاورة شيئًا مشابهًا تمامًا لما نعرف أنه حروفًا مزدوجة من النص اللاتيني.

إذن عندما يكون لديكم شكل مزدوج ربما يبدو مختلفًا للغاية عن الحروف الفردية. وأعتقد أنه سوف تكون هناك حاجة إلى إجراء الكثير من الأبحاث من أجل فهم نطاق ونوع الأشكال المنصهرة التي يمكن أن تسبب اللبس مع الأشكال المنصهرة في نفس النص. وهو مكان يصبح فيه النظر في العنوان والتسمية الكاملة اكثر أهمية من العديد من النصوص الأخرى.

لننتقل إلى الشريحة التالية، رجاءً. إذن إليكم مفهوم آخر للنصوص وهو أن بعص النصوص مرتبطة ببعضها الآخر. وربما تكون لها نفس الميزات أو ميزات مشابهة أو يمكن أن تكون مرتبطة بسبب التطورات التاريخية أو كلاهما. إذن بالنسبة للغة اللاتينية واليونانية واللغة السيريلية والأرمنية، فهناك مصدر مشترك لكل هذه النصوص. ونتيجة ذلك هو أن هناك العديد والعديد من الأشياء ذات المنشأ المشترك. ولقد رأينا بعض الأمثلة فيما بين اللاتينية واليونانية بالفعل. وبالنسبة للغة الصينية واليابانية والكورية، فإن لديكم تلك اللغات، لديكم مخزون كامل متداخل. وجميعها بها قدر من الاستخدام أو استخدام بارز لنصوص الهان.

وعلى وجه العموم، حتى أجزاء ذلك فهي ليست بشكل مباشر نص الهان فإن لها سمات مماثلة من حيث تنضيد الحروف. فنصوص البراهمية الجديدة المستخدمة في الهند والدول المجاورة ذات منشأ مشترك. حيث لديها نفس الشكل والهيئة. والعديد من الأشكال مشتركة. ومن الجدير بالملاحظة أنه يمكنك اعتبار ذلك أمرًا غير محبب، لكن كان لدينا تطور كانت في فيه الخطوط بشكل خاص المستخدمة في واجهات المستخدمين، على سبيل المثال، خطوط مثل Segoe UI فهي الخطوط التي توضح الميل القوي لتناسق الحروف المكتوبة في سائر النصوص المستخدمة.

إذن في الغالب سوف تجدون اللغة اليونانية معروضة بنمط مختلف إلى حد ما بشكل عام من اللاتينية، أو تجدون الشكل الشائع تاريخيًا من الأرمينية والذي يبدو مختلفًا للغاية. ولكن إذا ما تطرقنا إلى تلك الخطوط، نجد أنه اختارت جميعًا أن تستخدم مجموعة من الأشياء المتساوقة حيث تبدو اليونانية قريبة للغاية من اللاتينية إلى حد كبير، وحتى الأرمينية تبدو قريبة من اللاتينية إلى أبعد حد، وهو ما يزيد من إمكانية تسبب المسميات والعناوين في اللبس واختلاط الأمر على المستخدمين.

لننتقل إلى الشريحة التالية، رجاءً. لإعطائكم فكرة عن الأمر، فإن النصين المستخدمين في الهند أي الكانادية والتيلوغو لهما بالأساس 34 شكلاً متطابقًا. فبالنسبة للغة اللاتينية واللغة السيريلية، فإن مجموعة الأشكال المتطابقة هو 28. وبالنسبة لمجموعة أخرى من الحروف الهندية، الديفنغارية واللغة القوموقية، فإن مجموعة الأشكال المتطابقة هو 26. فهذه تداخلات مرتفعة حقًا، ومن السهل فعليًا إنشاء تسمية أو مسمى في واحد من النصوص التي تبدو قابلة للتمييز بالكامل عن تسمية وعنوان في نص آخر. وعلى الجانب الآخر، لديكم نصوص يمكن عزوها بقوة مثل التايلندية واللاوية، لكنها تطورت على المستوى التاريخي إلى استخدام أنماط خطوط متباينة إلا حد كبير، ومن ثم ليس من السهل إنشاؤها وليس سهلاً أن تمثل نفس المسمى أو العنوان، أي العنوان أو المسمى الذي لا يمكن تحديد النص الذي كُتب به.

وكما ناقشنا، فإن اللاتينية واليونانية على وجه الخصوص بهما العديد من أشكال الحروف الكبيرة المشتركة، وهي تصبح مسببة للبس في حالة عرض التسمية أو العنوان على المستخدم، وهي الحرف الكبير الذي لا يمكننا السيطرة عليه. وعلى وجه الخصوص، فإن جميع المسميات غير أسماء النطاقات المدوَّلة يمكن أن تشبه الحروف الكبيرة باللغة اليونانية.

لننتقل إلى الشريحة التالية، رجاءً. من الأشياء الهامة للغاية التي يجب العلم بها هي أن عدد المتباينات المحددة لمسمى أو عنوان أو متباينات الحجب، قد تكون كبيرة بشكل لا يمكن توقعه. ولقد رأينا أمثلة على المسميات والعناوين العربية التي قد تجدون فيها عدد المتباينات المحجوبة في العدد البالغ 10,000، أو إذا كانت تسمية أو عنوانًا أطول، حتى إلى مليون. وهذا يعني أنه عندما تحدث سرمد حول إنشاء قائمة بجميع المتباينات المحجوبة، فهو من الأشياء العامة أو حتى الحالات النموذجية من غير الممكن حدوثها وحسب.

إذن يتوجب عليكم استخدام نوع مختلف من الأساليب، ولكن استخدام السرد هو نوع من أساليب القوة الغاشمة التي يمكن أن تنفع فقط في المسميات المكونة من حرفين وليس أكثر من ذلك. ببساطة ليس من الممكن لنا أن نقارن على المستوى الفردي جميع المتباينات في مقابل جميع المتباينات الأخرى. ومن ثم يجب عليكم القيام بشيء مختلف. وقد يكون من الممكن القيام بعملية مرحلية من مرحلتين تجدون فيها طريقة تلقائية لكي تزيلوا من المقارنة جميع أزواج المسميات التي لا يحتمل أن تكون مسببة للإرباك واللبس استنادًا إلى أنكم قد قررتم بأن بها ما يكفي من الميزات الواضحة لوسمها بأنها غير مسببة للإرباك واللبس. بعد ذلك نتمنى أن يقلل ذلك من مجموعة المسميات القابلة للمراجعة بحيث يمكنكم القيام بمراجعة أكثر تركيزًا لهم بصفة فردية.

الشريحة التالية من فضلك. ومن ثم فإن النقطة الرئيسية التي يمكن أن تكون فيما يخص ذلك التخفيض بالنسبة للمسميات والعناوين هي الحصول على عملية تلقائية بنوع ما اتكون مستمدة من البيانات وتعمل بشكل أساسي إلى حد ما بما يشبه الطريقة التي نكتشف بها التضاربات فيما بين المتباينات المحجوبة. لننتقل إلى الشريحة التالية، رجاءً. والطريقة التي نعتقد بها في الوقت الحالي أن هذا قد يكون أسلوبًا ناجحًا تتمثل في القيام بشيء مماثل لآلية المتباينات حيث لدينا في الوقت الحالي بالنسبة للمتباينات، لدينا متباين إما أنه مغلق أو غير ذلك. فهذا سؤال إجابته بنعم أو لا. ولاستبدال ذلك بشيء نعرّف فيه شيئًا مثل إمكانية اللبس أو متباين ولكن إعطائه مزايا مختلفة عن عدم التشابه إلى متطابق. وهذا من شأنه أن يكون بسبب التطابق سوف يتعامل مع الأشياء المسببة للإرباك بأنه مجموعة فائقة من المتباينات.

والآن فقد ذكرت من قبل أن المسميات المسببة للإرباك ليست متعدية وانتقالية. وهذا الأسلوب سوف يتظاهر بأن المسميات المسببة للبس كانت متعدية ومن ثم تتسبب في لبس وإرباك أكثر. وهذا أمر مرغوب في مستوى ما. ومن ثم فإنكم تقومون بتشغيل أداة استنادًا إلى هذا النوع من البيانات ويتم تسليمك شيء عبارة عن مجموعة فائقة لما يمكن أن يكون في حقيقة الأمر مسميات

مسببة للإرباك واللبس. ويمكنكم التفكير في هذا باعتباره تعريفًا فضفاضًا للمتباينات وسوف تشتمل على بعض مما لا يجب اعتباره من المسميات المسببة للبس والإرباك في حقيقة الأمر، ولكن ذلك سيكون من فقط من أجل المراجعة التفصيلية في المرحلة الثانية المقرر استبعادها.

لننتقل إلى الشريحة التالية، رجاءً. أما بالنسبة لكل من اللاتينية واليونانية واللغة السيريلية والأرمنية، لدينا بعض بيانات نقاط الأكواد الفردية في صورة منتجات ثانوية للجذور في وضع وتطوير قواعد إنشاء المسميات. كما أن لبعض السجلات وثائق سياسة قاموا فيها بتحديد المسميات المسببة للبس والإرباك ذات نقاط الأكواد الفردية. وهناك مجموعة من البيانات المتاحة من يونيكود، وهي بيانات المسميات المسببة للبس والإرباك. والمشكلة في ذلك تتمثل في أنها لم تم الاحتفاظ بها بشكل متسق. ويجري مراجعة اتجاهها في منتدى يونيكود وهي بحاجة إلى مصادقة ومحاذاة مع تعريف منطقة الجذر للمتباينات.

وفي الوقت الحالي، ثمة تضاربات غير مفيدة. وعلى الرغم من ذلك، وبصرف النظر عن البيانات التي تنظرون فيها، هناك بشكل عام بيانات ضعيفة للغاية بالنسبة لأي من التسلسلات. أي هل هناك نقاط أكواد متعددة في صف يكون من المسميات المسببة للبس والإرباك ذات تسلسل مختلف من نقاط الأكواد؟

على سبيل المثال، فإن النصوص التي تتم كتابتها في مجموعات، على غرار جميع نصوص البراهمية الجديدة، حيث سيتوجب عليك مقارنة مجموعة بمجموعة، ومن ثم لديك تسلسل في مقابل تسلسل. إلا أن هناك أمثلة ليست مقتصرة فقط على تلك النصوص، كما هو الحال في نظام الترميز ASCII، فإذا ما كتبت corn. في مقابل com. فربما يمكنك خداع نسبة معينة من المستخدمين.

الشريحة التالية من فضلك. إذن حتى إذا كنت تقوم بأعمال الفحص المسبق، فما تزال لديك مهمة الوصول إلى قرار بالموافقة أو الرفض حل هل هي من المسميات المسببة للبس والإرباك أم لا؟ ونحن نتحدث حول تطوير الإرشادات التوجيهية. حيث يجب أن نكون واضحين للغاية في أن هذا الأمر لا يمكن أن يكون خوارزمية. فإذا كان بإمكانكم تحديد خوارزمية، فيمكننا كتابة أداة من أجل القيام بذلك. ومن ثم سوف يتوجب أن يكون هناك قرار نهائي متروك لهيئة مراجعة تشابه السلسلة حيث يتوجب عليهم اتخاذ قرار بات. فسوف يتوجب عليهم على الأرجح استشارة خبراء نصوص لأن ما يهمنا هو كيفية نظر القراء الأصليين للغاية في أي تسمية أو عنوان.

ونحن نفكر في استبعاد استخدام مجموعات التركيز لأنه تمثل عملية باهظة التكلفة. كما أنه لا تتوفر منهجية جيدة في الوقت الحالي. وفي النهاية، فإننا ندرك الحاجة بأنه حتى وإن كانت لدينا عملية فحص مسبق، فقد يتوجب أن تكون لهيئة مراجعة تشابه السلسلة القدرة على الإضافة الواضحة لمرشحين إضافيين قابلين للمراجعة. وعند الانتهاء من العمل، سيكونون قد انتهوا من وضع وتطوير مجموعة تنافس على طول الأسطر التي ذكرها سرمد في السابق.

لننتقل إلى الشريحة التالية، رجاءً. إذن نحن في منتصف كتابة تقرير أولي وتجميع شيء مثل مجموعة بيانات أولية، ومن المقرر أن يلي ذلك فترة تعليق عامة على التقرير الأوليّ. وإذا ما رأى المجتمع أن المنهجية المقترحة مقبولة، فسوف نجري تشغيلاً تجريبيًا حسب الفحص المسبق أو عملية الفحص المسبق على عينة ما من المسميات والعناوين وأيضًا المسميات والعناوين المفوضة بالفعل لكي نعرف فقط ما إن كانت تبدو بأنها مسببة للإرباك وبعد ذلك تعديل مجموعات البيانات حسب مقتضى الحال.

وبعد ذلك في نهاية المطاف، سوف نحصل على تعقيبات وآراء المجتمع حول نتائج الاختبارات الأولية تلك والمنهجية النهائية. الشريحة التالية من فضلك. إذن هل تودون بدء طرح الأسئلة أم ماذا؟

سرمد حسين: حسنًا، لقد نفد الوقت المخصص للأسئلة. وإن كان هناك أي منها، ربما يمكننا تناول سؤال أو سؤالين سريعًا. أو أن بيتينان قد شاركت معكم عنوان بريد إلكتروني في مربع الدردشة. برجاء إرسال أسئلتكم إلى ذلك البريد الإلكتروني وسوف نتأكد من الرد عليكم أو لا تترددوا في المجي والتحدث إلينا. إذن أنيل، اسمحوا لنا بأن نتناول الأسئلة، رجاء. كما يوجد ميكروفون في الجانب الأيمن في حال أردت المجيء إن كنت هنا وتريد التحدث، فشكرًا لك. ولا يمكننا سماعك.

أنيل كومار جين: أنا أنيل، للعلم وإثبات ذلك في محضر الجلسة، وسؤال يتعلق بافتراض أن لجنة مراجعة تشابه السلسلة قد رأت النطاق والنطاق المقدم طلب له وقاموا بالموافقة عليه. لكن ربما عندما يبدأ في الدخول إلى مرحلة الاستخدام الفعلي في النظام ونجد أنه مماثل لأي من النصوص الحالية أو أي

نص آخر ونجد أن هناك شكاوى، فما هو الإجراء عند تلك المرحلة الخاصة هل بالتراجع أم باتخاذ إجراء ما؟ فما هي الأنظمة المتاحة لدينا فيما يخص ذلك؟ شكرًا.

سرمد حسين: في البداية، يجب ألا يحدث هذا الأمر في الأساس. وفي نهاية المطاف، فسوف يتوقف الأمر في حقيقة الأمر على عدة عوامل، على هيئة مراجعة تشابه السلسلة، فهناك أيضًا نظام إضافي للضوابط والتوازنات الرقابية في المنظومة. على سبيل المثال، هناك مجموعة كاملة من إجراءات الاعتراض والتي يمكن استخدامها أيضًا من خلال المجتمع الشامل بما يتجاوز الهيئة في التعامل مع بعض المشكلات التي تبدو واضحة، ولكن إذا لم تحددها هيئة مراجعة تشابه السلسلة لأي سبب من الأسباب. ومن ثم فإن هناك نظام للضوابط والتوازنات الرقابية الأخرى في عملية تقدم الطلبات والتي يمكن استخدامها أيضًا، آمل ذلك. ولكن بمجرد أن يتم تفويض ذلك، فسوف يتم المضي قدمًا فيها بشكل واضح، شكرًا لكم.

أنيل كومار جين: شكرًا.

سرمد حسين: سوف نتناول تعليق واحد أخير وسوف يتوجب علينا إنهاء هذه الجلسة بما يتجاوز الوقت المحدد.

نبيل بن عمار: شكرًا. أنا نبيل للعلم وإثبات ذلك في محضر الجلسة. بحسب ما أعلم، فقد قمنا بالفعل بهذا العمل من أجل النص العربي. وكانت هو المتباين، ولكن فيما يخص التشابه الآمن أو إمكانية حدوث اللبس والإرباك، فإن هناك شيء جديد. وهو موجود داخل نفس النص، إذ يمكننا الحصول على كلمتين متشابهتين، لكنهما مكتوبتان بلغتين مختلفتين. ومن ثم أود أن أسأل عم إن كان لدينا في هذه الهيئة مجموعة فرعية تعمل بشكل أساسي على النص العربي، لأن له بعض المواصفات والخصوصيات التي لا نجدها في النصوص الأخرى. شكرًا.

سرمد حسين: إذن وكما قال سرمد، سوف يتوجب علينا في نهاية المطاف وضع بعض مجموعات البيانات بما يتجاوز مجموعة المتغيرات وحسب التي تم وضعها من أجل النظر في التشابه. فقد قامت بعض هيئات قواعد إنشاء المسميات لمنطقة الجذر بالفعل بوضع بعض مجموعات البيانات الخاصة بالتسبب في اللبس والإرباك لتكون جزءًا من عملها. والبعض الآخر لم يفعل ذلك. فهذه إذن هي الأشياء التي ننظر فيها بشكل واضح وهذا هو المكان الذي يلزم فيه توفير تلك المجموعات من البيانات، وسوف نعمل بالفعل مع المجتمع من أجل وضعها. وبذلك أعتذر منكم لأننا لم نخصص وقتًا كافيًا للأسئلة والأجوبة، لكن برجاء التواصل معنا من خلال البريد الإلكتروني المقدم واسمحوا لي أن أقدم الشكر لكم جميعًا على حضور الجلسة. شكرًا لك أسما على التحدث وسوف ننهي هذه الجلسة. برجاء وقف التسجيل.

[نهاية التدوين النصي]