



คู่มือสำหรับการยอมรับสากล
(Universal Acceptance)



รับ



ตรวจสอบ



จัดเก็บ



ประมวลผล



แสดงผล

“การยอมรับสากล” คืออะไร

การยอมรับสากล คือสถานะที่ชื่อโดเมน และชื่ออีเมลที่ถูกต้องตามกฎหมาย ถูกปรับ ตรวจสอบ จัดเก็บ ประมวลผล และแสดงผลได้อย่างถูกต้องและเป็นไปในแนวทางเดียวกัน โดยทุกแอปพลิเคชัน ทุกอุปกรณ์ และทุกระบบที่สามารถเชื่อมต่อหรือใช้งานบนอินเทอร์เน็ต

เนื่องด้วยความเปลี่ยนแปลงอย่างต่อเนื่องของการเปิดใช้ชื่อโดเมนระดับบนสุดใหม่ ระบบหลายระบบไม่รู้จักหรือไม่สามารถประมวลผลชื่อโดเมนใหม่ ๆ เหล่านั้นได้ เช่น ชื่อที่มีความยาวมากกว่า 3 ตัวอักษร หรือ อยู่ในรูปแบบที่ไม่ใช่ ASCII หรือ การใช้อักขระภาษาท้องถิ่นเป็นชื่อโดเมน เป็นต้น เหตุการณ์ดังกล่าวเกิดขึ้นกับชื่ออีเมลที่สร้างจากชื่อโดเมนใหม่เหล่านี้ด้วยเช่นเดียวกัน

คณะทำงานขับเคลื่อนการยอมรับสากล (Universal Acceptant Steering Group – UASG) ก่อตั้งขึ้นโดย ไอแคนน์ (ICANN - Internet Corporation for Assigned Names and Numbers)

นำโดยชุมชนผู้มีส่วนได้เสียในอุตสาหกรรมอินเทอร์เน็ต มีภารกิจในการสร้างความตระหนักรู้ กำหนดประเด็นปัญหา และแก้ไขปัญหาต่าง ๆ อันเกี่ยวข้องเนื่องกับการยอมรับสากล ของชื่อโดเมน เพื่อวัตถุประสงค์ให้ผู้ใช้งานอินเทอร์เน็ตทั่วโลกใช้งานชื่อโดเมนได้อย่างเหมาะสมและได้รับประสบการณ์ที่ดีที่สุด

สำหรับข้อมูลเพิ่มเติมเกี่ยวกับ UASG และข่าวสารล่าสุด โปรดเข้าไปที่ : www.uasg.tech

ซอฟต์แวร์และบริการออนไลน์สนับสนุนการ

ยอมรับสากล ได้ เมื่อสามารถดำเนินการตามรายการข้างต้น

ทั้งหมดได้สำหรับ ทุกชื่อโดเมน และทุกชื่ออีเมล

หมายเหตุ: ในเอกสารนี้ การรับ การตรวจสอบ และการประมวลผล จัดเป็นความสามารถที่แยกจากกัน แต่ในทางปฏิบัติอาจเป็นกระบวนการที่คาบเกี่ยวกันได้

รับ



การรับ

คือกระบวนการที่ชื่ออีเมลหรือชื่อโดเมน ถูกรับเข้ามาในลักษณะอักขระเรียงต่อกันตามลำดับเป็นชื่ออีเมลหรือชื่อโดเมน จาก ส่วนรับข้อมูลจากผู้ใช้ (UI - user interface) หรือจากไฟล์ หรือ ส่วนรับข้อมูลจากโปรแกรมประยุกต์ (API - application program interface) ของซอฟต์แวร์ประยุกต์/แอปพลิเคชัน หรือบริการออนไลน์

ข้อเสนอแนะจาก UASG

- ส่วนรับข้อความที่ต้องการให้ผู้ใช้ใส่ชื่อโดเมนหรือชื่ออีเมลต้องสามารถรองรับข้อความในรูปแบบ Unicode ที่มีความยาวอักขระได้ถึง 256 ตัวอักษร
- หากต้องการ สามารถเพิ่มให้ผู้ใช้ใส่ข้อความรูปแบบรหัส ASCII ที่สอดคล้องกับชื่อโดเมน Unicode (หรือที่เรียกว่า Punycode) ได้ แต่ควรให้มีการแสดงผลแบบ Unicode เป็นหลัก และแสดงผลแบบ Punycode เมื่อเห็นว่าจำเป็นประโยชน์แก่ผู้ใช้งาน

ตรวจสอบ



คือกระบวนการตรวจสอบความถูกต้องของ
ชื่ออีเมลหรือชื่อโดเมนที่ได้รับหรือถูกส่งมา

โปรแกรมเมอร์จำนวนมากถูกปลุกฝังให้ตรวจสอบความถูกต้องของชื่อโดเมนระดับบนสุดโดยการนับจำนวนตัวอักษร หรือการตรวจสอบว่าอักขระเป็นรหัส ASCII หรือไม่ ซึ่งไม่จำเป็นอีกต่อไปเนื่องจากปัจจุบันชื่อโดเมนระดับบนสุดสามารถอยู่ในรูปแบบ Unicode (ไม่ใช่รหัส ASCII) ที่มีความยาวได้มากกว่าสามตัวอักษร

ข้อเสนอแนะจาก UASG

- ไม่ควรเริ่มกระบวนการตรวจสอบยกเว้นเป็นความต้องการสำหรับการดำเนินการของแอปพลิเคชันหรือบริการเท่านั้น วิธีการนี้เป็นวิธีที่ง่ายที่สุดในการทำให้ชื่อโดเมนที่ถูกต้องใช้ได้ ถูกรับเข้าสู่ระบบ
- หากต้องการตรวจสอบ แนะนำให้:
 - ▶ ทวนสอบชื่อโดเมนระดับบนสุดจากตารางที่เชื่อถือได้ ที่:
<http://www.internic.net/domain/root.zone>
<http://www.dns.icann.org/services/authoritative-dns/index.html>
<http://data.iana.org/TLD/tlds-alpha-by-domain.txt>
ดูรายละเอียดเพิ่มเติมที่ <https://tinyurl.com/sac070>.
 - ▶ ทำการสอบถามชื่อโดเมนกับระบบชื่อโดเมน (Domain Name System - DNS)
 - ▶ จัดให้มีการใส่ที่อยู่อีเมลสองครั้งเพื่อป้องกันปัญหาการพิมพ์ผิด
 - ▶ ตรวจสอบอักขระในชื่อเฉพาะเพื่อให้แน่ใจว่าไม่มีอักขระที่ระบุว่าเป็นอนุญาต ("DISALLOWED") หรือไม่ ได้ถูกใช้งาน ในรหัส Unicode อ่านข้อมูลเพิ่มเติมที่ <https://tools.ietf.org/html/rfc5892>
 - ▶ จำกัดขอบเขตของการตรวจสอบชื่อ ตามข้อกำหนดการตั้งชื่อ (whole-label rules) ในเอกสารในเอกสาร RFC5894 ให้น้อยที่สุด อ่านข้อมูลเพิ่มเติมที่ <https://tools.ietf.org/html/rfc5894>
 - ▶ ถ้าอักขระที่เรียงประกอบกันเป็นชื่อโดเมน มี อักขระคล้ายจุด เช่น ควรแปลงอักขระดังกล่าวให้กลายเป็น จุด . ก่อนทำการตรวจสอบ

จัดเก็บ



การจัดเก็บ

หมายความรวมถึงการจัดเก็บชื่อโดเมน

และชื่ออีเมลระยะยาวและ/หรือการจัด

เก็บชั่วคราว ระหว่างการประมวลผล

ไม่ว่าระยะเวลาในการจัดเก็บจะเป็นเท่าใด

ควรจัดเก็บในรูปแบบที่

RFC กำหนด เป็นหลัก

หรือจัดเก็บในรูปแบบอื่นที่สามารถแปลง

เป็นรูปแบบที่ RFC กำหนด

ข้อเสนอแนะจาก UASG

- แอปพลิเคชันและบริการออนไลน์ควรรองรับมาตรฐาน **Unicode** ที่เหมาะสม
- ควรจัดเก็บข้อมูลในรูปแบบ UTF-8 (Unicode Transformation Format) ทุกเมื่อที่ทำได้ บางระบบอาจต้องการรองรับอักขระในรูปแบบ UTF-16 ด้วย แต่โดยทั่วไปแล้วควรใช้รูปแบบ UTF-8 เป็นลำดับแรก และไม่ควรรู้รูปแบบ UTF-7 หรือ UTF-32
- คำนี้ถึงการใช้งานตลอดทั้งกระบวนการก่อนแปลงชื่อโดเมนไปมาระหว่างรูปแบบ ASCII (A-Label) และ Unicode (U-Label) ในขั้นตอนการจัดเก็บ หากต้องการอาจพิจารณาจัดเก็บเฉพาะ U-Label ในฐานข้อมูลหรือไฟล์ เพราะทำให้ง่ายต่อการจัดเก็บและค้นหา อย่างไรก็ตามหากระบบต้องทำงานร่วมกับแอปพลิเคชันหรือบริการเก่าที่ไม่รองรับ Unicode ควรพิจารณาจัดเก็บทั้งสองรูปแบบ
- การจัดเก็บควรระบุว่าเป็นชื่ออีเมลและชื่อโดเมนอย่างชัดเจนเพื่อความสะดวกในการเข้าถึง ยกตัวอย่างเช่น การจัดเก็บชื่ออีเมลและชื่อโดเมน ในชื่อฟิลด์ “ชื่อผู้เขียน” ของเอกสาร หรือ “ข้อมูลติดต่อ” ในล็อกไฟล์ ทำให้สูญเสียความหมายที่แท้จริงของชื่ออีเมลและชื่อโดเมนไป

ประมวลผล



การประมวลผลชื่ออีเมลและชื่อโดเมนเกิดขึ้นเมื่อแอปพลิเคชันหรือบริการออนไลน์ดำเนินการ (เช่น ลืบค้น หรือจัดเรียงลำดับรายการ) หรือทำการแปลงรูปแบบข้อมูล (เช่น การจัดเก็บชื่อ ASCII ในรูปแบบ Unicode) ซึ่งอาจมีการตรวจสอบความถูกต้องของชื่อเพิ่มเติมระหว่างกระบวนการดังกล่าว ชื่ออีเมลและชื่อโดเมนอาจถูกประมวลผลได้อีกหลากหลายรูปแบบ* สิ่งผลสนับสนุนว่าการมีข้อตกลงที่ชัดเจน เพื่อให้ระบบสามารถเข้าใจและแยกแยะข้อมูลได้อย่างถูกต้อง เป็นสิ่งที่จำเป็น

ข้อเสนอแนะจาก UASG

- เนื่องด้วยมาตรฐาน Unicode มีการปรับปรุงเพิ่มเติมอยู่เสมอ ผู้พัฒนาแอปพลิเคชันหรือบริการควรตรวจสอบการเพิ่มขึ้นของ code point ใหม่ ๆ เพื่อให้มั่นใจว่ารหัส unicode ใหม่ ๆ เหล่านี้จะไม่กระทบต่อแอปพลิเคชันหรือบริการและ “ทำลาย” ประสบการณ์ของผู้ใช้ เช่น หากฟอนต์ไม่อยู่ในระบบปฏิบัติการ อาจจะทำให้แสดงผลอักขระไม่ได้ (โดยมากระบบจะแทนที่อักขระนั้นด้วยช่องว่าง) อย่างไรก็ตาม กรณีนี้นี้ไม่ควรส่งผลกระทบต่อระบบหยุดทำงาน
- ใช้ API ที่รองรับรหัส Unicode
- ใช้เอกสารล่าสุด เกี่ยวกับโปรโตคอล Internationalized Domain Names in Application (IDNA) [<http://tools.ietf.org/html/rfc5891>] และตาราง [<http://tools.ietf.org/html/rfc5892>] สำหรับชื่อโดเมนภาษาท้องถิ่น
- ประมวลผลในรูปแบบ UTF-8 เป็นหลัก
- ควรตรวจสอบให้แน่ใจว่าระบบสามารถจัดการตัวเลขได้ตามที่คาดหวังตัวอย่างเช่น ตัวเลข ASCII และอักษรแสดงตัวเลขในภาษาท้องถิ่นเอเชียควรได้รับการประมวลผลแบบตัวเลข
- ควรปรับปรุงแอปพลิเคชันและเครื่องแม่ข่าย/บริการไปพร้อม ๆ กัน ถ้าด้านหนึ่งใช้ Unicode ในขณะที่ด้านหนึ่งไม่ใช้ จะทำให้ต้องแปลงข้อมูลทุกครั้งที่มีการส่งไปมาระหว่างกัน
- ทบทวนโค้ดให้มั่นใจว่าจะไม่เกิดกรณีถูกโจมตีด้วยข้อมูลที่มึขนาดใหญ่มากกว่าที่ระบบรับได้ (buffer overflow attack) ในระหว่างการเปลี่ยนรูปแบบอักขระ อาจทำให้ขนาดของข้อความขยายขึ้นหรือลดลงอย่างมาก

* ตัวอย่าง: การระบุบุคคลในประเทศนิวซีแลนด์ด้วยการค้นหาจากโดเมนรหัสประเทศระดับบนสุด .nz; ค้นหาเภสัชกร (pharmacists) โดยค้นหาจาก ชื่ออีเมล user@*.pharmacist

แสดงผล



การแสดงผลเกิดขึ้นเมื่อชื่อโดเมนหรือชื่ออีเมลถูกแสดงต่อผู้ใช้ โดยปกติการแสดงผลเกิดขึ้นตรงไปตรงมาหากระบบปฏิบัติการรองรับอักขระชุดนั้น และสายอักขระถูกจัดเก็บในรูปแบบ Unicode อย่างไรก็ตาม การเปลี่ยนแปลงรูปแบบสายอักขระสำหรับแอปพลิเคชันที่เฉพาะเจาะจงบางระบบ อาจเกิดขึ้นได้

ข้อเสนอแนะจาก UASG

- ควรแสดงทุก code points ของ Unicode ที่รองรับโดยระบบปฏิบัติการนั้นๆ หากแอปพลิเคชันใช้ชุดฟอนต์เฉพาะของตนเอง ควรดำเนินการเพื่อสนับสนุนการแสดงผล Unicode ให้ครอบคลุมชุดฟอนต์เหล่านั้นจากระบบปฏิบัติการด้วย
 - เมื่อพัฒนาแอปพลิเคชันหรือบริการ หรือเมื่อบริหารจัดการระบบทะเบียนชื่อโดเมน ควรพิจารณาถึงภาษาที่อนุญาตให้ใช้ได้ และตรวจสอบให้แน่ใจว่าระบบปฏิบัติการและแอปพลิเคชันรองรับภาษาเหล่านั้น
 - ควรเปลี่ยนข้อมูลให้อยู่ในรูปแบบ Unicode ก่อนการแสดงผล ตัวอย่างเช่น ผู้ใช้งานควรเห็น “everyone. みんな” ไม่ใช่ “everyone.xn--q9jyb4c” (การเปลี่ยนข้อมูลนี้เป็นตัวอย่างการประมวลผลแบบ UA-ready)
 - แสดงผล Unicode เป็นหลัก และใช้ Punycode เฉพาะเมื่อให้ประโยชน์ต่อผู้ใช้เท่านั้น โดยอาจจัดแสดงผลแบบ Punycodeed เมื่อนำมาใส่ไปวางบนข้อความที่เป็น Unicode
 - ชื่ออีเมลที่ผสมอักขระหลายประเภทเข้าด้วยกันจะกลายเป็นสิ่งที่พบเห็นโดยทั่วไป อักขระสองตัวที่แตกต่างกันอาจจะดูเหมือนกันสำหรับบางคน แต่เป็นอักขระที่แตกต่างกันสำหรับคอมพิวเตอร์ อย่าสันนิษฐานว่าการผสมนี้เกิดจากความตั้งใจประสงค์ร้าย เช่นการฟิชซิง และการแสดงผลควรอยู่ในรูปแบบที่ไม่เป็นผลเสียกับผู้ใช้ภาษาอื่นที่ไม่ใช่ภาษาอังกฤษ (ละติน)
- อ่านเพิ่มเติมเกี่ยวกับประเด็นด้านความปลอดภัยของ Unicode ได้ที่ <http://unicode.org/reports/tr36/>
- ควรใช้ Unicode IDNA Compatibility Processing เพื่อให้การทำงานของระบบเป็นไปอย่างผู้ใช้คาดหวัง อ่านเพิ่มเติมได้ที่ <http://unicode.org/reports/tr46/>
 - ควรระวังอักขระที่ไม่ได้กำหนดไว้หรือไม่ได้รับอนุญาต อ่านเพิ่มเติมได้ที่ <https://tools.ietf.org/rfc/rfc5892.txt>

การเตรียมพร้อมเพื่อรองรับการยอมรับสากล (Universal Acceptance-Ready)

การทวนสอบ Source Code และ การทำ Unit Testing

สำรวจซอสโค้ดว่ามีการใช้เทคนิคการเขียนโปรแกรมไลบรารี และ ส่วนต่อประสานกับโปรแกรมประยุกต์ (API) ที่ถูกต้องเท่านั้น

จากนั้นจึงตรวจสอบการทำงานของแอปพลิเคชันหรือบริการว่าสามารถดำเนินการตามข้อกำหนดที่ปรากฏด้านบน (การรับ การตรวจสอบ ฯลฯ)

วิธีนี้เหมาะสำหรับผู้พัฒนาระบบและผู้ให้บริการออนไลน์ เท่านั้น

UASG

ได้ติดต่อไปยังผู้พัฒนาแอปพลิเคชันและผู้ให้บริการรายใหญ่ โดยตรง

เพื่อผลักดันให้ผู้พัฒนาและผู้ให้บริการ

ทำการทวนสอบ source code

และทดสอบเพื่อการยอมรับสากล

รวมถึงร่วมแบ่งปันรายการข้อกำหนดสำหรับใช้ในการพัฒนากรณีทดสอบ

การทดสอบโดยผู้ทดสอบ

ทำการทดสอบหลายอย่างกับชื่อโดเมนใหม่ๆ และชื่อที่ไม่อยู่ในรูปแบบ ASCII

เช่นการใส่ชื่ออีเมลเมื่อลงทะเบียนใช้บริการออนไลน์

และการตรวจสอบว่าชื่ออีเมลนั้นถูกรับเข้าระบบหรือไม่

เนื่องจากว่ามีบริการออนไลน์จำนวนมาก

และชื่ออีเมลในรูปแบบที่เกิดจากการผสมกันหลากหลายรูปแบบ (ASC,II, Unicode, Punycode) เช่น

ประชาสัมพันธ์@ชื่อไทย.tech

วิธีนี้ต้องทดลองผสมผสานระหว่างแอปพลิเคชัน

บริการ ชื่ออีเมล และชื่อโดเมน

เพื่อให้ครอบคลุมกรณีที่อาจเกิดขึ้นในการใช้งาน

จริงการทดสอบนี้ทำได้โดยทุกคนแต่เป็นงานที่ต้องใช้แรงงาน ในการทดสอบให้ครอบคลุมทุกกรณี

UASG ส่งเสริมการทดสอบลักษณะนี้

โดยการรวบรวมและประกาศรายชื่อเว็บไซต์

แอปพลิเคชัน ชื่ออีเมล และชื่อโดเมน

ที่เหมาะสมสำหรับใช้เป็นชื่อในการทดสอบ

การทดสอบอัตโนมัติ

การใช้สคริปต์หรือคำสั่งอัตโนมัติในการทดสอบชื่อเว็บไซต์ที่หลากหลาย

วิธีนี้ต้องมีการทำงานด้านเทคนิคล่วงหน้า

แต่สามารถขยายขอบเขตการวัดและติดตามผลได้

ดีตัวอย่างที่เกิดขึ้นจริง คือ APNIC

ได้รับการมอบหมายจาก ICANN

ให้ทำการทดสอบการยอมรับสากลของโดเมนระดับบนสุดใหม่

<https://tinyurl.com/new-gtld-ua>

เมื่อปี ค.ศ.2015

และใช้การทดสอบแบบอัตโนมัติในงานดังกล่าว

UASG

กำลังศึกษาวิธีการทดสอบการยอมรับสากลแบบ

อัตโนมัติ และจะแจ้งผลให้ทราบต่อไป