



Guia Rápido da Aceitação Universal

O que significa “Aceitação Universal”?



ACEITAR



VALIDAR



ARMAZENAR



PROCESSAR



EXIBIR

Software e serviços on-line são compatíveis com a Aceitação Universal quando eles oferecem os recursos listados acima para todos os nomes de domínio e e-mails.

UA (Universal Acceptance, Aceitação Universal) é o estado em que todos os nomes de domínio e endereços de e-mail válidos são aceitos, validados, armazenados, processados e exibidos de maneira correta e consistente por todos os aplicativos, dispositivos e sistemas que utilizam a Internet.

Devido ao cenário tão dinâmico de nomes de domínio, muitos sistemas não reconhecem nem processam adequadamente novos nomes de domínio, principalmente porque eles podem ter mais de três caracteres ou estar em um formato diferente de ASCII. O mesmo acontece com endereços de e-mail que incorporam essas novas extensões.

O UASG (Universal Acceptance Steering Group, Grupo de Gestão de Aceitação Universal), estabelecido pela ICANN (Internet Corporation for Assigned Names and Numbers, Corporação da Internet para Atribuição de Nomes e Números), é uma iniciativa movida pela comunidade e que abrange todo o setor dedicado à aquisição de conhecimento e identificação e resolução de problemas associados à aceitação universal de nomes de domínio. O objetivo desse trabalho é ajudar a garantir uma experiência consistente e positiva dos usuários da Internet globalmente.

Para obter mais informações sobre o UASG e recentes acontecimentos, acesse: www.uasg.tech.

ACEITAR



Aceitar é o processo pelo qual um endereço de e-mail ou um nome de domínio é recebido como uma cadeia de caracteres de uma interface de usuário, arquivo ou API (Application Program Interface, Interface entre Programa e Aplicativo) usado por um aplicativo de software ou serviço on-line.

Recomendações do UASG

- Qualquer elemento da interface do usuário que exigir que o usuário digite um nome de domínio ou endereço de e-mail precisa ser compatível com Unicode e cadeias de caracteres com até 256 caracteres.
- Os usuários devem ter permissão, mas não obrigação, de inserir texto compatível com codificação ASCII (também conhecido como "Punycode") no lugar do seu Unicode equivalente. No entanto, o Unicode deve ser exibido por padrão, sendo que o texto com Punycode é exibido para o usuário apenas quando for vantajoso para ele.

VALIDAR



O processo pelo qual um endereço de e-mail ou nome de domínio, recebido ou emitido, é verificado para garantir que a sintaxe está correta. Muitos programadores foram treinados para validar seguindo a heurística que exige verificar se um domínio de primeiro nível tem o número “correto” de letras ou se as letras são do mesmo conjunto de caracteres ASCII. Essas heurísticas são se aplicam mais devido à introdução de nomes de domínio com mais de três caracteres e caracteres Unicode (diferente de ASCII).

Recomendações do UASG

- As validações não devem ocorrer a menos que sejam exigidas para a operação do aplicativo ou do serviço. Essa é a maneira mais fácil de garantir que todos os nomes de domínio válidos sejam aceitos nos sistemas.
- Se a validação for obrigatória, considere o seguinte:
 - ▶ Verifique a porção TLD de um nome de domínio em comparação a uma tabela oficial:
<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>
Consulte também o SAC070: <https://tinyurl.com/sac070>.
 - ▶ Consulte o nome de domínio com relação ao DNS (Domain Name System, Sistema de Nomes de Domínio).
 - ▶ Exija que o endereço de e-mail seja informado mais de uma vez para descartar os erros de digitação.
 - ▶ Valide os caracteres nos rótulos para determinar se o rótulo U não contém pontos de código "NÃO PERMITIDO" ou pontos de código que não estão atribuídos na sua versão em Unicode. Visite: <https://tools.ietf.org/html/rfc5892>.
 - ▶ Limite a validação de rótulos para uma quantidade pequena de regras de rótulos inteiros definidas nas RFCs (Request for Comments, Solicitações de Comentários). Visite: <https://tools.ietf.org/html/rfc5894>.
 - ▶ Se uma cadeia de caracteres semelhante a um nome de domínio contiver o caractere de ponto final ideográfico ‘。’, ele deve ser convertido para um ‘.’ antes de realizar a validação.

ARMAZENAR



Armazenar refere-se ao armazenamento a longo prazo e/ou transiente de nomes de domínio e endereços de e-mail. Independentemente do tempo de vida dos dados, eles devem ser armazenados nos formatos definidos nas RFCs (preferencialmente) ou em outros formatos que podem ser transformados em formatos definidos na RFCs.

Recomendações do UASG

- Os aplicativos e serviços devem ser compatíveis com os padrões adequados de Unicode.
- As informações devem ser armazenadas em UTF-8 (Unicode Transformation Format, Formato de Transformação Unicode) sempre que possível. Alguns sistemas também podem exigir compatibilidade com o UTF-16, mas geralmente há uma preferência pelo UTF-8. UTF-7 e UTF-32 devem ser evitados.
- Consider all end-to-end scenarios before converting A-Labels to U-Labels and Considere todos os cenários completos antes de converter rótulos A em rótulos U, e vice-versa, ao armazenar. Pode ser interessante manter apenas rótulos U em um arquivo ou banco de dados, porque isso simplifica as pesquisas e classificações. No entanto, a conversão pode ter implicações na interoperabilidade com aplicativos e serviços mais antigos que não usam Unicode. Considere armazenar nos dois formatos.
- Marque claramente os endereços de e-mail e nomes de domínio durante o armazenamento para facilitar o acesso. Instâncias em que os endereços de e-mail e nomes de domínio foram arquivados no campo "author" (autor) de um documento ou "contact info" (informações de contato) em um arquivo de registro resultaram na perda da origem enquanto um endereço.

PROCESSAR



O processamento ocorre sempre que um endereço de e-mail ou nome de domínio é usado por um aplicativo ou serviço para executar uma atividade (por exemplo, para pesquisar ou classificar uma lista) ou é transformado em um formato alternativo (por exemplo, para armazenar ACSII como Unicode). É possível que seja realizada validação adicional durante o processamento.

Os nomes de domínio e endereços de e-mail podem ser processados de maneiras ilimitadas*, o que reforça a necessidade de haver convenções que garantam que os dados estão sendo compreendidos e classificados de modo consistente.

Recomendações do UASG

- Uma vez que o padrão Unicode é expandido continuamente, os pontos de código não definidos quando o aplicativo ou serviço foi criado devem ser verificados para garantir que eles não “interromperão” a experiência do usuário. Fontes ausentes no sistema operacional subjacente podem fazer com que alguns caracteres não sejam exibidos (muitas vezes o caractere “ ” é usado para representá-los), mas essa situação não deve resultar em uma falha fatal.
- Use APIs compatíveis com Unicode.
- Use os documentos mais recentes do protocolo [<http://tools.ietf.org/html/rfc5891>] e tabelas [<http://tools.ietf.org/html/rfc5892>] de IDNA (Internationalized Domain Names in Applications, Nomes de Domínio Internacionalizados em Aplicativos) para IDNs (Internationalized Domain Names, Nomes de Domínio Internacionalizados).
- Processe no formato UTF-8 sempre que possível.
- Certifique-se de que o produto ou recurso lide com os números da maneira esperada. Por exemplo, numerais ASCII e representações de números ideográficas asiáticas devem ser tratados como números. [RFC5892, link acima]
- Faça o upgrade de aplicativos e servidores/serviços juntos. Se o servidor for Unicode e o cliente não for Unicode, ou vice-versa, será necessário converter os dados para cada página de código sempre que os dados forem transferidos entre o servidor e o cliente.
- Faça análises da codificação para enviar ataques de buffer overflow. Ao fazer a transformação de caracteres, as cadeias de texto podem aumentar ou diminuir significativamente.

Ejemplos: Identificar personas en Nueva Zelanda al buscar dentro del ccTLD .nz; identificar farmacéuticos al buscar direcciones de correo electrónico usuario@.farmacéutico.

EXIBIR



A ação exibir ocorre sempre que um endereço de e-mail ou um nome de domínio é renderizado em uma interface de usuário. Exibir nomes de domínio e endereços de e-mail é geralmente algo bastante simples quando as escritas usadas são compatíveis com o SO subjacente e as cadeias de caracteres são armazenadas em Unicode. No entanto, algumas transformações específicas de aplicativos podem exigir algo a mais.

Recomendações do UASG

- Exiba todos os códigos de pontos Unicode compatíveis com o sistema operacional subjacente. Se um aplicativo tiver seus próprios conjuntos de fontes, é necessário oferecer um suporte abrangente a Unicode para a coleta de fontes disponíveis no sistema operacional.
- Ao desenvolver um aplicativo ou um serviço ou ao operar um registro, considere os idiomas compatíveis e certifique-se de que o SO e os aplicativos aceitem esses idiomas.
- Converta dados diferentes de Unicode para Unicode antes de exibir. Por exemplo, o usuário final deve ver “todos.みんな”, em vez de “todos.xn--q9jyb4c”. (Essa conversão é um exemplo de processamento compatível com a UA).
- Exibir Unicode por padrão. Usar texto com Punycode para o usuário apenas quando isso for vantajoso para ele. Aumente a exibição de Unicode com texto em Punycode exibido ao passar o mouse sobre ele como uma forma de mitigação.
- Considere que endereços com misturas de escritas serão mais comuns. Alguns caracteres Unicode podem parecer iguais ao olho humano, mas diferente para os computadores. Não presuma que as cadeias de caracteres com mistura de escritas são destinadas a fins maliciosos, como phishing, e se a interface do usuário chamar a atenção do usuário para as cadeias de caracteres, certifique-se de que isso seja feito de maneira que não prejudique os usuários de escritas não latinas. Saiba mais sobre considerações de segurança com Unicode em: <http://unicode.org/reports/tr36/>.
- Use o processamento de compatibilidade de IDNA para Unicode a fim atender às expectativas dos usuários. Para saber mais, acesse: <http://unicode.org/reports/tr46/>.
- Esteja atento a caracteres não atribuídos ou não permitidos. Saiba mais na RFC 5892: <https://tools.ietf.org/rfc/rfc5892.txt>.

Torne-se compatível com a Aceitação Universal

Revisões de código fonte e teste de unidade

O processo de inspeção do código fonte e verificação de que apenas as técnicas de programação, bibliotecas de software e interfaces (também conhecidas como “APIs”) corretas foram usadas. Depois disso, o administrador poderá verificar se o aplicativo ou serviço funciona testando-o com as capacidades específicas (aceitar, validar etc.) listadas anteriormente. Esse método é geralmente usado somente por desenvolvedores de aplicativos e provedores de serviços on-line.

Como parte do trabalho de conscientização do UASG, o grupo está entrando em contato diretamente com os desenvolvedores de aplicativos e os maiores provedores de serviços on-line para incentivá-los a fazer as revisões de código fonte e testes

para aceitação universal e compartilhar uma lista de critérios que podem ser usados para desenvolver os casos de teste.

Testes manuais

Exigem a execução de vários testes em domínios novos e que não utilizam ASCII, como o envio de um endereço de e-mail ao registrar para um serviço on-line e verificação se ele foi aceito. Uma vez que existem vários possíveis serviços on-line que aceitam inscrições, bem como possíveis novas combinações de endereços de e-mail, esse método exige a tentativa de diferentes combinações de aplicativos, serviços, endereços de e-mail e/ou nomes de domínio para fornecer um escopo maior de casos de uso. Esse método pode ser executado por qualquer pessoa, mas é mais trabalhoso.

O UASG também está ajudando a promover esse método

desenvolvendo uma lista dos principais sites, aplicativos, endereços de e-mail e nomes de domínio apropriados para esses testes.

Testes automatizados

O uso de scripts automatizados ou diretivas para testar uma variedade de URLs. Esse método exige um trabalho técnico mais direto, mas pode ser dimensionado com mais facilidade para trabalhos maiores de medição e monitoramento. Um exemplo do mundo real é a recente investigação de gTLDs realizada pelo APNIC em nome da ICANN: <https://tinyurl.com/new-gtld-ua>.

O UASG está examinando alguns métodos de testes automatizados para a aceitação universal e divulgará suas descobertas quando estiverem disponíveis.

