“Ok Google, já chega”: privacidade e reconhecimento de fala ininterrupto em celulares

“Ok Google, that’s enough”: privacy and ceaseless speech recognition in mobile phones

Miguel Said Vieira *

RESUMO

Este artigo discute as implicações para privacidade da tecnologia de reconhecimento de fala ininterrupto em celulares. Apresenta um breve histórico da evolução das tecnologias de reconhecimento de fala, indicando sua proximidade com os setores militar e de inteligência, e o paralelismo entre os avanços nas técnicas de processamento e no *hardware* computacional. Examina avanços tecnológicos da Google que conduziram à implementação do reconhecimento de fala ininterrupto em celulares; identifica algumas violações de privacidade possibilitadas por isso, e conclui relacionando-as a tipologias de privacidade, e à noção de sociedade de controle.

Palavras-chave: Reconhecimento de Fala; Dispositivos Móveis; Google; Privacidade; Sociedade de Controle.

ABSTRACT

This paper discusses the privacy implications of ceaseless speech recognition technology in mobile phones. It presents a brief history of the evolution of speech recognition technologies, indicating its proximity to the US military and intelligence sectors, and the parallelism between the advances in processing techniques and computer hardware. It examines technological advances by Google that led to the implementation of ceaseless speech recognition in mobiles; identifies some privacy violations made possible by this, and concludes relating them to privacy typologies, and to the notion of control society.

Keywords: Speech Recognition; Mobile Devices; Google; Privacy; Society of Control.

INTRODUÇÃO

Em junho de 2004, Richard Stallman (2004), o pai do movimento do *software* livre, foi convidado pela Google a fazer uma apresentação – uma das famosas “Google talks”: série de palestras (realizadas nos escritórios da empresa) cujo canal de Youtube já acumula 42 milhões de visualizações. Ao final de sua fala, respondendo a perguntas relacionadas à segurança e anonimato, Stallman comentou que se recusava a usar um celular, e explicou seu motivo: “Não quero carregar um dispositivo pessoal de rastreamento”. Não era a primeira vez que Stallman tocava no assunto — em 2002,

* Doutor em Educação pela Universidade de São Paulo (USP). Professor adjunto, Universidade Federal do ABC (UFABC). Endereço: Av. dos Estados, 5.001, sl. A646-1, CEP 09210-580 Santo André, SP. E-mail: miguel.vieira@ufabc.edu.br.

ele já chamara a atenção para o fato em sua página pessoal (STALLMAN, 2002); e em 2008, ele ampliaria suas preocupações: o problema de celulares não é apenas serem dispositivos de rastreamento, mas também de vigilância, uma vez que seus microfones podem ser ativados remotamente — pelo aparato estatal, por exemplo (STALLMAN, 2008).[1]

Embora Stallman tenha em geral fundamentado essas preocupações em notícias divulgadas na grande mídia (que davam conta desse uso dos celulares por forças policiais, ou de leis que o autorizavam), ele era frequentemente tratado como um paranoico, ou mesmo teórico da conspiração (HOLWERDA, 2012); não apenas por seus detratores – como Eric Raymond (2012), que o tachou de "fanático" em relação a suas posições sobre *software* livre –, mas até mesmo por quem o admirava. Em 2013, porém, 11 anos após a primeira vez que Stallman registrara esses temores em seu *site* pessoal, as revelações de Edward Snowden mostrariam que ele era plenamente justificado. A frase "Stallman was right" [Stallman estava certo] tornou-se uma espécie de mantra de arrependimento, estampada em camisetas e nomeando comunidades do Reddit.[2]

À luz dessas revelações, sua palestra na Google em 2004 é duplamente simbólica. Em primeiro lugar, porque elas evidenciaram o quanto a Google e o Android (seu sistema operacional para celulares) foram alvos e, em grande medida, "colaboradores" (voluntários ou não) da espionagem estatal global conduzida pela Five Eyes (aliança entre os órgão de inteligência dos EUA, Austrália, Canadá, Nova Zelândia e Reino Unido), encabeçada pela NSA. A agência estadunidense hackeara a comunicação entre *datacenters* da Google (GELLMAN; SOLTANI, 2013); incluía a empresa no Prism, seu principal programa secreto de espionagem — de participação não voluntária, mas que envolvia pagamentos para custeio das atividades de espionagem (MACASKILL, 2013); e mantinha forças-tarefas dedicadas a explorar vulnerabilidades de cada sistema operacional de celulares — incluindo o Android (ROSENBACH; POITRAS; STARK, 2013).

Em segundo lugar, porque, ainda que a Google tenha se posicionado publicamente contra as iniciativas de vigilância *estatal* (como as reveladas por Snowden), ela continuou sendo — ao lado de Facebook e Apple — um dos grandes atores da vigilância global de cunho *mercantil*.[3] Seu modelo de negócios é construído fundamentalmente em torno da coleta, acúmulo e processamento de informações sobre seus usuários, para aumentar a eficácia (e reduzir os custos) da publicidade comportamental dirigida a eles e vendida pela Google (VIEIRA, 2014, p. 289-318); um modelo particularmente bem-sucedido: estima-se que, em 2013, a empresa abocanhava metade de todo o valor gasto nesse mercado (GOOGLE, 2014; JOHNSON, 2013).

Este artigo aborda um dos mecanismos pelos quais a Google – assim como diversas outras empresas – tornou-nos, por meio de avanços tecnológicos, mais sujeitos à vigilância constante. O mecanismo em questão é o reconhecimento de fala ininterrupto em celulares, que seria implementado pela primeira vez... justamente em 2013.

---

[1] Stallman voltaria a esses temas diversas vezes — mas esses são os registros mais antigos em sua página pessoal da preocupação com os riscos de, respectivamente, rastreamento e vigilância por celulares.

[2] Disponível em: <http://trenchant.spreadshirt.com/rms-was-right-A13230701>; e em: <https://www.reddit.com/r/StallmanWasRight/>.

[3] Para uma reflexão sobre a importância da vigilância mercantil, ver Vieira e Evangelista (2015).

BREVE HISTÓRICO DO RECONHECIMENTO AUTOMÁTICO DE FALA

O reconhecimento de fala está longe de ser uma novidade. Os estudos pioneiros, que estabeleceram a relação entre as características do som e sua representação espectral, datam da década de 1930; em 1952, pesquisadores do Bell Labs desenvolveram um sistema para reconhecimento do som da fala de dígitos isolados; e entre as décadas de 1970 e 1990, os desenvolvimentos do campo rumaram ao reconhecimento de fala contínua (isto é, de palavras emitidas sem pausa significativa entre elas, como em uma conversa normal),[4] envolvendo o reconhecimento de padrões linguísticos e de vocabulários cada vez mais amplos (JUANG; RABINER, 2005).

Significativamente, os setores militar e de inteligência dos EUA estão entre os principais atores da pesquisa e desenvolvimento nessa área (FROOMKIN, 2015); a Darpa investiu abertamente nela desde o começo da década de 1970, e os esquemas revelados por Snowden envolvem uso pesado de reconhecimento de fala pela NSA – que, desde 1999, pelo menos, tem patentes ligadas a reconhecimento de fala. A patente registrada em 1999 foi "descoberta" por Julian Assange (à época apenas um *hacker* australiano pouco conhecido), e envolve técnicas para processar o conteúdo da fala reconhecida, removendo traços de oralidade pouco úteis na fala escrita (frases fáticas, repetições) e – mais importante – identificando temas centrais da conversação (DREYFUS, 1999).[5]

No decorrer dos avanços no campo, houve uma progressiva sofisticação das técnicas de processamento envolvidas, que passaram a incluir a modelagem estatística das linguagens e as redes neurais. Essa sofisticação andou em paralelo com o crescimento contínuo da capacidade de processamento dos dispositivos computacionais disponíveis; as redes neurais, em particular, envolvem requisitos pesados de memória e processamento, e dependem necessariamente da computação paralela. A título de exemplo, em meados da década de 1970, o melhor computador disponível para pesquisadores levava 100 minutos para processar apenas 30 segundos de fala (HUANG; BAKER; REDDY, 2014) – o reconhecimento de fala contínua era possível, mas não em "tempo real", de forma ininterrupta; nas décadas seguintes, porém, a capacidade de processamento cresceria a ponto de possibilitar não só o reconhecimento ininterrupto, mas também para atacar tarefas ainda mais difíceis – como a capacidade de reconhecer a fala em ambientes ruidosos, independentemente de quem está falando (voz, sotaques etc.), e usando mecanismos complexos de aprendizado de máquina (*deep learning*). Esses avanços levaram à aplicação do reconhecimento de fala em dispositivos cada vez mais próximos aos usuários finais: controle por voz em computadores pessoais, carros e TVs, por exemplo – já em 1995, a Microsoft disponibilizava no sistema Windows uma API para reconhecimento de fala em aplicativos.

Além dos efeitos práticos da "Lei de Moore" (o aumento contínuo e exponencial da capacidade dos *chips*, sem aumento correspondente nos seus custos), avanços "qualitativos" no *hardware* computacional também favoreceriam o avanço do

[4] Esse é o sentido técnico de "reconhecimento de fala *contínua*" na literatura da área. O sentido de "reconhecimento de fala *ininterrupto*" explorado neste artigo é mais amplo: não só o reconhecimento da *fala contínua*, mas também o reconhecimento que é conduzido *de forma ininterrupta*.

[5] Chama a atenção o fato de que essa última tarefa (a identificação do tópico de uma conversa por meio de palavras-chave) seja extremamente útil não só para o reconhecimento de fala no contexto da inteligência, mas também na vigilância mercantil voltada à publicidade.

reconhecimento de fala – e um exemplo central no caso analisado aqui é o surgimento e popularização de processadores de múltiplos núcleos. Quando a "Lei de Moore" parecia ter atingido seu limite (pois era cada vez mais difícil aumentar a *performance* de um processador apenas reduzindo o tamanho e aumentando o número de seus transistores), a indústria passou a investir na combinação de diversos núcleos em um mesmo processador; os ganhos de *performance*, então, passavam a ser obtidos por meio da computação paralela – um encaixe perfeito para o reconhecimento de fala via redes neurais.

O SALTO PARA OS DISPOSITIVOS MÓVEIS

Essa virada também implicou na redução da energia necessária e do calor gerado pelos processadores (dois dos principais fatores que limitaram o teto da *performance* em processadores de núcleo único), o que os tornou cada vez mais adequados para dispositivos móveis, uma vez que eles dispõem de fontes de energia e sistemas de refrigeração muito limitados: em linhas gerais, apenas uma bateria e um dissipador diminuto, sem contar com fontes de alimentação elétrica ou ventoinhas (como em PCs e *laptops*).

Ainda assim, os requisitos de processamento eram altos, e o consumo de energia era uma questão difícil de ser resolvida, particularmente em celulares; como descreve uma patente da Google, depositada em setembro de 2012 e concedida no ano seguinte:

> continually operating a microphone and continually running speech detection and recognition applications on a mobile computing device may be undesirable in many situations because the power required to operate the microphone and continually execute the detection and recognition applications can rapidly deplete the mobile computing device's battery (BRINGERT et al., 2013).

Da perspectiva do consumo de bateria (e possivelmente até do controle de temperatura do dispositivo), o reconhecimento de fala ininterrupto era algo proibitivo; na tecnologia descrita nessa patente, por conta disso, ele é ativado apenas quando o celular encontra-se conectado a um carregador.

Nessa época, a Google já investia bastante em reconhecimento de fala; antes da patente em questão, a empresa já havia obtido 40 outras envolvendo reconhecimento de voz.[6] Em 2008, disponibilizou um serviço de busca por voz, o Google Voice Search; e em 2009, lançou o Google Voice (ANDROIDGUYS, 2009), *rebranding* de um serviço de telefonia VoIP adquirido de uma empresa menor, mas que agora oferecia a possibilidade de transcrever mensagens de voz (ferramenta posteriormente integrada ao Gmail).

Esse investimento, porém só aumentaria: em agosto de 2016, a Google já possuía 174 patentes relacionadas a reconhecimento de fala. Em julho de 2012, foi lançado o Google Now – um ambicioso "assistente pessoal" disponível no Chrome e como aplicativo móvel (para Android e iOS), que integra vários produtos da empresa (como

[6] As informações mencionadas aqui se referem a patentes concedidas pela USPTO, o escritório estadunidense de propriedade industrial; foram obtidas pelo sistema de pesquisa da USPTO (Disponível em: <http://patft.uspto.gov>), buscando a expressão "speech recognition" no título, resumo ou reivindicações das patentes concedidas. O ano indicado nas citações é o ano de concessão, e não o de depósito ou solicitação da patente.

o Search e o Gmail, entre outros), além de serviços de informação desenvolvidos especificamente para o Now (indo de reservas de restaurante a avisos de eventos nas redondezas). Uma característica-chave do Now é sua interface baseada em linguagem natural, acessível por reconhecimento de fala (e por digitação).

E aqui saltamos para 2013. Nesse ano, a Google depositaria uma patente (MAANINEN, 2015) bastante distinta da que foi citada acima: propõe uma tecnologia de reconhecimento de fala de funcionamento ininterrupto, com base em um "*hardware* acelerador".

Não por acaso, foi nesse ano que a Google utilizou pela primeira vez (no Moto X, Droid Maxx e Droid Ultra, os primeiros dispositivos lançados pela Motorola após sua aquisição pela Google) um processador em que se tornou viável fazer reconhecimento de voz ininterrupto. Trata-se do Motorola X8 Mobile Computing System: uma espécie de SoC (*system on chip*) que inclui 8 núcleos; quatro para processamento gráfico, dois para aplicativos, e dois especiais, de baixíssimo consumo de energia. Desses dois núcleos especiais, um é dedicado a *contextual computing* (que controla o *display*, a tela de toque e os sensores do dispositivo: acelerômetro, proximidade, luz, giroscópio, temperatura, pressão, bússola) (MOORHEAD, 2013), e outro a reconhecimento de voz (com técnicas bastante avançadas, incluindo correção de ruído e o uso de três microfones simultâneos, no caso do Moto X). Não são núcleos típicos de processadores, mas customizados especificamente para processamento de sinais; a Motorola afirma que são fabricados sob especificações da Texas Instruments (mas é possível que a TI seja a fabricante direta).

O fato de esses núcleos consumirem pouquíssima energia permite que trabalhem continuamente, sem acordar os demais núcleos; ou seja, mesmo quando o telefone está em *standby* – no bolso, na mesa do escritório ou no criado-mudo; segundo a Motorola, sem o X8, as tarefas executadas por esses processadores requereriam o triplo de bateria (SEGAN, 2013). Isso, por sua vez, permite que o celular atenda a comandos de voz em linguagem natural – para o Google Now, por exemplo, que é ativado quando o reconhecimento de voz identifica a frase "Ok Google" sendo pronunciada nas proximidades do celular.[7]

IMPACTOS SOBRE A PRIVACIDADE

O funcionamento desses processadores paralelos não depende de alterações significativas no código do Android (o Moto X de primeira geração usava um Android padrão, embora acrescido dos aplicativos que fazem uso dos comandos de voz). Isso significa que boa parte do código – se não todo o código – que realiza o reconhecimento de fala ininterrupto é código fechado, que reside nos aplicativos proprietários da Google em cada dispositivo (e em parte também, possivelmente, em seus *firmwares*), e nos servidores da empresa. Por conta disso, é muito difícil determinar com certeza o que é feito com a fala que é reconhecida por esses celulares; mais ainda: embora a princípio eles só tenham funções ativadas quando reconhecem a frase "Ok Google", o fato é que eles estão ininterruptamente registrando e processando áudio – e podem muito bem estar realizando

[7] A soma de processamento contínuo de dados de sensores e linguagem natural também é uma combinação que caracterizava o Google Glass – que, como o Google Now, era ativado com o comando de voz "Ok Google".

reconhecimento de fala a todo o tempo, sem que possamos saber como essas informações são utilizadas pela empresa.

O processamento de áudio e voz ininterrupto abre também dois vetores específicos para violações à privacidade. O primeiro consiste na possibilidade de reconhecimento de *voz* (não apenas de *fala*) de maneira individualizada; as características da voz de uma pessoa poderiam funcionar como uma "*soft biometric*", identificando-a sempre que ela falar nas proximidades do celular. Essa identificação pode ter como objeto o usuário principal do telefone (a voz registrada com mais frequência), mas também outras pessoas, o que possibilitaria saber com quem o usuário principal se encontra e conversa. O segundo vetor específico envolve o processamento do áudio ambiental, e permitiria uma série de violações de privacidade relacionadas à localização e ao comportamento. A violação pode ocorrer de forma "passiva", pela comparação entre o ruído ambiental presente e o que foi registrado em uma série de diferentes locais; uma das patentes detidas pela Google é para sistemas que usam áudio geolocalizado, para, entre outras coisas, fazer exatamente isso: "*determining a particular geographic location associated with the particular mobile device*" (KRISTJANSSON; LLOYD, 2012). A violação também pode ocorrer de forma "ativa", por meio da emissão de sons específicos – que podem não ser audíveis ou facilmente identificáveis por humanos – que funcionem como *fingerprints* (identificadores únicos) de áudio; essa técnica já vem sendo utilizada na programação televisiva, por meio da inserção de "assinaturas" sonoras na trilha de um programa: um aplicativo que roda no celular e tenha acesso ao microfone pode identificar o programa sendo assistido, e oferecer publicidade com base nessa informação do indivíduo. A vigilância estatal também é possível por essa forma ativa: em um protesto, por exemplo, a polícia poderia reproduzir (em megafones ou alto-falantes) uma música ou ruído branco com um *fingerprint* sonoro específico; os celulares que reconhecessem tal assinatura identificariam pessoas presentes no protesto (a informação dos identificados poderia ser obtida por meio de colaboração com uma empresa como a Google, pelo uso de aplicativos maliciosos ou outros tipos de infiltração em celulares).

Parte do *marketing* em torno dos modelos de celulares em que surgiu essa tecnologia também girou em torno do Moto Maker: a possibilidade de comprar uma versão personalizada do aparelho (KLUG, 2013), incluindo seleção de cores, gravação física de um texto na superfície externa, e uma linha de texto customizada na tela de inicialização. Esse tipo de customização apenas aumenta a possibilidade – já inerente ao funcionamento dos aplicativos proprietários da Google para o Android – de identificar individualmente a pessoa que usa um dispositivo; mas, no limite, permitiria até que cada indivíduo que optasse por essa customização recebesse um aparelho com *hardware* ou *software* também previamente "customizado", de forma a garantir inequivocamente a identificação entre o dispositivo e o usuário que o comprou (incluindo seus dados pessoais) – possibilidade que poderia ser explorada (com ou sem a anuência da Google) em esquemas mais individualizados de vigilância estatal.

REFLEXÕES FINAIS

O avanço das tecnologias digitais tem trazido inúmeras transformações no cenário da privacidade – inclusive em relação ao que deve ser considerado como objeto de privacidade. As tecnologias ligadas à internet e a dispositivos móveis tornaram insuficientes muitas das tipologias e tentativas de classificação da privacidade. Exemplos de tais tipologias incluem as de Clarke (1997) e Kasper (2005). Clarke propõe as categorias de privacidade da pessoa (ligada a seu corpo, a extração de

tecidos e fluidos etc.), do comportamento pessoal (ligada a questões com potencial de discriminação, como práticas religiosas, sexuais ou políticas), da comunicação pessoal (ligada à comunicação por veículos como *e-mail*, telefone etc., e à ausência de escutas ambientais) e dos dados pessoais. Já a tipologia de Casper refere-se aos tipos de violação de privacidade: violação por extração, observação e intrusão. Tentando equacionar as transformações tecnológicas recentes na discussão da área, um grupo de pesquisadores europeus propôs uma nova tipologia (FINN; WRIGHT; FRIEDEWALD, 2012), que expande a de Clarke para sete categorias: privacidade da pessoa, do comportamento e da ação, das comunicações, dos dados e imagens, dos pensamentos e sentimentos, da localização e espaço, e de associação (a relação com grupos).

O reconhecimento de voz ininterrupto em celulares expõe seus usuários a uma gama muito ampla de violações de privacidade: elas englobam todas as sete categorias de privacidade propostas na mais ampla das tipologias acima, a de Finn, Wright e Friedewald – privacidade da pessoa (por meio da identificação da voz como um *soft biometric*), do comportamento e da ação (por meio, por exemplo, do registro sonoro de atividades religiosas ou parceiros afetivos e/ou sexuais), das comunicações (por meio do registro de conversas e ligações telefônicas), dos dados (registrados de múltiplas maneiras a partir do conteúdo da fala), dos sentimentos (por meio da detecção de estados emocionais identificados por inflexões de voz), da localização (pela comparação dos padrões de ruído ambiental, conforme descrito na patente citada de Kristjansson e Lloyd) e de associação (por meio do reconhecimento da voz de outras pessoas nas proximidades do celular). E como também foi discutido acima, em vários desses casos existe a possibilidade de que ocorra tanto a vigilância mercantil, como a estatal. É necessário considerar com cuidado os riscos envolvidos, e as implicações sociais e políticas implicadas por esse tipo de tecnologia.

O episódio e as posições de Stallman relatado na introdução deste artigo também sugere uma questão importante para reflexão sobre a relação entre privacidade e tecnologias. Algo aconteceu nas últimas décadas: não só passamos a carregar conosco pequenos dispositivos de vigilância, mas aceitamos que eles tenham passado a nos vigiar de forma ininterrupta – e de certa forma, *desejamos* isso, talvez pela praticidade associada a esses dispositivos: o reconhecimento contínuo de fala não é uma característica secreta, sub-reptícia dos celulares com essa tecnologia, mas algo intrínseco ao tipo de funcionalidade que eles oferecem. Ao que parece, estamos diante de mais uma confirmação da hipótese de Gilles Deleuze (1992): de que migramos da sociedade disciplinar, descrita por Foucault, para uma sociedade de controle, em que o monitoramento e a vigilância não são mais garantidos por instituições, mas se entranham mais e mais profundamente no tecido social.

Artigo recebido em 08/07/2016 e aprovado em 07/11/2016.

REFERÊNCIAS

ANDROIDGUYS. *Google voice app arrives in Android market.* 15 jul. 2009. Disponível em: <http://www.androidguys.com/2009/07/15/google-voice-app-arrives-in-android-market/>. Acesso em: 11 ago. 2016.

BRINGERT, Bjorn Erik et al. *Systems and methods for continual speech recognition and detection in mobile computing devices*, US 8.452.597 B2, 28 maio 2013. Disponível em: <http://pdfpiw.uspto.gov/.piw?Docid=08452597>. Acesso em: 10 ago. 2016.

CLARKE, Roger. Introduction to dataveillance and information privacy, and definitions of terms. *Roger Clarke's Website,* 15 ago. 1997. Disponível em: <http://www.rogerclarke.com/DV/Intro.html>. Acesso em: 11 ago. 2016.

DELEUZE, Gilles. Postscript on the societies of control. *October*, v. 59, p. 3-7, 1992. Disponível em: <http://www.jstor.org/stable/778828>. Acesso em: 11 ago. 2016.

DREYFUS, Suelette. Network: this is just between us (and the spies). *The Independent*, Londres, 14 nov. 1999. Disponível em: <http://www.independent.co.uk/arts-entertainment/network-this-is-just-between-us-and-the-spies-1126316.html>. Acesso em: 11 ago. 2016.

FINN, Rachel L.; WRIGHT, David; FRIEDEWALD, Michael. Seven types of privacy. In: GUTWIRTH, Serge et. Al (Ed.). *European data protection:* coming of age. [S. l.]: Springer, 2012. p. 3-32. Disponível em: <http://link.springer.com/chapter/10.1007/978-94-007-5170-5_1>. Acesso em: 10 ago. 2016.

FROOMKIN, Dan. Speech recognition is NSA's best-kept open secret. *The Intercept*, Nova York, 11 maio 2015. Disponível em: <https://theintercept.com/2015/05/11/speech-recognition-nsa-best-kept-secret/>. Acesso em: 11 ago. 2016.

GELLMAN, Barton; SOLTANI, Ashkan. NSA infiltrates links to Yahoo, Google data centers worldwide, Snowden documents say. *The Washington Post*, 30 out. 2013 Disponível em: <https://www.washingtonpost.com/world/national-security/nsa-infiltrates-links-to-yahoo-google-data-centers-worldwide-snowden-documents-say/2013/10/30/e51d661e-4166-11e3-8b74-d89d714ca4dd_story.html>. Acesso em: 11 ago. 2016.

GOOGLE. *2014 financial tables*: investor relations. Disponível em: <http://investor.google.com/financial/tables.html>. Acesso em: 19 abr. 2014.

HOLWERDA, Thom. *Richard Stallman was right all along*. OS News, 2 jan. 2012. Disponível em: <http://www.osnews.com/story/25469/Richard_Stallman_Was_Right_All_Along>. Acesso em: 11 ago. 2016.

HUANG, Xuedong; BAKER, James; REDDY, Raj. A historical perspective of speech recognition. *Communications of the ACM*, v. 57, n. 1, p. 94–103, 1 jan. 2014. Disponível em: <http://dl.acm.org/citation.cfm?doid=2541883.2500887>. Acesso em: 11 ago. 2016.

JOHNSON, Bradley. 10 things You should Know about the global ad market. *Advertising Age*, 8 dez. 2013. Disponível em: <http://adage.com/article/global-news/10-things-global-ad-market/245572/>. Acesso em: 19 abr. 2014.

JUANG, B. H.; RABINER, L. R. Automatic speech recognition: a brief history of the technology development. In: ELSEVIER encyclopedia of language and linguistics. Amsterdam: Elsevier Science, 2005. Disponível em: <http://www.idi.ntnu.no/~gamback/teaching/TDT4275/literature/juang_rabiner04.pdf>.

KASPER, Debbie V. S. The evolution (or devolution) of privacy. *Sociological Forum*, v. 20, n. 1, p. 69-92, mar. 2005. Disponível em: <http://doi.wiley.com/10.1007/s11206-005-1898-z>. Acesso em: 10 ago. 2016.

KLUG, Brian. Moto X review. *AnandTech*, 26 ago. 2013. Disponível em: <http://www.anandtech.com/show/7235/moto-x-review>. Acesso em: 26 jun. 2015.

KRISTJANSSON, Trausti; LLOYD, Matthew I. *Geotagged environmental audio for enhanced speech recognition accuracy*. US 8.265.928 B2, 11 set. 2012. Disponível em: <http://pdfpiw.uspto.gov/.piw?Docid=08265928>. Acesso em: 11 ago. 2016.

MAANINEN, Juha-Pekka. *Mobile speech recognition hardware accelerator*, US 9.153.230 B2, 6 out. 2015. Disponível em: <http://pdfpiw.uspto.gov/.piw?Docid=09153230>. Acesso em: 10 ago. 2016.

MACASKILL, Ewen. NSA paid millions to cover Prism compliance costs for tech companies. *The Guardian*, 23 ago. 2013. Disponível em: <https://www.theguardian.com/world/2013/aug/23/nsa-prism-costs-tech-companies-paid>. Acesso em: 11 ago. 2016.

MOORHEAD, Patrick. Motorola's "X8 Computing System", brought to You by Qualcomm and Texas Instruments? *Forbes*, 28 ago. 2013. Disponível em: <http://www.forbes.com/sites/patrickmoorhead/2013/08/28/motorolas-x8-computing-system-brought-to-you-by-qualcomm-and-texas-instruments/>. Acesso em: 26 jun. 2015.

RAYMOND, Eric. Why I think RMS is a fanatic, and why that matters. *Armed and Dangerous*, 2012. Disponível em: <http://esr.ibiblio.org/?p=4386>. Acesso em: 10 ago. 2016.

ROSENBACH, Marcel; POITRAS, Laura; STARK, Holger. iSpy: how the NSA accesses smartphone data. *Der Spiegel*, 9 set. 2013. Disponível em: <http://www.spiegel.de/international/world/how-the-nsa-spies-on-smartphones-including-the-blackberry-a-921161.html>. Acesso em: 11 ago. 2016.

SEGAN, Sascha. Motorola reveals more X8 chip details. *PC Magazine*, 30 jul. 2013. Disponível em: <http://www.pcmag.com/article2/0,2817,2422513,00.asp>. Acesso em: 26 jun. 2015.

STALLMAN, Richard. Engineering tech talk at Google. *Gnu.org*, 11 jun. 2004. Disponível em: <https://www.gnu.org/philosophy/google-engineering-talk>. Acesso em: 10 ago. 2016.

_____. Political notes from 2008: March-June – FBI uses cellphones to eavesdrop conversations. *Richard Stallman's Personal Site*, 11 mar. 2008. Disponível em: <https://stallman.org/archives/2008-mar-jun.html#11%20March%202008%20(FBI%20uses%20cellphones%20to%20eavesdrop%20conversations)>. Acesso em: 11 ago. 2016.

_____. Someone to watch over me. *Richard Stallman's Personal Site*, 2002. Disponível em: <https://stallman.org/watch-over-me.html>. Acesso em: 11 ago. 2016.

VIEIRA, Miguel Said. *Os bens comuns intelectuais e a mercantilização*. 2014. Tese (Doutorado em Educação) – Faculdade de Educação da Universidade de São Paulo, São Paulo, 2014. Disponível em: <http://www.teses.usp.br/teses/disponiveis/48/48134/tde-01102014-104738/>.

VIEIRA, Miguel Said; EVANGELISTA, Rafael de Almeida. A máquina de exploração mercantil da privacidade e suas conexões sociais. In: SIMPÓSIO INTERNACIONAL LAVITS: vigilância, tecnopolíticas, territórios, 3., Rio de Janeiro, 2015. *Anais*... Rio de Janeiro: Lavits, 2015. Disponível em: <http://lavitsrio2015.medialabufrj.net/anais/>. Acesso em: 29 abr. 2016.