

Big Data e política: aferindo o potencial preditivo da *world wide web* nas eleições europeias de 2014

Big Data and politics: determining the predictive potential of the world wide web in the European elections of 2014

Luísa Godinho*

*Professora, Universidade Autónoma de Lisboa

Resumo

Inversamente à tendência dominante nos estudos eleitorais previsionais em ambiente digital, focados no microblogue Twitter como universo de análise, neste artigo desenvolve-se uma abordagem ecológica integrada ao posicionamento da opinião pública portuguesa durante a campanha eleitoral para as eleições europeias de 2014. O potencial preditivo de três categorias é testado – partido/coligação; cabeça-de-lista e lista – assim como o é a representatividade da *world wide web*, enquanto Big Data cientificamente válida para a investigação em Ciências Sociais. A primeira categoria revelou maior valor empírico relativo e a adequabilidade das redes sociais para a antecipação de acontecimentos sai parcialmente reforçada, reiterando-se, no entanto, a necessidade de mais estudos, em particular no plano nacional, onde a investigação nesta área é praticamente inexistente.

Keywords: Big Data, Internet, eleições, opinião pública, posicionamento, previsões eleitorais.

Abstract

Inversely to the dominant trend in predictive electoral studies in a digital environment, focused on the microblogue Twitter as a universe of analysis, this article develops an integrated ecological approach to the positioning of Portuguese public opinion during the electoral campaign for the European elections of 2014. The predictive potential of three categories is tested - party / coalition; head-of-list and list - as well as is the representativeness of the *world wide web*, considered a Big Data scientifically valid for Social Sciences research. The first category revealed a higher relative empirical value and the suitability of social networks for the anticipation of events is partially reinforced. However, there is a need for further studies, especially at the national level, where research in this area is practically non-existent.

Keywords: Big Data, Internet, elections, public opinion, positioning, electoral forecasting.

Introdução

Em ótica estritamente comunicacional, candidatos políticos e eleitores, atores centrais dos regimes democráticos, são entendidos no seu duplo papel de emissores-recetores que trocam entre si signos passíveis de gerar o modelo habermasiano de deliberação (Habermas, 1962), processo que se crê ter vindo a transferir-se, desde finais do século XX, dos chamados media tradicionais para a *world wide web* (Koop & Jansen, 2005; Thimm, 2012), elevando esta última à categoria daquilo que Donath designou, não sem alguma propriedade, por *social machine* (Donath, 2014).

Hoje, mais de três biliões de pessoas em todo o mundo têm acesso à internet e este facto passou forçosamente a integrar a reflexão sobre as democracias contemporâneas, dependentes, mais do que das preferências dos eleitores, dos processos de formação da opinião que as antecedem e em que os media digitais participam de forma ativa e determinante. (Gimmler, 2001; Dahlberg, 2007; Rasmussen, 2009; Thimm, 2012)

Esta nova comunicação digital de massas, que Castells designou de *mass-self* (Castells, 2009), transferiu para a web uma parte que se crê substancial da formação da opinião individual, crucial tanto para a afirmação do valor da liberdade como para a sobrevivência dos próprios regimes democráticos (Polat, 2005; Best & Wade, 2009), gerando um novo ramo de investigação social interessado em compreender a forma como os eleitores participam no debate político online e decidem o voto.

A hipótese de a *world wide web* constituir um novo espaço de formação da opinião e de deliberação assume-se, por isso, como um tema centrípeto para investigadores de todo o mundo, sobretudo à medida que a tecnologia permitiu abordagens verdadeiramente inovadoras às problemáticas das ciências sociais, entre as quais a possibilidade de previsão dos resultados eleitorais a partir das trocas semânticas realizadas online. No presente estudo, esta última questão será explorada, propondo-se analisar o posicionamento eleitoral da opinião pública portuguesa expresso na *world wide web* no mês anterior às eleições europeias de 2014. O objetivo será determinar a relação eventualmente existente entre texto online e número de votos, em particular se a métrica do volume de referências, designada de *buzz*, é adequada a cumprir este desiderato. De modo a fazer avançar este tipo de pesquisas, normalmente centradas na categoria "partido" e, em raríssimos casos, na categoria "coligação" (Tumasjan et al., 2010), aqui se propõe o teste de duas novas categorias de análise - "lista" e "cabeça-de-lista" – cujo comportamento será avaliado mediante a sua correspondência com os resultados eleitorais.

Estes objetivos serão cumpridos através da resposta a quatro questões de investigação que passamos a enunciar:

Questão nº1

O volume de referências na web a um partido político ou coligação concorrente a um ato eleitoral pode ser considerado equiparável ao número de votos que virão a ser recolhidos por cada um dos candidatos?

E, em caso positivo, qual o volume de referências que apresenta maior valor preditivo: o volume médio do período em análise ou o verificado durante a semana que antecede as eleições?

Questão nº2

No caso das coligações, qual a métrica que apresenta maior valor preditivo: o volume de referências ao nome da coligação ou o somatório do *buzz* das forças políticas que a integram?

Questão nº3

Relativamente à lista apresentada por cada força política candidata, poderá o volume de referências aos elementos que a integram ser indicativo dos resultados eleitorais?

Questão nº4

Quanto à presença online dos cabeças-de-lista, qual a importância relativa de cada um, atendendo à notoriedade expressa pelo volume de referências? Será que o *buzz* dos cabeças de lista candidatos se equipara aos resultados eleitorais da força política que lideram?

A resposta a estas questões de investigação impõe o desenvolvimento de uma metodologia própria, apresentada no ponto 3, bem como o aprofundamento da questão teórica subjacente, relativa à conceção da *world wide web* como uma Big Data e ao seu contributo para a investigação em Ciências Sociais. Como veremos adiante, estes temas corporizam atualmente um importante ramo de investigação em particular no domínio da Comunicação Política.

A world wide web como Big Data

Para as ciências sociais, uma das valências mais interessantes da *world wide web* é enquanto novo campo de estudo, um grande texto aberto onde o sentido vai sendo expresso livremente. Sobretudo desde a primeira década do século XXI, quando a massificação do acesso online se efetivou, que este grande texto tem vindo a crescer exponencialmente, estimando-se que o universo digital, entendido como os dados digitais criados e copiados anualmente, duplique de tamanho a cada dois anos e atinja, no ano 2020, os 44 trilhões de gigabytes (Halfpenny & Procter, 2015).

Tamanhos volume e variedade de informação processada a uma velocidade tão expressiva obrigaram à criação de métodos de estudo e de gestão específicos, assim como a uma abordagem conceptual que permitisse passar do texto ao mega-texto, ou seja de uma unidade linguística até aqui passível de ser analisada manualmente por métodos tradicionais para uma base de dados cujas dimensão, variedade e ritmo de crescimento inviabilizam o seu estudo não mediado. Esta última perspetiva de estudo foi designada de *text as data* (Grimmer & Stewart, 2013; Benoît, et al 2013; Benoît, 2016;) e, segundo ela, o texto passa a ser entendido como um conjunto de dados, ou seja, "informação que pode traduzir-se por números, texto, som ou imagem (...), possui uma estrutura e passou por uma qualquer forma de processamento" (Halfpenny & Procter, 2015).

Uma conceção tão positivista do texto levanta inevitáveis objeções por parte das escolas fenomenológica, interpretativista e hermenêutica, cuja abordagem assenta no campo diametralmente oposto, centrado na polissemia da palavra e na riqueza interpretativa da própria receção, mas torna-se hoje difícil negar as potencialidades que a tecnologia aportou à análise textual, sobretudo quando estão em causa corpus cujas dimensão e complexidade impediriam a realização do próprio estudo não computadorizado.

Assumir a internet como uma base de dados de acesso aberto remete para o campo das tecnologias da informação, domínio científico que tradicionalmente se tem ocupado do tema e que nas últimas décadas se especializou no estudo de um tipo particular de bases de dados, aquelas cujas dimensão, variedade e velocidade de crescimento obrigam à utilização de novas técnicas de análise e de gestão. A estas bases de dados que desafiam as tradicionais técnicas de estudo, a ciência designou de Big Data e o conteúdo da *world wide web* constitui desta um dos exemplos mais expressivos (Halfpenny & Procter, 2015).

A Big Data trouxe às Ciências Sociais um conjunto de novas possibilidades assentes na adoção de modelos computacionais que acabaram por gerar um novo campo de estudo designado de Ciências Sociais Computacionais.

Este tipo de análises tem incidido sobre quatro distintos espectros de abordagem (Chang et al., 2013): macro, meso, micro e nano. O macro espectro corresponde à escala mais alargada, permitindo a identificação de padrões de interação nacionais, regionais ou setoriais; o meso-espectro corresponde a uma escala intermédia de análise, grupal, como seja os comportamentos de um determinado grupo de consumidores; o micro espectro corresponde à escala individual e o nano espectro à escala psicológica ou emocional.

Segundo outra perspetiva, relativa à atualização dos dados, os cientistas sociais têm ao seu dispor dois tipos de dados, cabendo-lhes selecionar o mais apropriado para a investigação: stocks e fluxos. Os stocks de dados consistem em tabelas provenientes de bases de dados existentes e que podem ser atualizadas em diferentes momentos (por exemplo, tabelas salariais que são atualizadas anualmente); os fluxos de dados,

por sua vez, são dados atualizados em permanência (por exemplo, os *clicks* de utilizadores de internet num *banner* publicitário). (Chang et al., 2013)

Big Data e Ciências Sociais

A Big Data constitui para a ciência moderna simultaneamente uma oportunidade e um desafio, embora o grau de utilização que as ciências sociais, em particular, atualmente dela fazem pareça ainda muito diminuto face às potencialidades antevistas. Estaremos porventura perante aquilo que Kuhn (Kuhn, 1962) designou de «mudança de paradigma», já que os ganhos científicos aportados pela Big Data, produto da mudança tecnológica, da convergência interdisciplinar, da analítica de dados e de novas práticas e ambientes organizacionais (Chang et al., 2013), permitem fortalecer as possibilidades de «generalização, controle e realismo» da investigação (Chang et al., 2013), assumindo-se como alternativa aos métodos tradicionais. Atualmente, podem ser identificadas cinco linhas de investigação em Big Data no campo das Ciências Sociais: uma linha focada na utilização da Big Data para o aprofundamento do estudo do contexto; uma segunda linha dedicada a explorar o potencial da Big Data para o melhor conhecimento de processos como os da personalização e da hiper-diferenciação e sua aplicação ao domínio da Gestão; uma terceira linha que visa utilizar a Big Data para o estabelecimento de padrões, a formulação de explicações e a previsão de acontecimentos; uma quarta linha dedicada a conhecer os impactos do comportamento humano nos mercados e uma quinta linha voltada para o estudo de fenómenos como os da privacidade e dos direitos (Chang et al., 2013).

A primeira linha de investigação visa aplicar a Big Data às áreas da Gestão, permitindo um melhor conhecimento nomeadamente do contexto em que o consumidor realiza as suas escolhas.

A segunda linha de investigação dedica-se a conhecer os processos de personalização e de hiper-diferenciação, aproveitando “a extraordinária quantidade de dados que podem, hoje, ser recolhidos acerca do consumidor” e utilizando-os para produzir “mudanças passíveis de garantir a maximização do valor quer na perspetiva do consumidor quer na do vendedor” (Chang et al., 2013: 11). Abordagens experimentais nesta área, com base em processos de recolha de dados por computador, de monitorização da decisão e de geo-localização de dados, têm permitido chegar a importantes inovações no design de produtos e serviços. A terceira linha de investigação dedica-se a aplicar os métodos de estudo da Big Data à análise de acontecimentos (event-focused analysis) bem como a produzir informação comportamental e económica ao nível individual, grupal e de rede a partir de dados recolhidos digitalmente. Este tipo de análises tem permitido a identificação de padrões e a produção de explicações e, aliado às análises previsionais, tornou possível a previsão de acontecimentos (Campbell, 2000; Clements & Hendry, 2011; Gelman, 2012; Bollen et al., 2010; Asur & Huberman, 2010; Gayo-Avello, 2012).

A quarta linha de investigação visa compreender os impactos do comportamento humano nos grupos sociais e nos mercados a partir dos novos mecanismos de recolha de dados. Exemplos de estudos são “o comércio e os leilões eletrónicos, os mercados digitais, os volumes de transação por cartão de crédito, os fluxos de tráfego e de transporte e a opinião” dos utilizadores (Chang et al., 2013: 11).

Uma quinta e última linha de investigação foca-se no estudo dos desafios à privacidade e aos direitos individuais criados pelo advento da Big Data.

Não obstante a variedade de problemáticas em estudo, a investigação em Big Data, a partir de dados de proveniências tão variadas como “os leilões online, as comunicações nas redes sociais, as redes de telemóvel, as plataformas de videojogos, os motores de pesquisa e os blogues” (Chang et al., 2013: 6), parece particularmente votada a abordagens multidisciplinares, sobretudo quando está em causa a realização de inferências no âmbito da chamada *business intelligence*.

A adequação desta multidisciplinaridade ao estudo da Big Data advém sobretudo do carácter interdisciplinar dos próprios dados, passíveis de gerar informação muito diversificada consoante a natureza das abordagens desenvolvidas.

“Alguém com uma especialização em Sociologia e em Psicologia pode não se aperceber das oportunidades de conhecimento” presentes numa Big Data se não utilizar técnicas de mineração de dados ou de machine learning, por exemplo (Chang et al., 2013: 6), o mesmo acontecendo nos campos da estatística e da economia.

A multidisciplinaridade no estudo de Big Data em Ciências Sociais incide nos planos teórico e metodológico e apresenta algumas limitações.

No plano teórico, verifica-se uma desadequação entre as teorias existentes e as problemáticas que carecem de explicação na análise de Big Data.

“As teorias mais relevantes provêm da Sociologia, das Ciências da Comunicação, da ciência cognitiva e da Psicologia Social. Mas saber de que modo se formam as relações entre participantes é do domínio da Economia de rede, da investigação em operações teórico-gráficas, computação e contabilidade. Muito do que está a acontecer pode ser descrito por medição através de representações abstratas como gráficos (...) ou tabelas” (Chang et al., 2013: 7).

A medição, de resto, constitui um desafio significativo no estudo de Big Data em Ciências Sociais. O problema advém do carácter inovador e diversificado dos aspetos a serem medidos, já que boa parte das questões que requerem medida, como sejam “a força das reações e comunicações humanas, a posição e a distância entre entidades numa rede, os custos da comunicação, as mudanças na opinião social, as alterações nas interações entre organizações e indivíduos, entre outras”, são inéditas em Ciências Sociais e requerem abordagens e métricas provenientes de vários campos do saber.

Em segundo lugar, existe ainda a desadequação dos métodos e das métricas tradicionais das Ciências Sociais ao estudo de Big Data.

“Os dados podem não apresentar as propriedades de amostragem» frequentes nas Ciências Sociais, para além de requererem novos métodos que permitam retirar sentido das relações causais observadas. “Parece, por isso, apropriado”, escrevem Chang et al., “desenvolver uma abordagem interdisciplinar e multi-perspetiva que permita chegar de fato aos resultados mais interessantes” (Chang et al., 2013: 6).

As questões metodológicas encontram-se entre os maiores desafios colocados hoje à utilização da Big Data no domínio das Ciências Sociais, existindo mesmo quem considere estarmos verdadeiramente perante um método novo alicerçado em novas estratégias de observação.

Esta tese tem por base quatro argumentos fundamentais: em primeiro lugar, considera-se que a Big Data permite diluir a distinção tradicionalmente existente entre os estudos de campo e os estudos laboratoriais, tornando desnecessária

“a criação de ambientes artificiais em laboratório que (...) mimetizam o mundo real num ambiente sintetizado. Pelo contrário, recolhe-se informação que representa os elementos fundamentais das ações e interações humanas (...) diretamente do mundo real como traços digitais do comportamento humano.”
(Chang et al., 2013);

em segundo lugar, crê-se que a Big Data aumentará significativamente aquilo a que os autores designam de realismo na pesquisa experimental, ao permitir um alargamento do espectro de investigação ao nível da dimensão das populações estudadas assim como dos temas em análise; em terceiro lugar, desafia a centralidade da teoria no processo científico, já que a análise em Big Data coloca o ênfase na qualidade dos dados recolhidos e dos métodos de análise e não na qualidade das teorias formuladas, alguns autores considerando estarmos perante uma abordagem de “grounded theory (...) mas suportada por Big Data e envolvendo a utilização de análises preditivas” (Chang et al., 2013); o quarto e último argumento assenta na ideia de que a utilização de Big Data em ciências sociais permite limitar consideravelmente o carácter intrusivo de algumas pesquisas ao mesmo tempo que garante maior precisão nos dados recolhidos e maior controle da parte do investigador.

Política e texto online

A utilização da *world wide web* enquanto Big Data socialmente representativa constitui uma das mais recentes correntes metodológicas no domínio das ciências sociais, seguindo o pressuposto de que a web constitui um espaço de representação discursiva cujo estudo permite a compreensão dos indivíduos, dos grupos e das interações sociais, bem como o estabelecimento de inferências cuja natureza varia consoante os domínios do conhecimento.

No domínio da Comunicação Política, que aqui nos ocupa, o estudo do texto online tem-se desenvolvido em três principais linhas de investigação: uma linha focada na monitorização dos discursos; uma segunda linha interessada na análise de sentimentos e reações a esses discursos e uma terceira linha orientada para a análise previsional a partir dos textos e das ações presentes na *world wide web* (Ruas-Araújo et al., 2016). A primeira linha de investigação dedica-se à análise dos discursos político-partidários online, particularmente os produzidos no âmbito de uma estratégia de comunicação política que inclua os media digitais entre os seus instrumentos de trabalho, utilizando-os como suporte de campanha e meio de interação direta com o eleitorado (Strandberg, 2013; Gibson, 2013; Jungherr, 2014). Apesar de frequente nos Estados Unidos da América, esta utilização é ainda limitada na Europa, assumindo aqui variações expressivas entre os países (Hansen & Kosiara-Pederson, 2014; Jungherr, 2014).

A segunda linha de investigação foca-se no estudo da receção da mensagem política e tem na análise de redes sociais um importante domínio de investigação, já que estas se constituem, hoje, como o espaço

primordial de reação do eleitorado às mensagens dos políticos. A participação das redes sociais no processo de deliberação política parece constituir, de resto, um dado consensual no atual panorama científico. Ruas-Araújo et al., por exemplo, reconhecem

“(...) existir um certo consenso sobre a contribuição das redes sociais na ativação do chamado capital social (Gil de Zúñiga et. al, 2011) e a modificação das condutas das pessoas que as utilizam, assim como a constatação de que as redes sociais começam a ocupar – ou, pelo menos, a partilhar – o espaço dos cibermeios de comunicação como narradores da realidade (García et al., 2012), tarefas até há pouco tempo limitadas exclusivamente ao domínio do jornalismo tradicional». (Ruas-Araújo et al., 2016)

No mesmo sentido manifestam-se Nulty et al. (Nulty et al., 2015: 4), ao considerarem que a utilização das redes sociais Facebook e Twitter permitiu nomeadamente aos candidatos “aumentar a sua exposição a muito baixo custo (...), o que ajudou candidatos menos conhecidos a saírem da escuridão”, para além de terem servido de “plataforma para os cidadãos comunicarem diretamente com os candidatos políticos”.

A terceira linha de investigação dedica-se ao estudo da produção e da receção da mensagem política online para, a partir destas, procurar desenvolver análises previsionais do sentido de voto do eleitorado. Esta área de estudos preditivos parte do pressuposto de que

“indivíduos racionais são capazes de tomar, em conjunto, uma decisão quase ótima, muitas vezes superando a que resulta da inteligência individual, o que quer dizer que a multidão, entendida como um ser inteligente, é mais dotada do que a maioria das suas componentes humanas entendidas singularmente” (Franch, 2010).

Esta sabedoria das multidões, como já lhe chamaram, depende, no entanto, como bem notou Surowiecki (Surowiecki, 2004), da diversidade e da independência dos seus atores, assim como de uma estrutura descentralizada, características que se aproximam fortemente das da internet. É justamente esta similitude que legitima os estudos preditivos a partir do universo online.

Nesta corrente de investigação podem ser hoje identificadas duas correntes: uma entusiasta quanto à validade preditiva das redes sociais (Tumasjan et. al, 2010; Asur & Huberman, 2010; O’Connor et. al, 2010; Tjong & Bos, 2012) e outra crítica (Goldstein & Rainey, 2010; Panagiotis, 2011; Gayo-Avello, 2011 e Jungherr, 2011).

Relativamente à corrente entusiasta, um conjunto de estudos realizados em vários países europeus, como são os casos da Grã-Bretanha, da França e da Alemanha, tem vindo a confirmar a validade dos media sociais como indicadores do posicionamento do eleitorado, acreditando no potencial preditivo destes.

A investigação demonstrou que não apenas o número de referências online aos candidatos políticos parece ser indicativo dos resultados eleitorais (Tumasjan et. al, 2010; Asur & Huberman, 2010; O’Connor et. al, 2010; Tjong & Bos, 2012) como também que as menções conjuntas a determinadas forças políticas parecem indicar a relevância social de atores como coligações e alianças eleitorais (Tumasjan et al., 2010).

Ruas-Araújo et al. precisam que

“(...) estudos anteriores defendiam já que, no caso do Facebook, o número de seguidores poderia considerar-se como um indicador válido de êxito eleitoral (Williams & Gulati, 2008), outros indicavam que tanto o Twitter como o Facebook eram previsores fiáveis e precisos (Carr, 2010) e outros contemplavam sobretudo o debate na internet como um veículo para a deliberação política (Koop & Jansen, 2009)”. (Ruas-Araújo et al., 2016)

A corrente crítica, no entanto, tem-se multiplicado em conclusões de sentido contrário, demonstrando que

“os partidos mais votados nem sempre eram os que registavam mais comentários ou atividade online (Bermingham & Smeaton, 2011). Pelo contrário: que as formações menos maioritárias eram as que mais utilizavam esta rede social. Um facto diretamente relacionado com a sua menor capacidade e maior custo de acesso aos meios tradicionais. Os autores deste trabalho também observaram que, mais do que os resultados de tipo quantitativo, o verdadeiro valor preditivo do Twitter – se é que existe algum – radicaria na observação qualitativa e não no volume de tweets ou comentários publicados”. (Ruas-Araújo et al., 2016)

A abordagem quantitativa que caracteriza a maioria destas análises previsionais assenta na contagem do número de menções a cada candidato político, aproximando-se, como nota Schoen et al. (Schoen et al, 2013), dos estudos com base em inquéritos.

Nos questionários eleitorais, os inquiridos são instados a indicar o partido ou candidato em quem pensam votar. Depois, essas escolhas são contabilizadas e a proporção de menções é interpretada como uma estimativa do número de votos que cada candidato receberá. Do mesmo modo, nas previsões a partir dos media sociais, as menções aos partidos e aos candidatos são contabilizadas e a proporção de menções é interpretada como sendo uma estimativa da percentagem de votos que cada partido ou candidato receberá. Não obstante a similitude entre os dois métodos de recolha de dados, algumas diferenças notórias recomendam prudência na equiparação, já que, contrariamente aos inquiridos, os utilizadores de internet não são interpelados no sentido de exprimirem as suas preferências, antes manifestam-se livremente online, podendo mencionar, elogiar, questionar ou criticar um partido ou candidato. As menções detetadas por qualquer software de recolha de referências online possuem, por isso, uma grande variedade de sentidos, não podendo ser associadas ao apoio a forças políticas.

Aberto o debate sobre esta hipótese de correspondência entre os mundos online e offline, um conjunto de questões surge para a investigação científica na área da comunicação política, o presente artigo situando-se justamente neste domínio, visando estudar o papel dos media digitais no processo eleitoral da União Europeia, em particular o caso das eleições portuguesas de 2014 para o Parlamento Europeu.

Interessa aqui testar a validade dos media digitais enquanto instrumento de aferição da vontade popular, propondo-se para tal desenvolver um estudo comparativo entre a notoriedade na web dos candidatos

portugueses às eleições europeias de 2014 e os resultados eleitorais. A demonstração de proximidade entre a notoriedade web dos candidatos e a distribuição dos votos permitirá não apenas corroborar a corrente favorável à capacidade preditiva eleitoral das redes sociais como também reforçar a hipótese de correspondência entre as realidades online e offline; a demonstração de distanciamento entre essas duas variáveis levará à conclusão contrária. Em qualquer caso, contribui-se para colmatar a necessidade de investigações neste domínio que, tal como escreve Ruas-Araújo et al., seguindo aliás a opinião da generalidade dos investigadores desta área, “contribuam para valorar o nível de predição das redes sociais” (Ruas-Araújo et al., 2016).

Metodologia: opção por uma abordagem de ecologia integrada

Do ponto de vista científico, a *world wide web* constitui um vasto universo de análise passível de ser analisado através de distintas abordagens, das quais duas se destacam: a abordagem modal e a abordagem ecológica integrada (Tufekci, 2014). A primeira, de tipo sincrónico, consiste na seleção de um único organismo para análise e no seu estudo aprofundado, sendo um bom exemplo o caso do estudo da rede social digital Twitter; a segunda, de tipo diacrónico, consiste na construção de uma amostra dos vários organismos existentes no ecossistema e no seu estudo alargado, sendo um bom exemplo os estudos de Franch (Franch, 2013).

Apesar de a abordagem modal ser dominante nas pesquisas nesta área, focando quase sempre a rede social Twitter devido à facilidade de acesso aos dados e à tecnologia, ela apresenta um conjunto de limitações passíveis de colocar em causa a validade científica dos estudos. Estas limitações são particularmente expressivas no caso de pesquisas que utilizam o Twitter como amostra do tecido social, já que tendem a tecer generalizações a partir de características específicas desta rede social, como sejam tipos particulares de conteúdos, de interação e de população.

No entanto, tal como recorda Tufekci, “estes desafios não significam que as análises a uma única plataforma não sejam válidas. Contudo, estas devem ter presente que não estão a examinar um sistema fechado e que podem existir efeitos que não sejam visíveis porque a informação relevante não está contida na plataforma” (Tufekci, 2014).

Tendo em conta a necessidade de encontrar uma amostra mais representativa do tecido social, na presente pesquisa optou-se por recusar a análise exclusiva do universo Twitter e desenvolver uma abordagem de ecologia integrada que reunisse uma variedade de organismos online, tais como as redes sociais Twitter, Facebook e Google+, sítios de órgãos de comunicação social, blogues, entre outros.

Uma tal amostra, por possuir uma composição mais diversificada, representa diferentes tipos de interação, de textos e de utilizadores, passíveis de melhor refletir a diversidade de comportamentos presente no universo online e, como tal, de possuir um grau mais elevado de representatividade do universo em estudo. Definido o modelo da abordagem, as etapas seguintes foram a seleção do software e a identificação das categorias de análise.

O software selecionado para realizar a pesquisa foi uma plataforma SaaS, de coleta, monitoramento e análise de texto colocado online proveniente de mais de 150 milhões de fontes em todo o mundo - tais como sites

noticiosos, blogs e fóruns e redes sociais como o Facebook, o Google+, o Foursquare, o Instagram e o Twitter.

O acesso a estas fontes é variável em extensão: no caso dos sites dos media tradicionais, analisa todos os existentes em Portugal; no caso dos blogs e fóruns, analisa uma amostra selecionada aleatoriamente; no caso do Twitter, a plataforma alcança todos os seus dados; no caso do Facebook, do Google+, do Foursquare e do Instagram analisa uma amostra selecionada aleatoriamente.

A análise neste dispositivo pode ser gerida mediante a seleção dos filtros considerados pertinentes para o estudo, como sejam o tipo de media, o país, a língua e a data.

Na presente investigação, dado ter-se optado por uma abordagem de ecologia integrada ao posicionamento político-partidário do eleitorado português durante o mês anterior à data das eleições europeias de 2014, definiram-se os seguintes parâmetros de análise:

Tipo de media: todos

País: Portugal

Língua: português

Métrica: buzz (número de menções)

Período de análise: de 24 de abril a 24 de maio de 2014

Quanto às categorias de análise, estas foram construídas de acordo com os objetivos do estudo, ou seja, num primeiro momento monitorizar a notoriedade dos candidatos portugueses no mês anterior às eleições europeias de 2014 e, num segundo momento, através de uma análise comparativa com os resultados eleitorais, aferir o grau de proximidade entre os dois valores, ou seja, o índice de preditibilidade da métrica buzz.

Três categorias foram criadas: a categoria "Partido/coligação", que inclui os partidos e coligações candidatos; a categoria "Cabeça de lista", relativa aos cabeças de lista de cada candidatura e a categoria "Lista", relativa aos primeiros dez nomes que compõem cada candidatura, excetuando o cabeça de lista.

Definidas as categorias de análise, passou-se à fase da codificação, que permitiu a construção da seguinte matriz de análise:

Tabela 1: Matriz de análise para o estudo do posicionamento da opinião pública portuguesa expressa na world wide web

Categorias	Indicadores (keywords)
Partidos e coligações	Bloco de Esquerda (BE)
	Coligação Democrática Unitária (CDU)
	Partido Socialista (PS)

		Aliança Portugal (AP)
Listas	Lista AP	Fernando Ruas
		Sofia Ribeiro
		Nuno Melo
		Carlos Coelho
		Cláudia Aguiar
		José Manuel Fernandes
		José Mendes Bota
		Fernando Costa
	Lista P S	Maria João Rodrigues
		Carlos Zorrinho
		Elisa Ferreira
		Ricardo Serrão Santos
		Ana Gomes
		Pedro Silva Pereira
		Liliana Rodrigues
Manuel dos Santos		
Cabeças-de-lista	Paulo Rangel (AP)	
	Marisa Matias (BE)	
	João Ferreira (CDU)	
	Francisco Assis (PS)	

Seguindo os estudos preditivos já realizados, a presente análise incidiu exclusivamente sobre o volume de menções (também designado de buzz), tendo abdicado expressamente de qualquer análise qualitativa. Trata-se, por conseguinte, de uma macro-análise de fluxos (Chang et al., 2013) que contabiliza o número de referências a cada categoria de análise publicadas online durante o período em estudo, permitindo conhecer o posicionamento do universo nacional português.

O período de análise escolhido foi o compreendido entre 24 de abril e 24 de maio de 2014, mês anterior às eleições europeias de 2014, tendo este sido dividido em sub-períodos semanais que permitissem identificar eventuais flutuações na notoriedade de cada candidato.

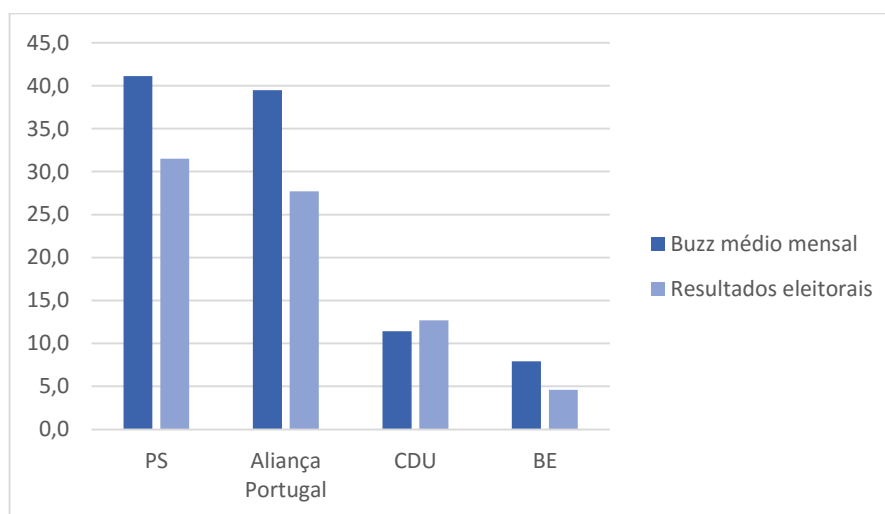
Resultados

Questão nº1

- a) O volume de referências na web aos partidos políticos ou coligações concorrentes às eleições europeias de 2014 pode ser considerado equiparável ao número de votos que virão a ser recolhidos por cada um dos candidatos?
- b) E, em caso positivo, qual o volume de referências que apresenta maior valor preditivo: o volume médio do período em análise ou o verificado durante a semana que antecede as eleições?

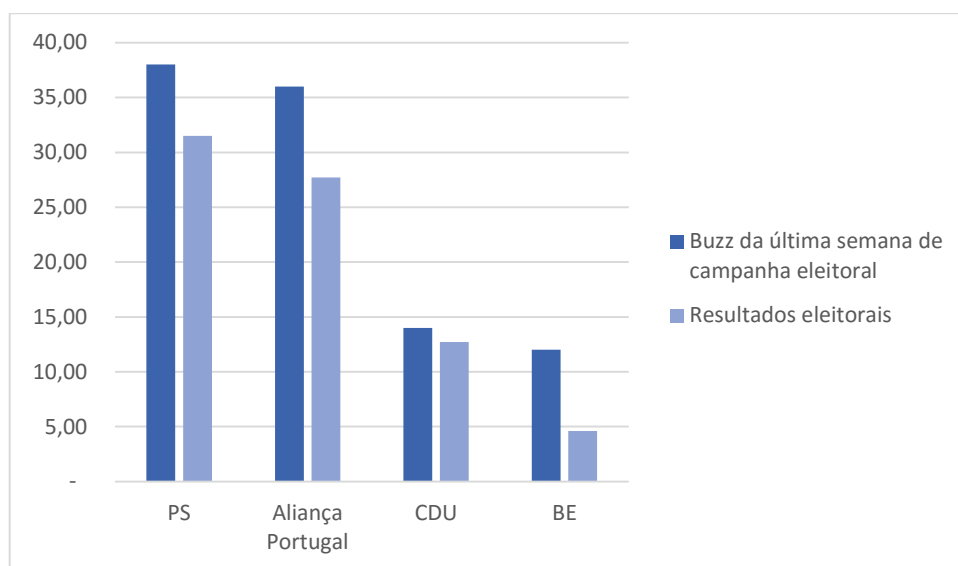
Os resultados obtidos permitem responder positivamente à questão nº1 a, já que a ordenação das forças políticas candidatas às eleições de 2014 para o PE, segundo o volume de referências na web, se assemelha ao ranking dos candidatos segundo o número de votos obtidos nesse mesmo sufrágio. Como podemos ver pelos gráficos nº 1 e 2, o PS é a força que conta com mais referências online e mais votos, seguido da coligação AP, da CDU e do BE.

Gráfico 1: Buzz médio mensal das principais forças políticas candidatas às eleições europeias de 2014 e respetivos resultados eleitorais



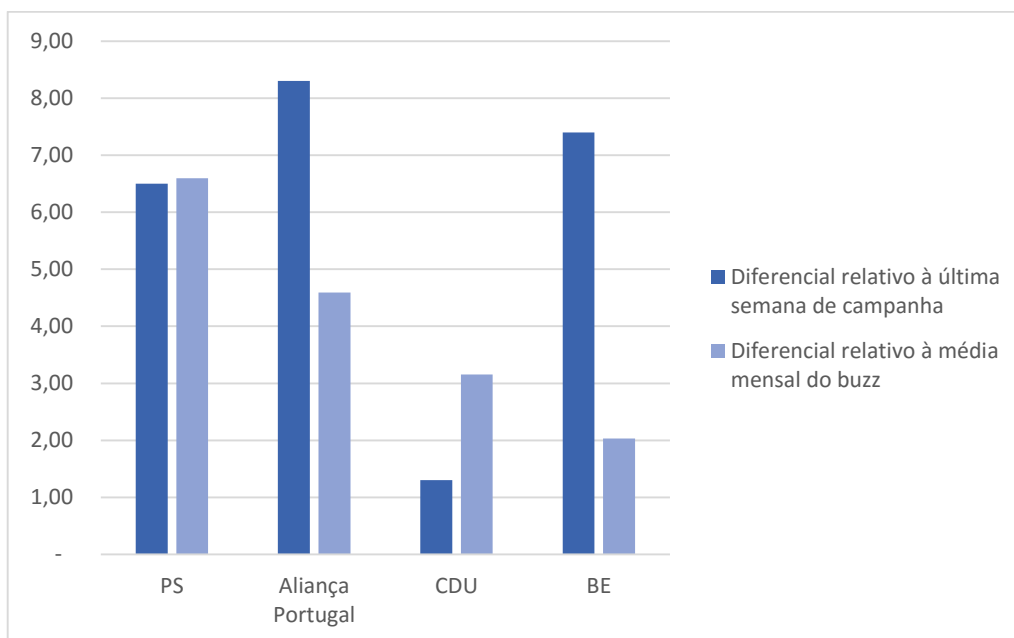
Fonte: elaboração própria.

Gráfico 2: Buzz médio de cada candidato às eleições de 2014 para o Parlamento Europeu durante a última semana de campanha e respetivos resultados eleitorais

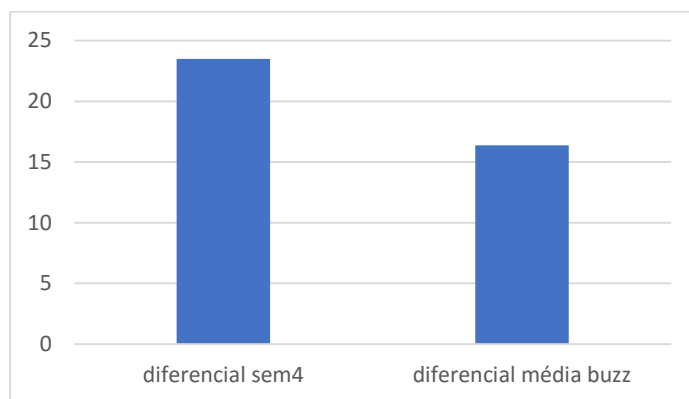


Fonte: elaboração própria.

Gráfico 3: Diferencial entre o buzz médio de cada candidato e os resultados eleitorais (dados desagregados)



Fonte: elaboração própria.

Gráfico 4: Diferencial entre o buzz médio dos candidatos e os resultados eleitorais (dados agregados)

Fonte: elaboração própria.

Nota: A coluna da esquerda representa o diferencial calculado a partir do buzz médio da última semana de campanha; a coluna da direita representa o diferencial calculado a partir do buzz médio mensal.

A relação entre estas duas variáveis altera-se, no entanto, consoante o período de análise, sendo negativa no caso do estudo da média de referências da última semana de campanha e positiva no caso do estudo da média de referências do mês imediatamente anterior às eleições. A análise comparativa dos gráficos nº 1 e 2 permite concluir que o *buzz* médio mensal é a métrica que produz valores que mais se aproximam dos resultados eleitorais, a métrica do *buzz* médio da última semana antes das eleições apresentando um diferencial superior relativamente à percentagem de votos recolhidos (ver gráfico nº3).

Neste sentido, a resposta à questão nº 1b é a de que o volume de referências que apresenta maior valor preditivo é o relativo ao *buzz* médio do mês anterior às eleições.

Estes dois resultados permitem formular algumas hipóteses explicativas. No caso da resposta à questão 1a - que vai no sentido de corroborar a validade preditiva da *world wide web*, equiparando a ordenação dos partidos políticos, segundo o volume de referências na web, à ordenação destes após o escrutínio - a equiparação de resultados aludida parece reforçar a tese de Surowiecki (Surowiecki, 2004) que, na linha de Galton (Galton, xx), veio sublinhar a qualidade das decisões resultantes de certo tipo de agregados populacionais, como é o caso da *world wide web*, relativamente às tomadas por qualquer indivíduo isoladamente, postulado que será desenvolvido e discutido na conclusão deste artigo.

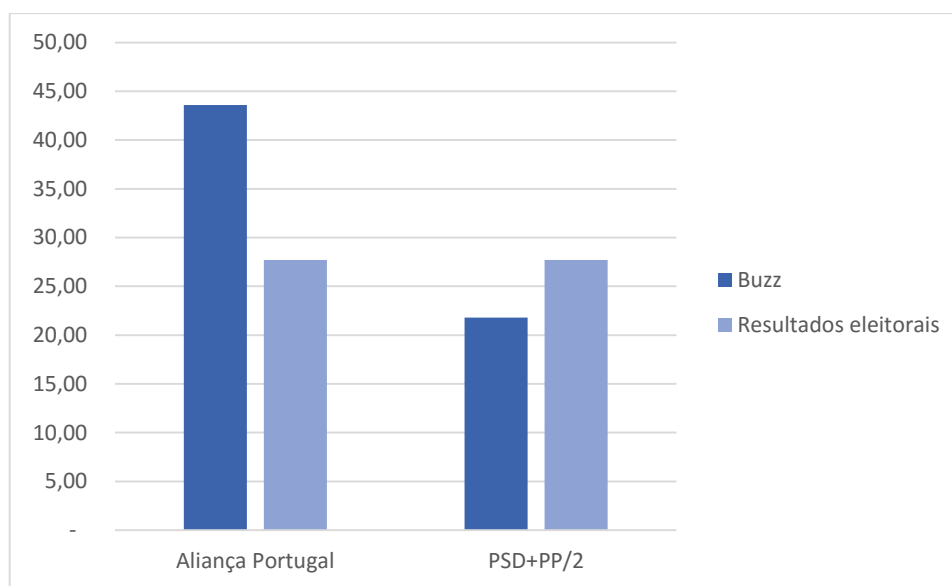
Já quanto à resposta à questão 1b - que faz prevalecer o valor preditivo da média mensal sobre o da semana que antecede o ato eleitoral - esta permite compreender a especificidade do comportamento da opinião pública online, colocando em causa a relevância preditiva das opiniões manifestadas nos dias anteriores ao sufrágio a favor da informação recolhida durante um período mais extenso. Neste sentido, o valor preditivo da tendência de médio-prazo parece sobrepor-se ao da tendência de curto-prazo.

Questão nº2

No caso das coligações, qual a métrica que apresenta maior valor preditivo: o volume de referências ao nome da coligação ou a soma do *buzz* das forças políticas que a integram?

A análise indica que a soma das referências às forças políticas que integram a coligação Aliança Portugal - PSD e CDS-PP – possui maior valor preditivo do que o *buzz* do nome da própria coligação. Tal como podemos ver no gráfico nº5, o *buzz* de Aliança Portugal apresenta valores 15,9% superiores ao número de votos obtidos por essa força candidata, o diferencial sendo menor no caso da análise do *buzz* dos partidos políticos que integram a coligação e que, apesar de inferior aos resultados eleitorais da AP, se situam apenas 5,9% abaixo.

Gráfico 5: Buzz médio mensal da expressão Aliança Portugal e dos partidos que constituem a coligação (PSD e CDS-PP) em comparação com os resultados eleitorais



Fonte: elaboração própria.

Estes resultados podem talvez ser explicados pelo facto de, não obstante a coligação candidata possuir uma designação própria, esta ser habitualmente criada no contexto do ato eleitoral, pelo que a sua notoriedade se revela naturalmente menor do que a dos partidos políticos que a integram, habituados a uma presença nos media durante os períodos não-eleitorais. Por este motivo, a utilização textual da sigla da coligação por parte dos internautas acaba por ser também menor, contrariamente ao que acontece com a sigla dos partidos políticos coligados, marcas já referenciadas na mente dos utilizadores digitais. Por outro lado, acresce ainda o facto de, nomeadamente para os simpatizantes de determinada força política, as siglas da coligação e do partido político serem percecionadas como sinónimas e utilizadas de forma indiscriminada no momento em que aqueles se exprimem online. Ainda assim, neste caso, a utilização textual da sigla do partido tenderá a ser mais natural do que a da coligação, dada a proximidade afetiva existente entre o eleitor-internauta e a força política em causa.

Em consequência, de modo a incrementar a validade dos resultados, a análise da opinião pública a partir das trocas semânticas realizadas online deverá, no caso das coligações e segundo os presentes resultados, centrar-se no estudo das siglas partidárias componentes da coligação e não no nome escolhido para esta última.

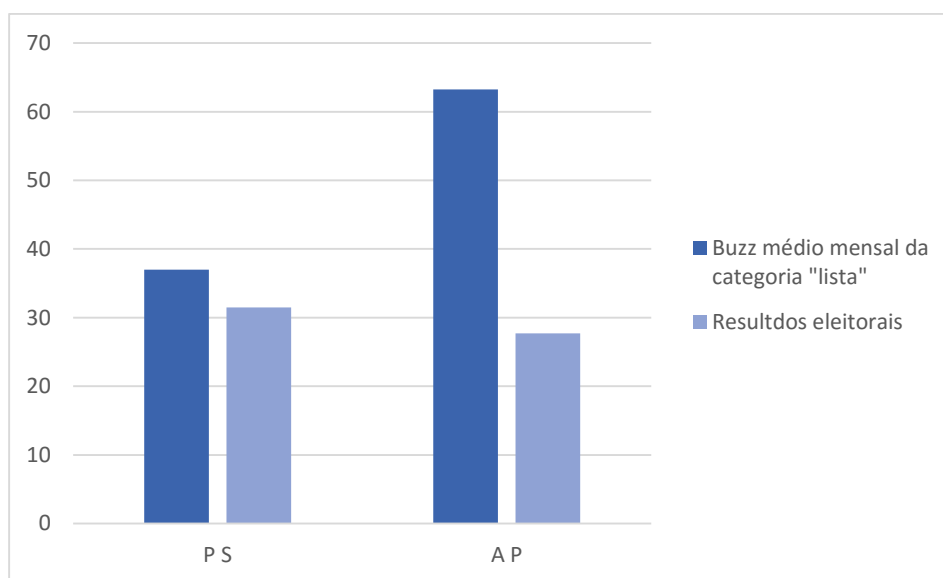
Questão nº3

Relativamente à lista apresentada por cada força política candidata, poderá o volume de referências aos elementos que a integram ser indicativo dos resultados eleitorais?

A avaliação do grau de credibilidade da categoria "Lista" incidu sobre as duas principais forças partidárias – PS e AP – já que o teste realizado à notoriedade dos elementos das listas da CDU e do BE produziu resultados tão residuais, quando não nulos, que impossibilitaram a própria realização da análise.

A análise realizada demonstrou ser a categoria lista uma dimensão com fraco índice de credibilidade, já que os seus valores não só contradizem os resultados eleitorais, colocando a AP à frente do PS, como se apresentam muito distantes da percentagem de votos obtidos por cada força candidata. Tal como podemos constatar pelo gráfico nº 6, o *buzz* médio dos elementos da lista do PS foi de 37%, quase seis pontos percentuais acima dos resultados eleitorais deste partido (31,5%), essa distância sendo ainda mais expressiva no caso da AP, cuja notoriedade dos vários elementos da lista (63,25%) se revelou quase 40% acima dos seus resultados eleitorais (27,7%).

Gráfico 6: Buzz médio mensal das listas do PS e da coligação Aliança Portugal nas eleições europeias de 2014 (em percentagens)



Fonte: elaboração própria.

A preponderância da utilização da categoria "lista" por parte dos internautas pode talvez ser explicada pelo efeito "personalização", segundo o qual a associação de um rosto e um nome a uma qualquer força política tenderia a aumentar o potencial afetivo desta, humanizando-a e aproximando-a do eleitor.

Seguindo esta tese, o elevado *buzz* dos nomes dos elementos das listas pode ser explicado pela existência de uma maior proximidade percebida por parte dos eleitores relativamente aos candidatos do que

relativamente aos partidos, o que facilitaria a utilização dos nomes dos elementos das listas durante a construção textual online.

Não obstante o efeito “personalização”, a proximidade percebida pelos eleitores relativamente aos candidatos traduz-se diferentemente em voto consoante as forças políticas sejam o PS ou o PSD. No caso do PS verifica-se um *buzz* menos elevado do que no caso do PSD, o que pode talvez ser explicado i) por uma menor notoriedade dos candidatos socialistas ii) por uma menor dedicação do PS à campanha em meios digitais ou, inversamente, por uma maior dedicação aos meios digitais por parte do PSD iii) por um benefício natural do PSD provocado pelo facto de ser o partido de apoio ao governo, o que afetaria positivamente o reconhecimento público dos candidatos sociais-democratas.

Questão nº4

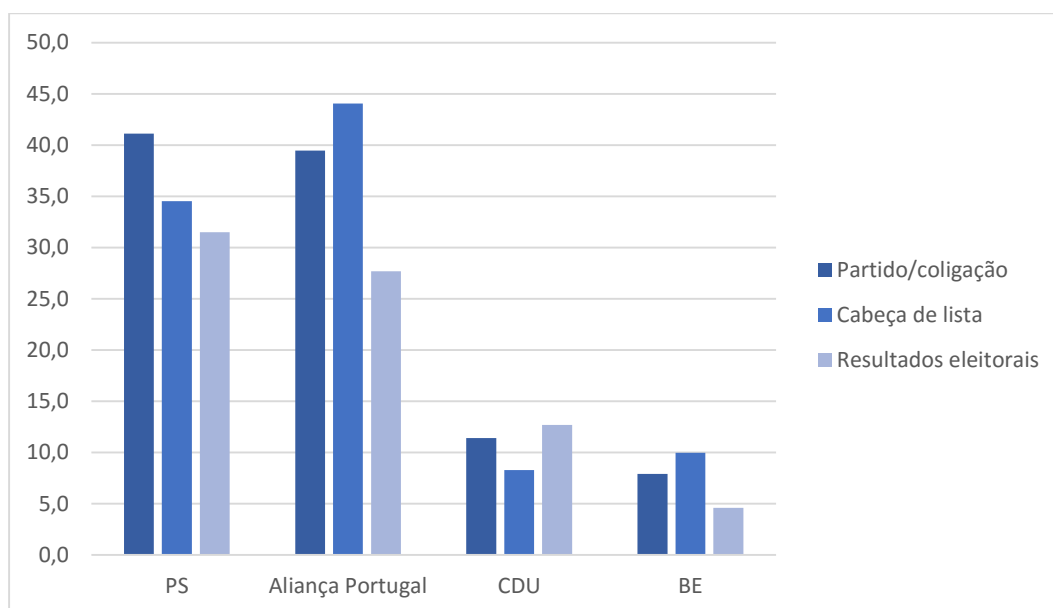
Quanto à presença online dos cabeças-de-lista, pode o seu *buzz* ser considerado preditivo dos resultados eleitorais?

A análise do grau de credibilidade da categoria “cabeça de lista” revelou resultados negativos, a gradação das forças políticas segundo este critério não acompanhando sequer os resultados eleitorais: a AP surge à frente, seguida pelo PS, pelo BE e pela CDU.

No mesmo sentido, a análise desagregada dos elementos que compõem a categoria (Francisco Assis, Paulo Rangel, João Ferreira e Marisa Matias) revelou resultados algo contraditórios, já que a proximidade da notoriedade de cada um relativamente aos resultados eleitorais oscilou consoante as forças políticas candidatas.

O caso do PS foi aquele que apresentou maior consonância entre a notoriedade do cabeça-de-lista e os resultados eleitorais, esta relação positiva dissipando-se nos casos das outras três forças políticas analisadas. A notoriedade do cabeça-de-lista da AP revelou-se 16,3% acima do volume de votos obtidos por esta coligação e, nos casos da CDU e do BE, apesar de esta diferença diminuir acentuadamente, continua ainda a existir, embora no primeiro caso seja 4,4% menor do que os resultados da coligação de esquerda e, no segundo caso, seja 5,4% maior do que os votos recolhidos pelo partido de extrema-esquerda.

Gráfico 7: Buzz médio mensal dos partidos/coligações e dos cabeças de lista e resultados eleitorais das eleições europeias de 2014 (em percentagens)



Fonte: elaboração própria.

Estes resultados apontam para explicações distintas consoante as forças políticas em causa. No caso do PS, o fator "referencial institucional" sobrepõe-se ao efeito "personalização", já que o recurso à sigla e ao nome do partido, por parte dos internautas, foi maior do que ao cabeça de lista. No caso do PSD, verifica-se o contrário.

Estas constatações permitem porventura compreender o tipo de identidade percecionada de cada um destes partidos, o PS possuindo maior referencial institucional e o PSD maior reconhecimento em dimensões como as da liderança e da personalização.

O caso dos chamados pequenos partidos é particularmente interessante nesta matéria, já que se verifica um fenómeno semelhante ao conjeturado acima para o PS e o PSD. A CDU é mais referenciada como força política do que através do nome do seu cabeça de lista, o contrário acontecendo com o Bloco de Esquerda, cujo cabeça de lista goza de um *buzz* mais elevado do que o do próprio partido.

Esta diferença pode, na linha das hipóteses explicativas avançadas acima, talvez ser explicada por questões como as da antiguidade da força política e da notoriedade do cabeça de lista. A CDU, cujo principal partido político é o PCP, o mais antigo partido político português entre os atuais, parece beneficiar significativamente de um referencial institucional, já que o *buzz* do partido é mais elevado do que o do cabeça de lista. Já no caso do Bloco de Esquerda, parece prevalecer o efeito "personalização", já que o cabeça de lista revela mais notoriedade do que o próprio partido.

Conclusões

A análise que precede corrobora algumas das teses existentes no debate científico sobre a validade preditiva das redes sociais e refuta outras, pelo que gostaríamos de as aprofundar nesta conclusão.

Em primeiro lugar, os resultados obtidos são algo encorajadores quanto à capacidade preditiva da análise de redes sociais, já que permitiram identificar, previamente à realização do sufrágio, a ordenação das forças políticas que viria a resultar das eleições. Esta conclusão vem reforçar a linha de estudos que, nos últimos anos, tem vindo a defender esta tese.

Os dados recolhidos na *world wide web* revelam consonância com os resultados das eleições portuguesas para o Parlamento Europeu, tendo expresso corretamente a ordenação eleitoral que veio a efetivar-se no dia 25 de maio de 2014. Os resultados permitiram a identificação do candidato vencedor, neste caso o PS, dos segundos partidos mais votados, o PSD e o PP, que se apresentaram coligados, bem como da ordenação eleitoral dos dois partidos seguintes, o PCP e o BE.

Esta consonância verificou-se em ambas as subcategorias testadas - o *buzz* médio mensal e o *buzz* médio da semana anterior às eleições - embora a distância relativamente aos resultados eleitorais se tenha revelado maior neste último caso, levando a concluir pela maior validade preditiva relativa da primeira subcategoria referida.

A congruência não se verificou, no entanto, no caso da percentagem de votos obtida por cada força política, já que os resultados da análise se revelaram sempre acima dos eleitorais.

Em segundo lugar e relativamente às categorias de análise testadas, a categoria "coligação" revelou fraca correspondência com os resultados das eleições, tendo o número de menções ao nome da coligação candidata - a Aliança Portugal - saído sobreavaliado relativamente ao número de votos recolhidos por esta força política. No entanto, quando testada a correspondência entre a soma do número de referências aos dois partidos políticos que integram a coligação e os resultados eleitorais, esta ofereceu resultados mais encorajadores, corroborando, de resto, estudos anteriormente realizados (Tumasjan et al., 2010).

As categorias "lista" e "cabeça de lista" revelaram fraca correspondência com os resultados eleitorais, embora, desconhecendo-se a existência de outros estudos sobre a validade empírica destas dimensões, aquelas mereçam ser de novo testadas em investigação futura.

Do que precede se conclui pela validade da *world wide web* na indicação de tendências de voto e, em particular, da categoria de análise "partido", que se revelou a mais adequada ao estudo do posicionamento da opinião pública no momento pré-eleitoral.

A saliência desta categoria parece depender em forte medida da notoriedade do nome e da sigla em causa, dimensão que, por sua vez, resulta de fatores como a longevidade e a intensidade da exposição pública das mesmas. Esta hipótese explicativa sai particularmente reforçada do presente estudo, já que o *buzz* das siglas partidárias com maior longevidade, como são os casos das do PS e do PCP, apresenta maior valor preditivo do que o de forças políticas constituídas mais recentemente, como são os casos da Aliança Portugal, formada especificamente para o ato eleitoral em questão, e do Bloco de Esquerda, um dos mais recentes partidos portugueses.

Por outro lado, face às outras categorias testadas, a categoria "partido" parece refletir o reconhecimento e a valorização pública da dimensão institucional das forças políticas candidatas a eleições, aspetos que se sobrepõem a questões como a notoriedade do cabeça de lista ou dos elementos da lista. A saliência do fator

“institucionalização”, por sua vez, pode apontar para a valorização social de dimensões como a estabilidade, a estandardização e a segurança.

Esta saliência do fator “institucionalização” permite ainda realçar a importância histórica dos media tradicionais na construção diacrónica da identidade organizacional, questão que os designados novos media tenderão a ampliar para novas dimensões comunicativas. No entanto, não obstante o efeito amplificador dos meios digitais, resulta claro desta análise o papel decisivo do *capital de notoriedade inicial* na construção da saliência percetiva de qualquer força política candidata a eleições.

Em terceiro e último lugar, o presente estudo permitiu reforçar a demonstração da validade científica da *world wide web* como universo de análise em Ciências Sociais, apontando para a consolidação da hipótese da designada “sabedoria das multidões” (Surowiecki, 2004; Simmons et al. 2011; Hertwig 2012; Koriat 2012), a qual, seguindo uma linha de estudo com quase dois séculos de existência (Mackay, 1843; Galton, 1907; Hogarth, 1978; Treynor, 1987; Surowiecki 2004; Kittur and Kraut 2008; Leimeister et al. 2009) defende a qualidade das decisões resultantes da agregação da informação em grupos relativamente àquelas tomadas por membros individualizados.

Esta “sabedoria das multidões” parece ter encontrado na *world wide web* um espaço particularmente apto ao seu desenvolvimento, já que, tal como acontece nos grandes agregados humanos, esta se caracteriza por fatores como a diversidade, a independência, a descentralização e a agregação dos seus elementos, sendo estes os critérios que diferenciam as designadas *multidões sábias* (do inglês “wise crowd”) das que o não são, designadas *irracionais* (do inglês “irrational crowd”).

No entanto, tal como ficou demonstrado pelo presente estudo, o otimismo quando à validade científica da *world wide web* como universo de análise deverá ser moderado na afirmação da validade preditiva dos meios digitais em matéria de resultados eleitorais.

A precaução acima aludida não visa, contudo, coartar o entusiasmo científico existente em torno desta área recente, parecendo evidente a necessidade de realização de mais estudos e o teste de novas categorias de análise, sobretudo no plano nacional, em que rareia investigação neste campo.

Referências bibliográficas

- AAPOR (2009). *An Evaluation of the Methodology of the 2008 Pre-Election Primary Polls*. Lenexa, KS: American Association of Public Opinion Research.
- Abejón, P.; Sastre, A. & Linares, V. (2012). Facebook y Twitter en campañas electorales en España. Anuario Electrónico de Estudios en Comunicación Social. *Disertaciones*, 5(1). Disponível em: <http://erevistas.saber.ula.ve/index.php/Disertaciones/>. Consultado em 24 de Agosto de 2016.
- Adamic, L. A. & Glance, N. (2005). The political blogosphere and the 2004 US election: Divided they blog. In *Proceedings of the 3rd International Workshop on Link Discovery*, 36-43. Chicago, IL.
- Adar, E., Weld, D., Bershady, B. & Gribble, S. (2007). “Why We Search: Visualizing and Predicting User Behavior.” In *Proceedings of the 16th International Conference on World Wide Web*, 161–70. WWW '07. New York, NY, USA: ACM.
- Aday, S.; Farrell, H.; Lynch, M.; Sides, J.; & Freelon, D. (2012). *Blogs and Bullets II: New Media and Conflict after the Arab Spring*. United States Institute of Peace.

- Albrecht, S.; Lübcke, M. & Hartig-Perschke, R. (2007). Weblog Campaigning in the German Bundestag Election 2005. *Social Science Computer Review*, 25(4): 504-520.
- Allison, P. (2001). *Missing data*. Thousand Oaks, CA: SAGE.
- Ampofo, L., Anstead, N. & O'Loughlin, B. (2011). Trust, Confidence and Credibility. *Information, Communication & Society* 14 (6): 850–71.
- Asur, S. & Huberman, B. (2010). Predicting the future with Social Media. *International Conference on Web Intelligence and Intelligent Agent Technology*, vol.1, 492-499. Disponível em: <http://arxiv.org/abs/1003.5699>. Consultado em 24 de agosto de 2016.
- Bakker, T. & de Vreese, C. (2011). Good News for the Future? Young People, Internet Use and Political Participation. *Communication Research* 20(10): 1–20.
- Bamman, D.; Eisenstein, J. & Schnoebelen, T. (2012). Gender identity and lexical variation in social media. *Journal of Sociolinguistics*, 18, 39-52. Disponível em: <http://adsabs.harvard.edu/abs/2012arXiv1210.4567B>. Consultado em 28 de agosto de 2016.
- Banks, J. (1991). *Signaling Games in Political Science*, Harwood Academic.
- Barbera, P. (2013). Birds of the Same Feather Tweet Together. Bayesian Ideal Point Estimation Using Twitter Data. *APSA 2012 Annual Meeting Paper*. Disponível em: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2108098
- Bennett, W. & Alexandra S. (2011). Digital media and the personalization of collective action: Social technology and the organization of protests against the global economic crisis. *Information Communication & Society*, 14(6):770-799.
- Benoît, K., Conway, D. Lauderdale, B., Laver, M. & Mikhaylov, S. (2016). Crowd-sourced text analysis. Reproducible and agile production of political data. *American Political Science Review*, 110 (2). 278-295.
- Bermingham, A. & Smeaton, A. (2011). On using Twitter to monitor political sentiment and predict election results. *Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP), IJCNLP 2011*, 2–10.
- Bernstein, M. (2013). Quantifying the Invisible Audience in Social Networks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 21-30. New York: Association for Computing Machinery.
- Best, M. & Wade, K. (2009). The Internet and Democracy Global Catalyst or Democratic Dud? *Bulletin of Science, Technology & Society*, Vol 29(4), 255-271.
- Best, S. & Krueger, B. (2005). Analyzing the Representativeness of Internet Political Participation. *Political Behavior* 27(2): 183–216.
- Böhringer, M., & Richter, A. (2009). Adopting Social Software to the Intranet: A Case Study on Enterprise Microblogging. In *Proceedings of the 9th Mensch & Computer Conference*, 293-302. Berlin.
- Bolker, J. (1995). Model systems in developmental biology. *BioEssays* 17(5): 451–455.
- Boyd, D. & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, 15 (5): 662–679.
- Borondo, J., Morales, A., Losada, J. & Benito, R. (2012). Characterizing and modeling an electoral campaign in the context of Twitter: 2011 Spanish Presidential Election as a Case Study. *Chaos*, 22, 023138.
- Butts, C. (2009). Revisiting the foundations of network analysis. *Science*, 325(5939): 414.

- Cameron, M., Barrett, P. & Stewardson, B. (2013). Can Social Media Predict Election Results? Evidence from New Zealand. Working Paper in Economics 13/08, University of Waikato.
- Campbell, J. E., & Garand, J. C. (Eds.). (2000). *Before the vote: Forecasting American national elections*. Thousand Oaks, CA: Sage.
- Campos, F. y Rúas, J. (2015) (Eds.). Las redes sociales digitales en el ecosistema mediático. La Laguna (Tenerife): *Cuadernos Artesanos de Comunicación*, 92. Disponible em: <http://www.cuadernosartesanos.org/2015/cac92.pdf>. Consultado em 6 de agosto de 2016.
- Carr, A. (2010). *Facebook, twitter election results prove remarkably accurate*. Fast Company. Disponible em: <http://bit.ly/dW5gxo>. Consultado em 20 de julho de 2016.
- Carrera, M.; Olmos, J. & Cebollero, M. (2013). Psicología y lenguaje en política: los candidatos a la Presidencia del Gobierno y su estilo lingüístico. *Anuario de Psicología*, 43, 1, 39-52.
- Casero, A. (2015). Estrategias y prácticas comunicativas del activismo político en las redes sociales en España. *Historia y Comunicación Social*, 20(2), 535-550. Disponible em: http://dx.doi.org/10.5209/rev_HICS.2015.v20.n2.51399. Consultado em 6 de marzo de 2016.
- Castells, M. (2009). *Communication power*. Oxford/New York: Oxford University Press.
- Ceron, A., Curini L., Stefano M., Giuseppe P. (2013). Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' policy preferences. An application to Italy and France, *New Media & Society*, 16(2), 340-358. DOI: 10.1177/1461444813480466
- Cha, M.; Haddadi, H.; Benevenuto, F. & Gummadi, K. (2010). Measuring user influence in twitter: The million follower fallacy. *ICWSM-10*: 10-17.
- Chadwick, A. (2008) Web 2.0: New challenges for the study of e-democracy in an era of informational exuberance. *I/S: Journal of Law & Policy for the Information Society* 4(3): 9-42.
- Chang, R. et al (2013). Understanding the paradigm shift to computational social science in the presence of Big Data. *Support Systems*, <http://dx.doi.org/10.1016/j.dss.2013.08.008>.
- Choy, M., Cheong, M., Laik, M., Shung, K. (2011). A sentiment analysis of Singapore Presidential Election 2011 using Twitter data with census correction. URL: <http://arxiv.org/abs/1108.5520>.
- Choy, M., Cheong, M., Laik, M., Shung, K. (2012). US Presidential Election 2012 Prediction using Census Corrected Twitter Model. URL: <http://arxiv.org/ftp/arxiv/papers/1211/1211.0938.pdf>
- Chung, J. & Mustafaraj, E. (2011). Can collective sentiment expressed on twitter predict political elections? *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence*, San Francisco, CA, USA.
- Clements, M. & Hendry, D. (eds.) (2011). *Oxford Handbook of Economic Forecasting*. Oxford: Oxford University Press.
- Congosto, M. & Aragón, P. (2012). Twitter, del sondeo a la sonda: nuevos canales de opinión, nuevos métodos de análisis. *Más Poder Local*, 12, 50-56.
- Conover, M., Goncalves, B., Ratkiewicz, J., Flammini, A. & Menczer, F. (2011). Predicting the Political Alignment of Twitter Users. *IEEE International Conference on Privacy, Security, Risk and Trust and International Conference on Social Computing*, 192-199.
- Cottle, S. (2011). Media and the Arab uprisings of 2011: Research notes. *Journalism* 12(5): 647-659.
- Cox, G. (1997). *Making votes count: Strategic coordination in the world's electoral systems*. New York: Cambridge University Press.

- Crespi, I. (1988). Pre-election Polling: Sources of Accuracy and Error. *Fourth International AAAI Conference on Weblogs and Social Media*, 10-17.
- Dahlberg, L. (2007). Rethinking the fragmentation of the cyberpublic: from consensus to contestation. *New Media Society* 9, 827.
- Donath, J. (2014). *The social machine: Designs for living online*. Cambridge, MA: MIT Press.
- Drezner, D. & Henry F. (2007). Introduction: Blogs, politics and power: a special issue of Public Choice. *Public Choice*, 134(1-2): 1-13.
- Durand, C., Blais, A. & Larochelle, M. (2004). The Polls in the 2002 French Presidential Election: An Autopsy. *Public Opinion Quarterly* 68(4): 602-622.
- Falck, O., Gold, R. & Heblich, S. (2012). *E-Lectons: Voting Behavior and the Internet*. IZA Discussion Paper.
- Farrell, H. & Drezner, D. (2008). The power and politics of blogs. *Public Choice*. 134(1-2): 15-30.
- Fernández-Cabana, M., García-Caballero, A., Alves-Pérez, M., García-García, M. & Mateos R. (2013). Suicidal traits in Marilyn Monroe's Fragments: an LIWC analysis. *Crisis*, 34, 2, 124-130.
- Fields, S. & Johnston, M. (2005). Whither Model Organism Research? *Science* 307(5717): 1885-1886.
- Fox, S. (2010). *Four in ten seniors go online*. Pew Research Center, URL: <http://www.pewinternet.org/Commentary/2010/January/38-of-adults-age-65-go-online.aspx>.
- Franch, F. (2012). (Wisdom of the Crowds): 2010 UK Election Prediction with Social Media. *Journal of Information Technology & Politics* DOI: 10.1080/19331681.2012.705080
- Freelon, D. (2014). On the interpretation of digital trace data in communication and social computing research. *Journal of Broadcasting & Electronic Media*, 58(1), 59-75.
- Freeman, S. & Bleifuss, J. (2006). Was the 2004 Presidential Election Stolen?: Exit Polls, Election Fraude the Official Count. *Seven Stories*.
- García, A.; García, I. & Varona, D. (2012). Incidencia de las redes sociales vs. cibermedios en las Elecciones en España, 2011. *Enl@ce Revista Venezolana de Información, Tecnología y Conocimiento*, 9, 2, 11-29.
- García, M., del Hoyo, M. & Fernández, C. (2014). Jóvenes comprometidos en la Red: el papel de las redes sociales en la participación social activa. *Comunicar*, 43, 35-43.
- Gayo-Avello, D.; Metaxas & Mustafaraj (2011). Limits of Electoral Predictions Using Twitter. Association for the Advancement of Artificial Intelligence. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, 490-493. Disponível em: <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2862>. Consultado em 5 de maio de 2016.
- Gayo-Avello, D. (2011). *A warning against converting social media into the next literary digest*. CACM.
- Gayo-Avello, D. (2012). No, You Cannot Predict Elections with Twitter. *IEEE Internet Computing*, vol. 16, no 6. pp. 91-94, Nov.-Dec. 2012.
- Geddes, B. (1990). How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics. *Political Analysis* 2(1): 131-150.
- Gelman, A., Silver, N., Edlin, A. (2012). What Is the Probability Your Vote Will Make a Difference, *Economic Inquiry*, 50(2): 321-326.
- Ghannam, J. (2011). *Social Media in the Arab World: Leading up to the Uprisings of 2011*. Center for International Media Assistance.

- Ghonim, W. (2012). *Revolution 2:0: A Memoire Call to Action*. New York: Houghton Mifflin Harcourt.
- Gibson, R., Lusoli, W. & Ward, S. (2008). Nationalizing and normalizing the local? A comparative analysis of online candidate campaigning in Australia and Britain. *Journal of Information Technology & Politics* 4(4): 15–30.
- Gibson, R. and McAllister, I. (2013). Online Social Ties and Political Engagement. *Journal Of Information Technology & Politics*, (10)1.
- Gil De Zúñiga, H.; Jung, N. y Valenzuela, S. (2011). Social media use for news and individual's social capital, civic engagement and political participation. *Journal of Computer-Mediated Communication*, 3, 319-333.
- Gilbert, S. (2001). Ecological Developmental Biology: Developmental Biology Meets the Real World. *Developmental Biology* 233(1): 1–12.
- Gimmler, A. (2001). Deliberative democracy, the public sphere and the internet. *Philosophy & Social Criticism* 27: 21-39.
- Gloor, P., Krauss, J., Nann, S., Fischbach, K., Schoder, D. (2009). Web Science 2.0: Identifying Trends through Semantic Social Network Analysis. *CSE4. 2009 International Conference on Computational Science and Engineering*: 215–222.
- Goidel, K. (2011). *Political Polling in the Digital Age: The Challenge of Measuring and Understanding Public Opinion*. New Orleans: LSU Press.
- Golbeck, J. (2011). Computing Political Preference among Twitter Followers. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1105-1108. New York: ACM Digital Library. Disponível em: <http://hci2.cs.umd.edu/trs/2010-20/2010-20.pdf>. Consultado em 2 de julho de 2016.
- Golbeck, J. (2012). The Twitter Mute Button: A Web Filtering Challenge. *CHI '12: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2755-2758. New York: ACM Digital Library.
- Golder, S. & Macy, M. (2011). Diurnal and seasonal mood vary with work, sleep and day length across diverse cultures. *Science* 333(6051): 1878–1881.
- Goldstein, P. & Rainey, J. (2010). The 2010 elections: Twitter isn't a very reliable prediction tool. *Los Angeles Times*. Disponível em: <http://lat.ms/fSXqZW>. Consultado em 2 de junho de 2016.
- González, B. (2015). E-Actividad de los políticos en Facebook: políticos gallegos y europarlamentarios españoles 2010-2014. In J. Rúas-Araújo, A. Silva-Rodríguez & I. Puentes- Rivera (Eds.), *De los Medios a los Metamedios de Comunicación* (pp. 923-944). Pontevedra: XESCOM / Universidade de Santiago de Compostela. Disponível em: <https://delosmediosalosmetamedios.wordpress.com/actas/> Consultado em 18 de junho de 2016.
- Grimmer, J. & Stewart, M. (2013). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts, *Political Analysis*, 21(3), 267-297. doi:10.1093/pan/mps028.
- Gueorguieva, V. (2007). Voters, MySpace and YouTube: The Impact of Alternative Communication Channels on the 2006 Election Cycle and Beyond. *Social Science Computer Review*, 26(3): 288-300.
- Gunsch, M.; Brownlow, S.; Haynes, S. & Mabe, Z. (2000). Differential linguistic content of various forms of political advertising. *Journal of Broadcasting & Electronic Media*, 44, 27-42.
- Habermas, J. (1962). *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*. Cambridge: Polity Press (1989).

- Hanse, K. & Kosiara-Pederson, K. (2014). Cyber-Campaigning in Denmark: Application and Effects of Candidate Campaigning. *Journal of Information Technology & Politics*, 11:206–219.
- Hargittai, E. (2008). Whose Space? Differences Among Users and Non-Users of Social Network Sites. *Journal of Computer-Mediated Communication* 13(1): 276–297.
- Hilbert, M. (2009). The maturing concept of e-democracy: From e-voting and online consultations to democratic value out of jumbled online chatter. *Journal of Information Technology & Politics* 6(2): 87–110.
- Hillygus, D. (2011). The Evolution of Election Polling in the United States. *Public Opinion Quarterly* 75 (5): 962–81.
- Holmes, D., Alpers, G., Ismailji, T., Classen, C., Wales, T., Cheasty, V., Koopman, C. (2007). Cognitive and emotional processing in narratives of women abused by intimate partners. *Violence Against Women*, 13, 1192-1205.
- Honeycutt, C. and Herring, S. (2009). Beyond microblogging: Conversation and collaboration via Twitter. In *42nd Hawaii International Conference on System Sciences*, 1-10, Hawaii.
- Hopkins, D., King, G. (2010). A Method of Automated Nonparametric Content Analysis for Social Science. *American Journal of Political Science* 54(1): 229–247.
- Huber, J. & Hauser, F. 2005. Systematic mispricing in experimental markets evidence from political stock markets. In *10th Annual Workshop on Economic Heterogeneous Interacting Agents*. Essex, UK.
- Jansen, H. & Koop, R. (2005). Pundits, Ideologues and Ranters: The British Columbia Election Online. *Canadian Journal of Communication* 30(4): 613–632.
- Jansen, B.; Zhang, M.; Sobel, K. & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology*, 60: 1 20.
- Java, A.; Song, X.; Finin, T. & Tseng, B. (2007). Why we twitter: understanding microblogging usage and communities. In *Proceedings of the 9th WebKDDe 1st SNA-KDD 2007 workshop on Web mining and social network analysis*, 56 65. San Jose, CA: ACM.
- Jenner, R. & Wills, M. (2007). The choice of model organisms in evo–devo. *Nature Reviews Genetics* 8(4): 311–314.
- Jensen, M. & Anstead, N. (2013). Psephological Investigations: Tweets, Votes and Unknowns in the Republican Nomination Process, *Policy & Internet*, 5(2): 161-182
- Jensen, M., Jorba, L. & Anduiza, E. (2012). Introduction. In *Digital Media and Political Engagement Worldwide: A Comparative Study*, eds. E. Anduiza, M. Jensen & L. Jorba. New York: Cambridge University Press, 1–15.
- Jérôme, B., Jérôme, V. & Lewis-Beck, M. (1999). Polls fail in France: forecasts of the 1997 legislative election. *International Journal of Forecasting* 15: 163-174.
- Jungherr, A., Jürgens, P. & Schoen, H. (2012). Why the Pirate Party Won the German Election of 2009 or The Trouble With Predictions: A Response to Tumasjan et.al.: Predicting Elections With Twitter: What 140 Characters Reveal About Political Sentiment. *Social Science Computer Review*. 30, 2, 229-234.
- Jungherr, A. (2014). The logic of political coverage on Twitter: Temporal dynamics and content. *Journal of Communication* 64(2): 239-259.
- Kacewicz, E.; Pennebaker, J. W.; Davis, M.; Jeon, M. y Graesser, A. C. (2013). Pronoun Use Reflects

- Standings in Social Hierarchies. *Journal of Language and Social Psychology*, 1-19. SAGE Publications. DOI: 10.1177/0261927X13502654. Disponível em: <http://jls.sagepub.com/content/early/2013/09/18/0261927X13502654> Consultado em 11 de junho de 2016.
- Kairam, S., Morris, M., Teevan, J., Liebling, D. & Dumais, S. (2013). Towards Supporting Search over Trending Events with Social Media. In *Seventh International AAAI Conference on Weblogs and Social Media*.
- Kim, D. (2007). *Identifying Opinion Leaders by Using Social Network Analysis: A Synthesis of Opinion Leadership Data Collection Methods and Instruments*. (Unpublished doctoral dissertation). Scripps College of Communication, Ohio University.
- King G., Pan, J., & Roberts, M. (2012). How Censorship in China Allows Government Criticism but Silences Collective Expression. *American Political Science Review*, 107(2), 326-343. doi:10.1017/S0003055413000014.
- Knigge, P. (1998). The Ecological Correlates of Right-Wing Extremism in Western Europe. *European Journal of Political Research* 34: 249-279.
- Koop, R. & Jansen, J. (2009). Political Blogs and Blogrolls in Canada: Forums for Democratic Deliberation? *Social Science Computer Review*, 27, 2, 155-173.
- Kwak, H., Lee, C., Park, H. & Moon, S. (2010). What Is Twitter, a Social Network or a News Media? In *Proceedings of the 19th International Conference on World Wide Web*, 591–600. WWW '10. New York, NY, USA: ACM.
- Vázquez, D. (2015). La utilidad de las redes sociales como herramientas investigadoras en el marco mediático. In J. Rúas-Araújo, A. Silva-Rodríguez & I. Puentes-Rivera (Eds.), *De los Medios a los Metamedios de Comunicación*. Pontevedra: XESCOM / Universidade de Santiago de Compostela. Disponível em: <https://delosmediosalosmetamedios.wordpress.com/actas/> Consultado em 8 de abril de 2016.
- Larsson, A., Moe, H. (2012). Studying political microblogging: Twitter users in the 2010 Swedish election campaign. *New Media & Society* 14(5): 729–747.
- Lasorsa, L. & Holton, A. (2012). Normalizing twitter: journalism practice in an emerging communication space. *Journalism Studies* 13(1): 19–36.
- Laver, M., Benoit, K. & Sauger, N. (2006). Policy competition in the 2002 French legislative and presidential elections. *European Journal of Political Research* 45: 667-697.
- Lazer, D.; Pentland, A.; Adamic, L.; et al. (2009). Computational Social Science. *Science* 323(5915): 721–723.
- Lenhart, A., Purcell, K., Smith, A. & Zickuhr, K. (2010). *Social Media and Young Adults*. Washington, D.C.: Pew Internet and American Life Project.
- Lerman, K. & Ghosh, R. (2010). Information Contagion: An Empirical Study of the Spread of News on Digg and Twitter Social Networks. *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, 10 (2010): 90-97.
- Lewis-Beck, M. (2005). Election forecasting: Principles and practice. *The British Journal of Politics and International Relations*, 7: 145-164

- Lindsay, R. (2008). *Predicting polls with Lexicon*.
<http://languagewrong.tumblr.com/post/55722687/predicting-polls-with-lexicon>.
- Lynch, M. (2012). *The Arab uprising: the unfinished revolutions of the new Middle East*. New York: Public Affairs.
- Lynch, M.; Freelon, D. & Aday, S. (2014). *Syria's Socially Mediated Civil War*. United States Institute of Peace.
- Madge, C., Meek, J., Wellens, J., Hooley, T. (2009). Facebook, social integration and informal learning at university: It is more for socializing and talking to friends about work than for actually doing work. *Learning, Media and Technology* 34(2): 141–155.
- McKenna, L. and Pole, A. (2007). What do bloggers do: an average day on an average political blog. *Public Choice*, 134(1-2): 97-108.
- Martins, T. & Azevedo, I. (2015). Tendências no Twitter. Um Estudo de Caso. In A. Rocha, A. Martins, G. Paiva Dias, L.P. Reis & M. Pérez Cota (Eds.), *Sistemas & Tecnologias de Informação*. Atas da 10a. Conferência Ibérica de Sistemas & Tecnologias de Informação (pp. 922-927). Águeda, Portugal: AISTI / Universidade de Aveiro. Disponível em: <http://www.aisti.eu/cisti2015/index.php/es/proceedings> Consultado em 24 de marzo de 2016.
- Meiman, J. & Freund, J. (2012). "Large Data Sets in Primary Care Research." *The Annals of Family Medicine* 10 (5): 473–74.
- Metaxas, P. & Mustafaraj, E. (2010). From obscurity to prominence in minutes: political speech and real-time search. *Web Science Conference 2010*, April 26-27, 2010, Raleigh, NC, USA. Available at: <http://bit.ly/h3Mfld>.
- Míguez, M., Fernández, M. & Alves, M. (2013). La actividad en Twitter de los partidos políticos en las elecciones gallegas. In M. Ledo & M. Vasallo de Lopes (eds.), *Comunicación, Cultura & Esferas de Poder* (pp. 3.505-3.521). Santiago de Compostela: IBERCOM/ AGACOM. Disponível em: http://www.estudiosaudiovisuais.org/lusofonia/?page_id=844&lang=es. Consultado em 17 de maio de 2016.
- Mislove, A. et al (2011). Understanding the Demographics of Twitter Users. *Proceedings of the 5 International AAAI Conference on Weblogs and Social Media (ICWSM'11)*. Association for the Advancement of Artificial Intelligence, 554-557. Disponível em: <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2816> Consultado 5 de março de 2016.
- Mitchell, A. & Hitlin, P. (2014). *Twitter Reaction to Events Often at Odds with Overall Public Opinion*. Pew Research Center. Accessed at: <http://www.pewresearch.org/2013/03/04/twitter-reaction-to-events-often-at-odds-with-overall-public-opinion/>.
- Morozov, E. (2009). Iran: Downside to the 'Twitter revolution'. *Dissent* 56(4): 10–14.
- Mustafaraj, E., Finn, S., Whitlock, C. & Metaxas, P. (2011). Vocal Minority versus Silent Majority: Discovering the Opinions of the Long Tail. *Privacy, Security, Risk and Trust (PASSAT), 2011 IEEE Third International Conference on Social Computing (SocialCom)* 103– 110.
- Nedeau, R., Lewis-Beck M. & Bélanger, E. (2012). Proxy models for election forecasting: The 2012 French test. *French Politics* 10(1): 1–10.

- Noelle-Neumann, E. (1974). The spiral of silence: a theory of public opinion. *Journal of Communication* 24: 43–51.
- Nulty, Paul, Theocharis, Yannis, Popa, Sebastian Adrian, Parnet, Olivier and Benoit, Kenneth (2016) Social media and political communication in the 2014 elections to the European Parliament. *Electoral Studies* (in press). Disponível em <http://dx.doi.org/10.1016/j.electstud.2016.04.014>. Consultado em 3 de julho de 2016.
- O`Connor, B. et al. (2010). From tweets to polls: Linking text sentiment to public opinion time series. *Proceedings of 4th International AAAI conference on Weblog and Social Media (ICWSM)*, 122-129. Disponível em: <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/view/1536/1842> Consultado em 12 de maio de 2016.
- Orihuela, J. (2011). *Mundo Twitter*. Barcelona: Alienta.
- Östman, J. (2012). Information, expression, participation: How involvement in user-generated content relates to democratic engagement among young people. *New Media & Society* 14(6): 1004–1021.
- Outhwaite, W.; Turner, S.; Dunning, T. & Freedman, D., eds. (2007). Modeling Selection Effects. In *The SAGE Handbook of Social Science Methodology*. Thousand Oaks, CA: SAGE Publications.
- Pamela, J. & Shannon L. (2011). *Politics and the Twitter Revolution: How Tweets Influence the Relationship between Political Leaders and the Public*. Lanham, MD, Lexington.
- Panagiotis et al. (2011). How (not) to Predict Elections. Privacy, Security, Risk Trust. *IEEE Third International Conference on Social Computing*, 165-171. Disponível em: <http://ieeexplore.ieee.org/xpl/login.jsp?reload=true&tp=&number=6113109&url=http%3A%2F%2Fieeex>. Consultado em 2 de julho de 2016.
- Pennebaker, J.; Mehl, M. & Nierderhoffer, K. (2003). Psychological aspects of natural language use: our words, our selves. *Annual Review of Psychology*, 54, 547-577.
- Pérez, J.; Puentes-Rivera, I. & García, A. (2013). Los candidatos en las elecciones gallegas en Twitter. En: Ledo, M.; Vasallo, M. (eds.), *Comunicación, cultura & esferas de poder*. Santiago de Compostela: IBERCOM/AGACOM, 3.572-3.590. Disponível em: http://www.estudiosaudiovisuais.org/lusofonia/?page_id=844&lang=es. Consultado em 3 de agosto de 2016.
- Perlmutter, D. (2008). Political Blogging and Campaign 2008: A Roundtable. *The International Journal of Press/Politics*, 13(2): 160-170.
- Pew Research Center (2012b). *Social media and voting*. URL: http://www.pewinternet.org/~/media/Files/Reports/2012/PIP_TheSocialVote_PDF.pdf
- Polat, R. (2005). The Internet and Political Participation: Exploring the Explanatory Links. *European Journal of Communication* 2005; 20; 435
- Procter, R. & Halfpenny, P. (Eds.) (2015). *Innovations in Digital Research Methods*. London: Sage.
- Ramírez-Esparza, N. et al. (2007). La psicología del uso de las palabras: Un programa de computadora que analiza textos en Español (The psychology of Word use: A computer program that analyzes texts in Spanish). *Revista Mexicana de Psicología*, 24, 85-99.
- Rasmussen, T. (2009). The significance of internet communication in public deliberation. *Javnost-The Public*, 16 (1), 17-32.

- Redondo et al (2007). The Spanish adaptation of ANEW (Affective norms for English Words). *Behavior Research Methods* 39 (3) 600-605.
- Rodríguez, R. & Ureña, D. (2011). Diez razones para el uso de Twitter como herramienta de comunicación política y electoral. *Comunicación y Pluralismo*, 10, 89-116.
- Romero, D., Meeder, B. & Kleinberg, J. (2011). Differences in the mechanics of information diffusion across topics: idioms, political hashtags and complex contagion on twitter. *Proceedings of the 20th international conference on World wide web*. ACM, 2011.
- Ruas-Araújo, J., Puentes-Rivera, I., Miguez-González, M. (2016). Capacidad predictiva de Twitter, impacto electoral y actividad en las elecciones al Parlamento de Galicia: un análisis con la herramienta LIWC. *Observatorio (OBS*) Journal*, vol.10 - no2, 055-087.
- Rubio García, R. (2014). Twitter y la Teoría de la Agenda Setting: mensajes de la opinión pública digital. *Estudios sobre el Mensaje Periodístico*, 20(1), 249-264.
- Salathé, M., Vu, D., Khandelwal, S. & Hunter, D. (2013). The Dynamics of Health Behavior Sentiments on a Large Online Social Network. *EPJ Data Science* 2 (1): 4. doi:10.1140/epjds16.
- Sanders, E. & van den Bosch, A. (2013). *Relating Political Party Mentions on Twitter with Polls and Election Results*, unpublished manuscript. URL: ceur-ws.org/Vol-986/paper_9.pdf
- Savigny, H. (2002). Public Opinion, Political Communication and the Internet. *Politics*, 22(1): 1–8.
- Schneider, S. (1996). Creating a Democratic Public Sphere Through Political Discussion. *Social Science Computer Review*, 14(3): 373-392.
- Schoen, H., Gayo-Avello, D., Metaxas, P., Mustafaraj, E., Strohmaier, M., Gloor, P. (2013) The power of prediction with social media, *Internet Research*, Vol. 23 Iss: 5, pp.528- 543. 10.1108/IntR-06-2013-0115
- Seegerberg, A. & Bennett, W. (2011). Social media and the organization of collective action: Using twitter to explore the ecologies of two climate change protests, *Communication Review*, 14(3): 197-215.
- Sexton, J. & Helmreich, R. (2000). Analyzing cockpit communications: The links between language, performance and workload. *Human Performance in Extreme Environments*, 5, 63-68.
- Shi, L., Agarwal, N., Agrawal, A., Garg, R. & Spoelstra, J. (2012). *Predicting US Primary Elections with Twitter*, unpublished manuscript. URL: <http://snap.stanford.edu/social2012/papers/shi.pdf>
- Slatcher, R. et al (2007). Winning words: individual differences in linguistic style among U.S. presidential and vice presidential candidates. *Journal of Research in Personality*, 41, 63-75.
- Smith, A. (2009). *The Internet's Role in Campaign 2008*. Pew Internet & American Life Project. Washington, DC: Pew Research Center.
- Spierings, N. & Kristof J. (2013). Getting Personal? The Impact of Social Media on Preferential Voting. *Political Behavior*, DOI 10.1007/s11109-013-9228-2
- Strandberg, K. (2013). A Social Media Revolution or Just a Case of History Repeating Itself? – The Use of Social Media in the 2011 Finnish Parliamentary Elections, *New Media & Society*. 15(8), 1329–1347.
- Surowiecki, J. (2004). *The wisdom of crowds*, New York: Doubleday.
- Tausczik, Y. & Pennebaker, J. (2010). The Psychological Meaning of Words: LIWCe Computerized Text Analysis Methods. *Journal of Language and Social Psychology*, 29, 1, 24-54.
- Thimm, C., (2012), Political Conflict, Microblogging and the changing Role of the Citizens: Examples from Germany, Tunisia and China. *Proceedings of the 4th International Conference on Mobile Media for*

- Development*, p. 388-398.
- Tjong, E. & Bos, J. (2012). Predicting The 2011 Dutch Senate Election Results With Twitter. Association for Computational Linguistics. *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, 53-60.
- Toni, A. (2011). *Los guardianes del mensaje*. Barcelona: UOC.
- Tourangeau, R. & Plewes, T. (2013) *Nonresponse in Social Science Surveys: A Research Agenda*, The National Academies Press: Washington, D.C.
- Tufekci, Z. & Wilson, C. (2012). Social Media and the Decision to Participate in Political Protest: Observations From Tahrir Square. *Journal of Communication* 62(2): 363–379.
- Tufekci, Z. (2014). Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls. In *ICWSM '14: Proceedings of the 8th International AAAI Conference on Weblogs and Social Media*, 2014.
- Tumasjan, A. et al. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. *Association for the Advancement of Artificial Intelligence (AAAI)*, 178-185. Disponível em: <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/view/1441>. Consultado em 22 de junho de 2016.
- Wei, L. & Hindman, D. (2011). Does the Digital Divide Matter More? Comparing the Effects of New Media and Old Media Use on the Education-Based Knowledge Gap. *Mass Communication and Society* 14(2): 216–235.
- Williams, C. & Gulati, G. (2008). What is a Social Network Worth? Facebook and Vote Share in the 2008 Presidential Primaries. *Annual Meeting of the American Political Science Association*, 1-17. Boston: MA.
- Woodly, D. (2007). New competencies in democratic communication? Blogs, agenda setting and political participation. *Public Choice*, 134(1-2): 109-123.
- Yu, B., Kaufmann, S. y Diermeier, D. (2008). Exploring the characteristics of opinion expressions for political opinion classification. *Proceedings of the 2008 International Conference on Digital Government Research*. Montreal, 82-91.
- Zúñiga, G., Jung, N. & Valenzuela, S. (2011). Social media use for news and individual's social capital, civic engagement and political participation. *Journal of Computer-Mediated Communication*, 3, 319-333. Disponível em: http://www.academia.edu/539230/Social_media_use_for_news_and_individuals_social_capital_civic_engagement_and_political_participation. Consultado em 5 de agosto de 2016.