

Chapter 2

DO EGYPTIANS STILL CARE ABOUT THE ARAB SPRING? COMPUTATIONAL CULTURAL ASSESSMENT OF ONLINE AND OFFLINE ACTIVISM

Emad Mohamed and Eid Mohamed

Introduction

This chapter analyses change and continuities in a post-Arab Spring era, with a focus on the Egyptian case, through a preliminary framework of analysis of our understanding of and approach to social, cultural and political change/transformation, and continuities, hopes and frustrations in Egypt, and by corollary in the Arab world, for the past thirty years. What mainly came out of it is that scholars should focus on change in a multilayered way to avoid the clichés carried by the analysis of ‘brutal’ change after the revolutionary ‘awakening’ or its consequent authoritarian ‘stagnation’. The main problem with this approach is that it only considers change in terms of rupture, transition, modernity and tradition. Considering change as a multi-influential, multi-scaled long-term process may help us analyse the visible and invisible transformations of Arab identities and social realities in a finer way.

One interesting methodological approach would be to comprehend change through the analysis of the changing use, reception and circulation of specific media and news stories (i.e. our research material). Here it can be a good segue into how the Egyptian revolution sparked a change. Some scholars argue that, in fulfilment of the process of change, there was a failure to complete a democratic transition because of the coup. Others argue that there was a change in the way Egyptians defied the laws set in place by the Mubarak regime, which restricted Egyptians from these forms of expression, and that this defiance constituted a major gain of the 2011 uprising. While there is no readily available information on how these changes apply today under the Sisi regime, it does not mean that such changes are not occurring. For a long time, under the repression of the Mubarak regime, activists remained underground. It was only after the fall of Mubarak that their activism came out into the public sphere.

There are several scales of analysis we can use to understand change in the Egyptian case: individual or collective, daily or historical (at the national

level) and local or global. We can superimpose these levels (and media such as literature, personal biographies) on one another in order to understand how past social narratives are constructed differently in the present. How Egyptians reconstruct and understand the past (not only their role as individuals, but also their perception of notions, such as the West, the Arabs, the revolution, etc.) to redefine their present agency could be one of the main research questions of this article.

The approach we employ in this research is both quantitative and qualitative. At the quantitative level, the chapter uses tools and algorithms from computational linguistics and data mining to extract useful information from the corpus from the Egyptian newspaper *Almasry Alyoum* (The Egyptian Today). This information is then used build a regression model that can be used for both explanation and prediction. The qualitative side of the analysis makes use of this quantitative analysis to gain, and provide, meaningful insights that can be used in the historical and socio-political aspect of the project. Choosing *Almasry Alyoum* is based on its being a successful independent newspaper which is able to challenge taboos, criticize the regime's performance and at the same time maintain professional standards of journalism in terms of accuracy, objectivity and credibility.

Below are the most interesting themes found in the post-2011 Arab Spring literature that we reflect upon.

A critique of how the Arab Spring had not been predicted

Many scholars note that overstatement of the robustness and stability of Arab authoritarianism led to a culturalist myth of Arab fatality or exceptionalism. This view has been shaped by researchers who took a holistic approach to the question of Arab people's agency and then their inability to have predicted the revolt. This is the main critique we find in the literature when talking about how to assess the post-Arab Spring change in Egypt. Some macro causal approaches, though, explain that change is simply not predictable, since it depends on 'elite mis-assessments of the situation' or their anticipation of contestation through ad-hoc reforms (Volpi 2013, p.969). The critique is based on the lack of attention to the salience of a cross-border Arab identity (Gause 2011) and an intense focus solely on regime change (Kohstall 2016), obstacles to democratization (Schwedler 2015) and organized/institutionalized social movements as the main actors of transformation (Morjé and Walters 2014). Again, we outline that some authors (Gause 2011; Kohstall, 2016; Volpi 2013) label the revolutions as failures because there was no regime change, or in some cases there were illegitimate regime changes. This is another point that is in accordance with the examined literature that points to lack of regime change as the sole cause of decline in the state of the country.

Indeed, neither the revolution itself nor its trajectory and current outcome were predictable. Such predictions are difficult if not impossible to make. Some researchers have been trapped by the surprise effect of the revolutions and built biased timelines in order to explain them. One critique that can be found in these

works is describing the use of media as something ‘new’ without taking into account the ‘ecologies of the anti-authoritarian uprising – that is, the availabilities of different forms of communication to different actors involved in the revolutionary processes at different points in time’ (Rinke and Röder 2011).

The same can be said about analysing ideological changes and social movements. Islamists, for example, were not suddenly changed by the revolution but by the specific context of transition that followed. It allowed them to expose the post-Islamist transformations they were experiencing even before the toppling of dictators (Bayat 2013b). Some authors explain that Western researchers have been trying to commodify changes experienced in the Arab world for the past fifteen years. They have done so by defining a specific, limited period of time – a timeline – they called ‘Arab Spring’ as a way to keep control over the intelligibility of these changes. This ‘temporal Othering’ renders ambiguous processes intelligible, manageable and ultimately securable under the rubric of democratization (Hom 2016). A longitudinal approach is recommended with a focus on individuals’ diverse relationships to the revolutionary process.

This idea of unfolding the Arab Spring as a process is also found in the analysis of the ways communication occurred ‘with the interplay of the media available to different actors, the cultural contexts of protest communications, and the different stages of the lifetimes of such movements’ (Rinke and Röder 2011, p. 1276). Still in the vein of focusing on parallel trajectories of change, others have underlined the importance of understanding state reformation efforts (and the idea of the nation) according to varying contexts of change. The various attempts to redefine the state and its consequent contentious politics could well be considered as opportunities to reactivate ‘latent identities, mobilize passive populations, and give rise to new ideologies and actors’ (Saouli 2015, p. 316). Although some suggest that ‘new technologies and social media which, capitalizing on past mobilizations, have been a new and relevant feature in the protests and which may change future mobilization patterns for good’ (Lynch 2014, p. 6), there is a growing consensus that agency and protests should also be looked at by understanding the continuity of relationships between individuals and their state, their communities, the public sphere, the police, the economy, etc. (Bogaert 2013) that allowed, triggered and sustained them (Hinnebusch 2015). The heterogeneity of timeframes at work in the aftermath of the uprisings also includes some reflections on continuing patterns inherited from colonialism (under new forms) (Rivetti 2015). Finally, cliché temporalities that need to be debunked include labels such as ‘Arab Spring’/‘Islamic Winter’, ‘leaderless and spontaneous mobilizations’, ‘the Arab street’, ‘Facebook revolution’ and ‘failed revolutions’.

The literature on how to comprehend change in post-Arab Spring has also revealed deep fractures and frustrations between Western and local Arab scholars. Of course, some are quite superficial. But the question of why some discourses are heard more than others and what consequences this de facto hegemony may have on the way they see, trust or distrust each other (and also the impact this division may have on the actors of Arab change themselves) seems to be a really interesting question. There are two main critiques emanating from the Arab world:

1. The behaviours of some Western researchers who focused in an opportunistic way on fashionable countries. Suddenly, a large number of writers were researching Egypt or Tunisia and writing about these countries without having visited them even once, and they interacted with local researchers as native informers (Abaza 2017) with no plan of building a common knowledge.
2. The flood of publications from US universities and think tanks has contributed to orientating research on the Arab Spring and influencing perceptions of change that are linked to specific topics, such as the victory or failure of Islamist parties, neutralization or return of the military (75 per cent of the scholarly production comes from outside the Arab World) (Maghlouth et al. 2015).

Feasible alternative: An exploration of everyday social change

Unfortunately, the coup on 3 July 2013 in Egypt has changed the scene of activism so much that it relatively impedes what we are proposing here about change and the expression of the willingness for change. This is not to say that there is no research available, since the literature we surveyed highlights the significance or the impact of the revolution – in anxieties, perceptions, hopes and dreams – on local communities and daily life. But the historical distance from that event goes against our investigation of evidence for the change we want to map out, not because it is not or was not there, but because we have a reversal of events that took place after the coup; whatever change happened has been forced into hiding – not removed, we believe, but hidden. Added to this are the restrictive measures adopted by the counter-revolutionary forces either to alienate and oppress revolutionary figures and movements or to suppress and manipulate revolutionary processes and incentives. Examples of these restrictive measures include, without being limited to, print media, state-/self-imposed censorship and state-controlled internet trolls. However, we do not think that there is no form of online activism taking place; there has to be. Considering the fearlessness and defiance that were evident under the former regime, we believe that it is hidden and underground. Unfortunately, there is insufficient research to back this up, since the available research does not study this; it just stops at reporting repression. Here, it is about an academic approach to change (in politics, identity, public spaces, etc.) from below. While protests are not dismissed and are still considered a major repertoire of change, researchers are more interested in the self-centred daily concerns that are pushing people to hope for change (through the streets, mobilizations, etc.) (Bayat 2013b).

The corpus and material gathered by individuals themselves has shown greater interest in expressing willingness (or fear) of change:

It is through their humor, satire, images and novels that we can get a sense of the discontent of youth that might lead the next revolution. The crowd re-articulates history as local, connected and malleable to the will and desire of political subjects, as opposed to the will of a corrupt dictator.

(Khalil 2012, p. 63)

There also is a lot on how people are rewriting changes they have experienced through autofictions and biographies, including, for example, *Tahrir Monologues* (Tahrir Monologues 2012).

An interesting idea is the importance of localizing change: how people retranslate and understand the revolutionary 'grand schemes' (Schielke 2015, p. 13) in their daily lives and what kind of anxieties and fears may arise from the discrepancy between the revolutions' promises and local experiences (Belakhdar et al. 2014; Rivetti 2015). Overall, the idea that informal daily politics, like the relationship between online and offline activism (Zayani 2015), gave us more information about how change occurred and how it has started to be widely shared since 2015–16.

Data and methods

In spite of the fact that it is really hard to monitor and analyse formal daily activities, there is an option of using a proxy to understand how people in countries of the Arab region interacted with the news of the day. Instead of analysing people's reactions and actions directly, we measure how much they interacted with media. For this purpose, we analysed a corpus of 169,000 news articles from newspapers. We focus the discussion on the number of times these articles were shared on Facebook to measure people's interest. Our working hypothesis is that people share only that which they feel strongly about. Sharing does not, however, mean that the sharers share the same views as the writers. Sharers may pass along those ideas because they strongly object to them. By analysing people's sharing patterns and interest, we hope to gain some insight into the mindset of the people of the region at a specific point in time.

Data

According to the web-intelligence website Similarweb.com, the website of the Egyptian newspaper *Al-Masry Al-Youm* (Almasryalyoum.com) ranks 2,699th on the list of the top visited websites in the world. It is also the 23rd most visited website in Egypt and, in the category of news and media, *Al-Masry Al-Youm* ranks 389th, which makes it one of the most important venues for news and political commentary in the Arab world in general and in Egypt in particular. Most of the traffic for the website comes from Egypt (69 per cent), followed by Saudi Arabia (7 per cent), the United States (3 per cent), the United Arab Emirates (3 per cent) and Kuwait (1.38 per cent).

Methods

To answer the questions posed in this chapter, we adopt a quantitative approach and supplement it with qualitative analysis. We use a pipeline of natural language

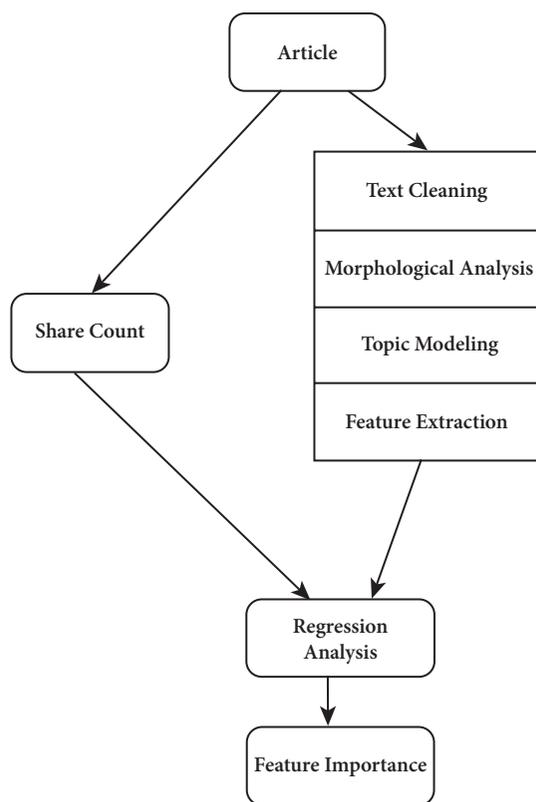


Figure 2.1 A flowchart of the research pipeline showing the processes a document goes through until feature importances are extracted.

processing and machine learning. The pipeline is outlined in Figure 2.1 and goes as follows:

Every article in our corpus undergoes the analysis pipeline outlined in the flowchart above. Each one of these steps deserves some consideration.

Text cleaning

The corpus we use, Almasryalyoum.com, is a website in html. In order to work with the data from the website, we extract the text from the html, making sure the encoding is uniform. This is probably the easiest kind of textual processing. This step also includes minimal processing, such as separating punctuation from text and separating numerical from alphabetical characters. This step is absolutely necessary, because any information-extraction task depends on plain text, and the form in which webpages naturally occur has many tags, comments and metadata that are not conducive to the kind of research we do.

Morphological analysis

Arabic is a morphologically complex language. While in many languages a word can be defined as a whitespace-delimited unit, such units in Arabic are far more complex. The word *wllmSryyn* ('as and for the Egyptians') comprises four units: the conjunction *w* ('and'), the preposition *l* ('to'/'for'), the definite article *Al* in its assimilated form and the noun *mSryyin* ('Egyptians'), which is itself made up of the adjective *mSry* ('Egyptian') and the genitive sound plural masculine marker *-yn*.

For morphological analysis, we use the memory-based segmenter and tagger developed by Mohamed (2012). Given a sentence such as:

وأعلن الرئيس ميادرتة في المؤتمر المشترك مع الرئيس الأمريكي

The morphological analyser produces the following (Table 2.1):

The purpose of morphological analysis is twofold: (1) mapping different forms of the same word together and (2) disambiguation. Mapping different forms of the same lemma can be exemplified by the word *ktAb* ('book'), which can appear in so many different forms: *kitAby* ('my book'), *Alkutub* ('the books') and the plural *kutub*. It does not make much sense to treat these as different words; hence the importance of mapping.

Disambiguation is usually a result of part-of-speech (POS) tagging, a process by which a grammatical category is assigned to each word/lemma/token. For example, the Arabic word *علي* can be either a proper noun (the name 'Ali') or the preposition *على* plus the first person singular pronoun *ي*, in which case it means 'on me'. The combination of morphological segmentation and POS tagging achieves this disambiguation.

As depicted in Figure 2.2, the word *wllmSryyn* has two levels of morphological analysis. At the second level, the token *mSryyn* can be divided into two

Table 2.1 The morphological analysis (segmentation + part of speech tagging) for one Arabic sentence.

Form	Analysis
وأعلن	wa/CONJ+AElan/PV
الرئيس	Al/DET+r}ys/NOUN
ميادرتة	mbAdrp/NOUN+h/POSS_3MS
في	fy/PREP
المؤتمر	Al/DET+m}tmr/NOUN
المشترك	Al/DET/m\$trk/ADJ
مع	mE/PREP
الرئيس	Al/DET/r}ts/NOUN
الأمريكي	Al/DET+>mryky/ADJ

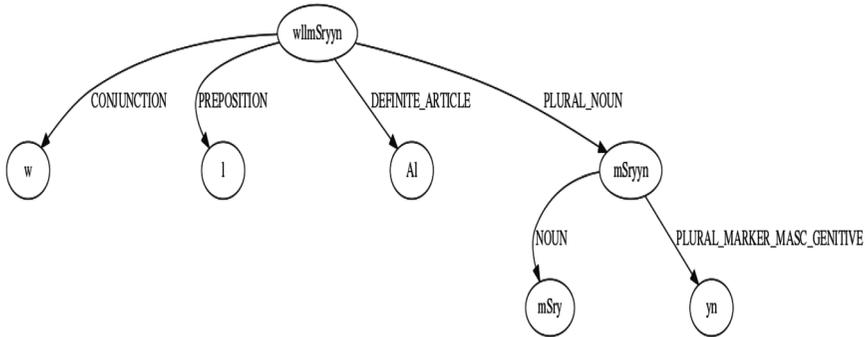


Figure 2.2 The morphological analysis of the Arabic word *wllmSryyn* ('and for the Egyptians').

segments: a noun and a suffix. Since this suffix is inflectional, we have decided, in our scheme of segmentation, not to split it off. We end up with the word having four segments, with the lexical unit being *mSryyn*. This level of segmentation is suitable for topic modelling since it reduces the vocabulary and at the same time maintains interpretability.

Topic modelling

Topic modelling is a statistical model that clusters documents into topics and is widely used in the humanities. It is capable of discovering the thematic structure of a large corpus, and it assumes that each document may contain a number of topics (Steyvers and Griffiths 2007). This is very important for our purposes, since an article that talks about both elections and torture in prisons may easily be assigned two topics. Usually, the researcher determines the number of topics wanted and the software arranges the documents according to the most prominent topics. The choice of number of topics is a question of trial and error and basically depends on finding the most topics with the least noise. Topic modelling is especially good when dealing with large bodies of text. With thousands of documents, it is usually unfeasible to go through them manually. Topic modelling analyses those documents and assigns clusters of words to them (we call these clusters 'topics'). A document may be associated with several topics with varying probabilities. For this study, we use the MALLET topic-modelling toolkit (McCallum 2002) because it is open source and easy to use.

We use the topic modelling output for three purposes: (1) obtaining a topic classification for all the topics in the corpus; (2) ranking the most important topics according to the *Al-Masry Al-Youm* website, which gives us an idea about what the newspaper focuses on as well as what it neglects; and (3) the topics produced by the topic modelling algorithm are fed into the random forest regressor as independent variables while the number of shares per document is the dependent variable. The objective here is to rank the importance of the topics as viewed by the newspaper readership and followers. The whole setting allows us to compare the interests of the editorial staff with those of the readership. We can then pick any

topic covered in the newspaper and see how it ranks among the netizens of social media as compared to how much focus the editors give to it. For the purpose of this analysis, we will focus only on those topics related to the Arab Spring.

Feature extraction

We measure the importance to the readership by the number of shares a document has. A document that is shared more is considered to be more important (and vice versa). Shares represent active engagement with the document, unlike readings or views, which are more passive. Reactions can also be active engagements, but they are not as indicative of importance as shares are. Comments can also be considered active engagement, but comments are only active at the personal level. A commenter wants to express her opinion, but a sharer wants others to express their opinions as well. If it is true that a Facebook user has an average of 400 friends, shares can also tell us much about the reach of the article shared, which cannot be accounted for by other modes of engagement.

The choice of shares over comments, the only two measures available in our datasets, can also be justified by looking at the data. In our dataset, people seem to be more keen on sharing than on commenting. The average number of shares per document is 364.19, with a standard deviation of 1,510, a minimum of 0 and a maximum of 66,696. The median number of shares per document is 3, which shows a skewed distribution. Obviously, there are some documents with a large number of shares, while 50 per cent are shared less than three times each. A third quartile of 82 makes the picture clearer.

When it comes to comments, the average number of comments per document is 0.34, with a standard deviation of 2.93, a minimum of 0 and a maximum of 428. Also, 90 per cent of the documents in the corpus do not have any comments associated with them, while only 27 per cent of the documents have not been shared at least once. For these reasons combined, we have decided to use the share count rather than the comment count as our indicator of topic importance.

While share count tells us much about how significant an article is, it tells us nothing about how essential it is to the editorial board of the newspaper in question. However, this significance can be inferred from the topic(s) of the article. If the topics rank highly in the topic modelling stage, this may indicate that the article expresses an important topic. This is vital since, by combining these two metrics, we can compare the interest of readership to the interest of the editorial board. Such a finding can be central not only for research purposes but also for commercial applications, since we can use this to tell the editorial board what topics readers are interested in, thus making them focus more on what the readers know is important and not what the editorial boards think will grab the readers' attention.

Regression analysis

Random forests is a machine-learning algorithm that can be used for both classification and regression. It is capable of handling large numbers of variables, even when the information in each variable is limited, without being prone to

over-fitting. Random forests are known to have high accuracy and to be ‘relatively robust to outliers and noise’ (Breiman 2001, p. 10). We have found random forests to be superior to ordinary least squares (OLS) multiple regression since we have tried several algorithms for the purpose of analysing the data for this chapter. One favourable feature of random forests is that they rank the independent variables in terms of their importance to the regression/classification model. To be specific to the problem at hand, random forests are capable of telling us which factors are most associated with sharing. Given a topic-classification scheme of the newswire articles (e.g. politics, religion, January 25th Revolution), random forests produce a ranking and a score for each. This is indispensable for our purposes, since this is exactly what we are looking to achieve. While there are several implementations available, we have used the scikit-learn implementation (RandomForestRegressor) for analysis in this chapter for its ease of use and programmability.

Results

Preliminary statistics

We use a corpus from the Egyptian newspaper *Al-Masry Al-Youm*. The corpus covers the period from 1 June 2016 to 31 May 2017, and is thus a full year of coverage. The corpus comprises 169,026 documents, where a document is any published item (e.g. news item, opinion piece, interview). The number of items per month in the corpus over this period is as follows: 14,944 (June 2016), 13,862 (July 2016), 14,625 (August 2016), 13,589 (September 2016), 14,891 (October 2016), 14,260 (November 2016), 13,489 (December 2016), 14,525 (January 2017), 12,787 (February 2017), 13,310 (March 2017), 14,804 (April 2017) and 13,940 (May 2017). In terms of word count, the one-year corpus has 41,260,327 words, not including punctuation, with an average of 246.11 words per document.

Topic Model Topics

We have decided to use forty topics for topic modelling, which is an easy number to handle. We have also tried up to one hundred topics, and found that we gain more fine-tuned details by increasing the number of topics, and that forty is a good approximation and valid generalization. The following are the forty topics suggested by MALLET, sorted in descending order of importance in the corpus. The more important topic appears first:

It should be noted that the order in which these topics appear reflects how important they are to the editors of *Al-Masry Al-Youm*. We can see also that the focus is mainly local and that international news takes a back seat, which is expected in a newspaper whose title translates to ‘The Egyptian Today’. This, of course, does not mean a binary classification between local and international news, since many news items are essentially both. What this tells us is that, across

Table 2.2 T# = Topic Number, MR = MALLET Rank, RG = Importance in Regression Model

T#	MR	RG	Description	T#	MR	RG	Description
29	1	3	The Problems facing Egypt	8	21	12	Theft and other crimes
34	2	6	The news of the Parliament	23	22	22	Health and hospitals
36	3	10	Local news from different districts	30	23	26	Europe, refugees, elections
32	4	13	Projects for boosting the economy	11	24	4	Food supplies: wheat, sugar and meat
25	5	8	The relations between Egypt and the other Arab countries	19	25	24	Tourism
33	6	1	Media coverage of President Al-Sisi	16	26	39	ISIS
18	7	14	Conduct of the people of Egypt	38	27	17	Egyptian ports
0	8	21	Police, incident and explosions: most probably foreign news	2	28	9	Military operations in Sinai
5	9	2	Family and marital issues	6	29	27	China, North Korea, Iran and Nuclear programmes
3	10	34	Women and children	26	30	7	Movies, TV series and songs
14	11	18	Education and universities	27	31	30	Agriculture, Nile, irrigation
37	12	23	Utilities and roads	17	32	20	Azhar, Mosques and fighting terrorism
7	13	28	US elections	31	33	32	Church, explosions and terrorism
4	14	36	Culture, arts and festivals	28	34	25	Sports news
12	15	15	Crime and accident news	22	35	37	European football: Barcelona, Manchester United, Bundesliga
24	16	5	Banks and the stock exchange	35	36	35	African football
9	17	19	Traffic	20	37	29	diet and sleep
10	18	16	January 25th Revolution, political parties, the Egyptian army and the Muslim Brotherhood	1	38	33	Cars, internet and technology
21	19	31	Ahli and Zamalek	15	39	40	Brazil Olympic Games
39	20	11	Criminal courts, possibly related to the Muslim Brotherhood	13	40	38	The Israeli occupation of Palestinian territories, Hamas and the West Bank

the entire corpus, there is more focus on the local part of the story, and a piece of international news may derive its influence on the local news.

Out of forty topics, the January 25th Revolution and its related news ranks 18th, which is somewhat in the middle. It is not priority news, but it is not in the bottom either.

Topic association

The focus of this article is the Arab Spring in Egypt, which is Topic 10 in Table 2.2, but a topic that hardly comes in isolation. When a writer mentions the January 25th Revolution, even when the revolution is the main focus of the article, she usually relates this topic to other relevant ones. She may talk about the detainees, violations of human rights, the economic situation in Egypt, among other topics. Since we have summarized the whole corpus into forty topics, we will now see which of these topics co-occurs with the revolution topic in the same document. For the sake of brevity, we have excluded any topic whose contribution to the document is less than 10 per cent. This guarantees that topics mentioned in passing are not given as much importance as those constituting an integral part of the document.

Table 2.3 lists the top ten topics associated with the Arab Spring. The ten topics together make up 70 per cent of all the topics associated with the Arab Spring in Egypt. The most important associated topic is Topic 29, the problems and challenges facing Egypt. Topic 29 alone represents 16.4 per cent of topic associations and, being negative in nature, tints the Arab Spring topic with a negative tone.

Table 2.3 Topic Associations: The top ten topics associated with Topic 10

Topic	Frequency	Percentage	Cum%	EditorRank	PeopleRank
29 Problems of Egypt	3665	16.4107	16.4107	1	3
18 Conduct of the people of Egypt	3204	14.3465	30.7572	7	14
34 Parliament	1414	6.3314	37.0886	2	6
5 Art, science and revolution	1377	6.1658	43.2544	9	2
7 US elections	1369	6.1299	49.3843	13	28
25 Egypt and other Arab countries relations	1185	5.306	54.6904	5	8
33 Media coverage of President Al-Sisi	1083	4.8493	59.5397	6	1
16 ISIS	858	3.8418	63.3815	26	39
4 Culture, arts and festivals	848	3.7971	67.1786	14	36
30 Europe and the refugees	768	3.4389	70.6175	23	36

The second most related topic is Topic 18, which is mainly about the people of Egypt, talking primarily about the characteristics of the people and what they have been through. This connects smoothly with Topic 29, and we have the triangle of revolution, people and problems.

Less important associated topics follow. First, we have Topic 34: the Parliament. The general sentiments in the documents about the Parliament seem to be negative ones, basically claiming that this is not the kind of assembly one elects after a revolution.

While it makes sense to look at the cluster of topics representing the Arab Spring in terms of their ranked importance, this is not the only way to examine them. One important aspect is the difference in interest between editors and readers. The editors may think some topic is important while the readership views it otherwise, and vice versa. It thus benefits our analysis to examine the losers and winners as topics move from their static status on the website and their shareability on social media.

The biggest winner is Topic 5 (family and marital issues). While the topic ranks ninth in terms of editor interest, it is the second most shareable topic in its category and thus earns seven ranks in the move. The next biggest winner is Topic 33 (Sisi and the media), which wins five ranks as it moves up from rank six to rank one. These two topics are the only winners. All the other topics lose rank as they move from the website to social network.

The biggest loser is Topic 4 (culture, arts and festivals). While this is the 14th most covered topic on the website, sharers do not seem to care much about it, since it ranks 36th out of forty on the sharing scale, thus losing 22 ranks. The second biggest loser is Topic 7 (the US elections and President Trump). US elections and Trump's presidency seem to be a favoured topic for the editors, ranking 13th out of forty, but readers do not find it worth much sharing on social media, and it thus loses 15 places and ranks 28th.

Below we will talk briefly about the top five of these topics:

Topic 29: Problems of Egypt

Topic 29 has the basic theme of 'Egypt is in crisis', with politics being the term of focus on all the crises facing Egypt: how young people are treated, detained in police custody and denied their rights and equal share of opportunities; the Muslim Brotherhood; the Egyptian parliament seen as unrepresentative of the people; the faltering economy; Egypt's rank in education and university ranking; Israel and Trump; the water crisis; the worsening conditions of tourism; automobile prices; religious discourse; and crises in the cinema industry and football. Some little focus is also paid to the conditions of the health sector in Egypt.

Topic 18: The deteriorating social demeanour in Egypt

Topic 18 can be summarized as 'how the conduct and social demeanour of the people of Egypt has changed for the worse'. In this broad topic, we find several

examples: how young people do not show enough respect to the elderly, how the streets of Egypt are filled with garbage because people do not clean up after themselves, how many people are on a mission to find fault with the government instead of doing the right things themselves, Egypt's cultural degeneration, religious fanatics and many other social phenomena that may indicate that Egypt is headed in the wrong direction. All this relates to the January 25th Revolution, either in terms of causation or correlation.

Topic 34: The Egyptian Parliament, elections and decrees

Topic 34 has to do with every elected body that makes resolutions. In this topic, we find discussions of several parliamentary bills and laws; decrees by the president of Egypt that will have to be ratified by the Parliament; the Parliament and the nomination of a new president for the State Council, the highest administrative court in Egypt; and even news about the elections in several professional syndicates in Egypt. The main theme is elections and resolutions, a topic that ties in well with the January 25th Revolution.

Topic 5: Art, science and the revolution

Topic 5 discusses the relationships between art, science and the revolution. We find, for example, an interview with the Egyptian pop singer Amina, who declares that she will not sing in Qatar because Qatar is against the June 30th Revolution; an article discussing the relationship between art, football and politics; the TV series *Mamoun and Associates*, which is a political satire and parody of the Mubarak regime; a discussion of the benefits of scientific revolutions and the harms of political ones; and many other themes related to how the political sphere in Egypt has affected art, sport and science.

Topic 7: US elections

Although we call this topic US elections, the truth is that it is about elections around the world in general, with the United States having the lion's share of coverage. We have pieces on the French elections and racism, the attempted military coup in Turkey and even the Korean elections. The largest part is concerned with Obama and Trump and compares their policies towards the Middle East in general and Egypt in particular.

Regression analysis

The first thing to be noticed is that R^2 is quite high, with a score of 0.86, which means that this may have a high predictive power. In other words, the model encompassing the thematic topics is a good predictor of sharing activity. If we know what topics a certain document contains, then we have a good idea of how

many shares it will get. Since we have forty topics, and each document is usually made up of several of these, with varying contributions, one needs to know which of these topics may trigger the sharing instinct. One good thing about random forest regressors is that they are able to rank the factors contributing to the sharing, assigning each factor its relative importance.

Conclusion

We have presented a new method of studying media, and while qualitative study is important, we have chosen to have a quantitative basis on top of which qualitative analyses can be realized. In the same vein, Bennett and Segerberg (2011, p. 21) stress the role that digital media technologies play in intertwining individualized and collective forms of protest, and how 'personalized communication technologies ... enable a large number of people to become linked to and recognized by a large number of others'. This individualized/collectivized protester figure has become a powerful actor in the socio-political scene, where Tahrir Square serves as an ideal example that encompasses both individual and collective participation and engagement, as well as local, national and even transnational causes.

Three lessons can be learned from this study. First, quantitative studies, especially ones that make use of large corpora, natural language processing and machine learning, are capable of producing results that are on par with, or even more insightful than, the ones produced by traditional methods. Second, digital humanities methods can be useful not only for the kind of analysis produced here but also in giving feedback to content creators. We have seen that regression analysis has shown that the editors' interests are in full agreement with those of the audience targeted by the editorial board. Editors can make use of this information to improve their content creation and writer selection processes. While we have not included the writer/editor/correspondent in this study, such a process is trivial and can tell us which writers attract a wider readership. Third, with the help of the analysis provided above, we are now in a better position to answer the question we started with: Do Egyptians still care about the Arab Spring? While this is a yes/no question, we prefer to give the more nuanced answer available through topic rankings as they relate to the Arab Spring. Both the editors of *Al-Masry Al-Youm* and the readership agree that the revolution is not a priority topic: the editors rank it 18th while the readership ranks it 16th. More important to the people are what may be described as the ramifications of the January 25th Revolution, especially what may be termed the problems facing Egyptians.

This study invites further research on the diverse tracks of the developments of these topics over time, perhaps at fixed intervals to see how time can affect the process. Also, the focus on Egypt, justified enough as a representative case study in the Arab world, allows for casting a wider net to include other Arab countries as well. Our regression models may also include other factors, such as the writers and the time of publication.

References

- Abaza, M. (2017). 'Academic tourists sightseeing the Arab Spring', *Jadaliyya*, 27 September. Available at: www.jadaliyya.com/pages/index/2767/academic-tourists-sight-seeing-the-arab-spring [accessed 1 September 2018].
- Bayat, A. (2013b). *Life as politics: How ordinary people change the Middle East*. Stanford, CA: Stanford University Press.
- Belakhdar, N., Eickhof, I., El Khawaga, A., El Khawaga, O., Hamada, A., Harders, C. and Sandri, S. (eds.). (2014). *Arab revolutions and beyond: Change and persistence*, Proceedings of the international conference, Tunis, 12–13 November 2013, Working Paper No. 11. Berlin: Center for North African and Middle Eastern Politics, and Cairo: Faculty of Economics and Political Science. Available at: https://www.polsoz.fu-berlin.de/polwiss/forschung/international/vorderer-orient/publikation/working_papers/wp_11/WP11_Tunis_Conference_FINAL_web.pdf [accessed 11 January 2020].
- Bennett, W. (2012). 'The personalization of politics: Political identity, social media, and changing patterns of participation', *The Annals of the American Academy of Political and Social Science*, 644, pp. 20–39. Available at: <http://www.jstor.org/stable/23316140> [accessed 4 May 2021].
- Bennett, W.L. and Segerberg, A. (2011). 'Digital media and the personalization of collective action', *Information, Communication & Society*, 14(6), pp. 770–99.
- Bogaert, K. (2013). 'Contextualizing the Arab revolts: The politics behind three decades of neoliberalism in the Arab World', *Middle East Critique*, 22(3), pp. 213–34.
- Breiman, L. (2001). 'Random forests', *Machine Learning*, 45(1), pp. 5–32.
- Gause, G. III. (2011). 'Why Middle East Studies missed the Arab Spring: The myth of authoritarian stability', *Foreign Affairs*, 90(4), pp. 81–90.
- Hinnebusch, R. (2015). 'Conclusion: Agency, context and emergent post-uprising regimes', *Democratization*, 22(2), pp. 358–74.
- Hom, A.R. (2016). 'Angst springs eternal: Dangerous times and the dangers of timing the "Arab Spring"', *Security Dialogue*, 47(2), pp. 165–83.
- Khalil, A. (2012). 'The political crowd: Theorizing popular revolt in North Africa', *Contemporary Islam*, 6(1), pp. 45–65.
- Kohstall, F. (2016). *Beyond regime change: Middle East Studies and academic cooperation in the wake of the Arab uprisings*. Arab-German Young Academy (AGYA) – Transformation Group Working Paper No. 5. Berlin: Arab-German Young Academy of Sciences and Humanities. Available at: http://agya.info/fileadmin/user_upload/Working_Groups-images/Transformation/WPS_Academia_in_Transformation/Paper_5_Middle_East_Studies_in_the_Wake_of_the_Arab_Uprisings_Kohstall.pdf [accessed 12 May 2021].
- Lynch, M. (2014). 'Media, old and new', in M. Lynch (ed.), *The Arab uprisings explained: New contentious politics in the Middle East*, pp. 93–110. New York: Columbia University Press.
- Maghlouth, N., Arvanitis, R., Cointet, J. and Hanaf, S. (2015). 'Who frames the debate on the Arab Uprisings? Analysis of Arabic, English, and French academic scholarship', *International Sociology*, 30(4), pp. 418–41.
- McCallum, A.K. (2002). 'MALLET: Machine learning for language toolkit'. Available at: <http://mallet.cs.umass.edu>
- Mohamed, E. (2012). Morphological Segmentation and Part of Speech tagging for Religious Arabic. Proceedings of the Fourth Workshop on Computational Approaches to Arabic Script-based Languages (CAASL4). San Diego, California, USA.

- Morjé, M.H. and Walters, M.R. (2014). 'Explaining the unexpected: Political science and the surprises of 1989 and 2011', *Perspectives on Politics*, 12(2), pp. 394–408.
- Rinke, E.M. and Röder, M. (2011). 'The Arab Spring – Media ecologies, communication culture, and temporal-spatial unfolding: Three components in a communication model of the Egyptian regime change', *International Journal of Communication*, 5, pp. 1273–85. Available at: <https://ijoc.org/index.php/ijoc/article/download/1173/603> [accessed 6 May 2021].
- Rivetti, P. (2015). 'Continuity and change before and after the uprisings in Tunisia, Egypt and Morocco: Regime reconfiguration and policymaking in North Africa', *British Journal of Middle Eastern Studies*, 42, pp. 1–11.
- Saouli, A. (2015). 'Back to the future: The Arab uprisings and state (re)formation in the Arab world', *Democratization*, 22(2), pp. 315–34.
- Schielke, S. (2015). *Egypt in the future tense: Hope, frustration, and ambivalence before and after 2011*. Bloomington: Indiana University Press.
- Schwedler, J. (2015). 'Comparative politics and the Arab uprisings', *Middle East Law and Governance*, 7(1), pp. 141–52.
- Steyvers, M. and Griffiths, T. (2007). 'Probabilistic topic models', *Handbook of latent semantic analysis*, 427(7), pp. 424–40.
- Tahrir Monologues. (2012). 'Tahrir monologues: The idea', 12 May. Available at: www.youtube.com/watch?v=JtyhaPi3tsk [accessed 10 December 2017].
- Volpi, F. (2013). 'Explaining (and re-explaining) political change in the Middle East during the Arab Spring: Trajectories of democratization and of authoritarianism in the Maghreb', *Democratization*, 20(6), pp. 969–90.
- Zayani, M. (2015). *Networked publics and digital contention: The politics of everyday life in Tunisia*. Oxford: Oxford University Press.