Going Viral
To all of those dedicated to an open Internet
Going Viral

Karine Nahon and Jeff Hemsley

polity
# Contents

Figures and Tables vi
Acknowledgments viii
Preface xi
Cover Art xvii

1 Introduction: Virality of pets and presidents 1
2 What Virality Is: I know it when I see it 15
3 What Makes Something Viral I: The control of networks through gatekeeping 41
4 What Makes Something Viral II: What is everyone looking at? 60
5 What Makes Something Viral III: Caught in the viral net! 81
6 Networked Changed Societies 98
7 Afterlife 124

Notes 143
References 156
Index 175
Figures and Tables

Figures
2.1 Sigmoid curve 22
2.2 Blog reported cumulative views 23
2.3 Daily views from blog reports 24
2.4 Social vs. promoted growth and decay 25
2.5 Example of a power-law distribution 27
2.6 “Librarians Do Gaga” 28
3.1 Keith Urbahn’s tweet 47
3.2 Brian Stelter’s tweet 47
3.3 Total retweets and direct retweets 48
3.4 Social network visualization of Keith Urbahn tweet 49
3.5 Tweets-per-second traffic (May 1, 2011) 52
5.1 Power-law distribution of blogs by traffic 86
5.2 Viral lifecycle of content in the blogosphere 87
5.3 Example of a power-law 91
6.1 TED talks views 101
7.1 Daily video views of “Yes We Can Obama Song” by Will.I.Am 138
7.2 Virality: A theoretical framework 139
Figures and Tables

Tables

2.1 Distinguishing related concepts from virality 39
3.1 Information flow milestones in the Keith Urbahn case (May 1, 2011) 45
Acknowledgments

There is so much more to a book than just the writing of it. As such, we are indebted to a large number of people, who answered our questions, shared their data with us, provided some of the graphics for the book, debated with us, and helped us crystalize ideas and concepts or just contributed their time. A special thanks to Gilad Lotan of SocialFlow, who provided us data and graphics related to the viral spread of the news of Osama bin Laden’s death and the video of KONY2012; Miguel Rios from Twitter allowed us to use his graphic of Twitter volume regarding Osama bin Laden’s death; Simon Rogers of the Guardian and Alon Halevi at Google assisted us with the WikiLeaks case; Daphne Koller answered our questions about Coursera; Sarah Wachter graciously allowed us to analyze her YouTube referral data of “Librarians Do Gaga” and Jacques Hebert at MotherJones blog also shared data with us regarding the video of Mitt Romney talking about 47 percent of Americans; Tal Gallili answered questions and shared data related the R-Bloggers website; and
Acknowledgments

Alex Constant helped us gather and interpret materials about the Rosa Parks case. Elliott Karstadt, Jennifer Jahn, and Neil de Cort at Polity Press worked behind the scenes on Going Viral and guided us through the publishing process. Arlene Pritzker, Elia Finkelstein, and Melody Lutz all assisted with language edits and feedback, and the anonymous readers provided thoughtful feedback that greatly improved the quality of the text. We thank all of you.

Another category of people who deserve our heartfelt thanks are members of our research teams, colleagues, and mentors whose efforts and input made this work possible. From the retroV research team (supported by a Google Research Award) we thank Muzammil Hussain, Chris Wells, Lance Bennett, and particularly Shawn Walker, whose legendary skill with collecting and stitching together complex data sets are at the heart of a growing list of research projects. Thank you to our colleagues at the Social Media Lab (SoMe Lab, funded by the National Science Foundation), made up of Robert Mason (the Million Dollar Man), Josef Eckert (GeoJoe), and again Shawn Walker (Data Daddy), for their patience, support, input, and assistance, as well as their work on the SoMe Tools project, out of which came the data used in the UC Davis pepper spray case. Further, Jeff thanks Robert Mason for support and guidance, Kirsten Foot for hard questions, and Malcolm Parks for giving virality a chance. Karine thanks Manuel Castells and the members of the working group at the Library of Congress, where the seed for this book was planted, and both of us extend our gratitude to John Thompson for his insistence that we nurture that seed and
Acknowledgments

patience in walking us through the process of writing a book.

Finally, and most importantly, we thank our spouses, Arlene and Michael, Karine’s kids (Daniel and Ari) and mother (Ody) for encouraging, understanding, supporting, and enduring our esoteric rants on the topic of virality.
Preface

The idea for this book came about while I was visiting the Library of Congress (LOC) in Washington, DC. For me, LOC is one of the temples of culture of societies. I came to Washington on the invitation of my colleague and friend Manuel Castells, who was participating in a one-week working group about the role of the Library of Congress over the next 25 years. The growing use of social media and the explosion of user-generated content create new challenges and opportunities for libraries. One of the most pressing questions for libraries is what should libraries attempt to collect and curate in this new media environment, given the ease of creating and sharing content, the tremendous and growing volume of this content, as well as the infeasibility of libraries collecting and saving it all. This is where virality comes in. While there, I gave a talk about virality and its impact on society.

My colleagues and I from the retroV research group at the University of Washington’s Information School had been researching virality extensively for four years,
particularly what it means for society. In the meeting at LOC, I argued that viral information is one indicator of what is important to a particular society at a particular moment. I suggested that as libraries struggle to decide what to curate, a good starting point would be to consider information items that have gone viral. This is because viral events arise out of the complex interactions of many actors—individuals, companies, governments—as well as the social and cultural contexts within which they are embedded. Virality can signal what is considered important and interesting to parts of a society at a particular time, and traces of viral content may also become a way of documenting the fabric of societies. As such, future generations may find the viral events of today valuable lenses into our time, providing insights that we ourselves will miss because of our own embeddedness in the world from which these viral events arise.

Social media makes it very easy to digitally express ourselves and share that expression instantly. It enables some social activity to happen faster and reach farther out into the networks that connect us all. Part of what makes this social infrastructure so interesting is that it is made up of vast interconnected crowds all doing their own thing. Except, sometimes they come together. Sometimes the crowd finds some bit of content remarkable. In fact, they find it so remarkable that within a few days they remark on it, and share it will all of their friends and followers. In a way, a viral event is the collective voice of the crowd saying “this is important!” Whether it’s a middle-aged Scottish vocalist who makes us challenge our stereotypes, a humorous protest video
that creates a public relations nightmare for a major airline, or an uplifting news story of a flash mob where musicians played the Beatles’ tune “Here Comes the Sun” in a busy Spanish unemployment office, viral events are the remarkable bits of culture that rise to the top. Once there, they can quickly focus the public’s attention on police brutality against peaceful protestors, the personal infidelity of high-ranking officials, or the suffering of communities after natural disasters.

After my talk, John Thompson, the editor of Polity Press and one of the working group participants, insisted that I write a book, arguing that our knowledge of virality should be available to the public at large. “A short book,” he said, “it shouldn’t take long, and yes . . . there is another condition, you need to be ready in four months so it can be published in September the next year.” This was July and I had just started my sabbatical. I was resistant because I had planned to write a book about Network Gatekeeping Theory, an area I have been developing over the last ten years.

About a month later I received a persuasive email from John, emphasizing the importance of a book about virality as a bridge between different fields and audiences, all of whom are embedded in this new social media environment and affected by viral events. As he saw it, a book about viral events filled an important niche at the intersection of sociology, information science, communications, political science, and network theory. I already knew that most books about virality discussed it from within a particular academic lens, from a popular point of view, or from a marketing perspective, so I agreed.
Now that I was on board, I was left with questions. What’s the best way to tell the story using the integrated lenses of social and exact science? How does one bring the story and knowledge about virality to both a general and academic audience at the same time? Finally, as I have mentioned, one of my key interests over the past decade has been about gatekeepers, actors who, through their own discretion, control information flows. As I saw it, gatekeepers played such a critical role in viral events that I could certainly write a book just about their role. But, the book also needed to be balanced by talking about people, their role in driving these processes, the crowd, social media, and the networks that connect us all. Also, a fair amount of literature discusses the technical aspects of how information flows in networks, and, while John and I did not want to create a technical book, I knew the arguments needed to be backed up by rigorous, qualitative, and quantitative empirical research.

At this point I contacted my PhD student, Jeff Hemsley. For the past few years he has been immersed in the study of virality. His perspective was more quantitative and bottom-up. I knew that his expertise would balance my top-down research and studies of virality. He saw network structures, crowds, and the emerging social infrastructure as being primary drivers for viral events. Having worked with him on other projects in our research groups, I knew that he would be the perfect person to bring these needed perspectives. Together we decided to take up the challenge and try to tell the story from both the top-down and the bottom-up perspectives. Interestingly enough, even though we have been
researching virality together for a few years, we were surprised at just how far apart our views were. The journey has been full of arguments, debates, agreements, persuasions, and compromises.

We have found that the process of argumentation has helped us to crystalize our thoughts as we strive for accuracy and rigorousness. It has led to moments of epiphany and new insights, but also to frustration and compromise when we find that sometimes there is no perfect word; that any word may be laden with alternate meanings.

How much of virality is controllable? To what extent is it predictable? Is it a process that can be designed or does it emerge out of the whims of the crowd? What are the roles of network and social structures? These are some of the questions that, at the start, we were deeply opposed about. Sometimes, writing a book with a partner is not an easy task, but it has its advantages. Through our discussions, we have both arrived at a “center point,” one that we hope will open the door to future studies on the topic. Yes, there are still a few things we disagree about, but the rewarding part of partnership is growth and understanding. I have probably learned as much from my PhD student as he has from me, and that is how it should be.

At the outset, we knew that if we created a cookbook of how to create viral events, it could sell many copies to marketers and practitioners, but our scientific truth was different. The audience we hope to reach is our own scientific community, but we also think that understanding how viral events work, how they impact society, and their value to future generations can better prepare
Preface

people to operate in the emerging social infrastructure. As such, we have tried to make this book as readable as possible in the hope that each one of you can gain insight into the events that we call viral events.

Karine Nahon
Tel-Aviv
March 2013
“Occupying Twitter: A retweet network.” Jeff created this visualization from Twitter data captured during the days leading up to, and just after, the Occupy Oakland protesters shut down the Oakland ports on November 2, 2011. Each dot is a Twitter user and each line is a case where one user retweeted another user. The large clusters of dots typically represent viral events about the port shut down or other news related to Occupy Oakland. The disconnected clusters at the lower part of the image represent large retweet events, some viral, by people not necessarily connected to the protesters, but who otherwise are tweeting about Occupy Oakland – sometimes in harsh terms. The data was drawn from the Social Media Lab’s corpus of Occupy Tweets.
Introduction: Virality of pets and presidents

Viral information is not new. When we look back in time, before the Internet, there are plenty of examples of fast-moving information flows that reached many people and happened as a result of people sharing – the key elements of virality. One likely candidate is the news of the arrest of Rosa Parks, on Thursday, December 1, 1955. She was arrested in Montgomery, Alabama, for not giving up her seat to a white person on a segregated bus (Parks and Haskins 1999). Instead of the Internet, people used phones, hand-bills, and word of mouth. It is estimated that within three days, roughly 40,000 blacks had heard about, and joined a boycott of the bus system in protest at Rosa Park’s arrest (Dove 1999). They walked, often miles, every day to their schools and jobs. And they continued to walk for more than a year until the segregation law was repealed.

Viral events are not new. What is new is that a viral video, a news story, or a photo can reach 40,000 people in hours, or even minutes, instead of days. And it isn’t just the speed and reach of these information flows that