Original Pirate Material

Reading the Digital Archive of the Underground Warez Scene

Brief Synopsis

When most people think of digital piracy, the phrases that likely come to mind are “Bittorrent”, “Napster”, and “The Pirate Bay”; the popular manifestations and accessible incarnations of home copyright violation. However, this is a poor reflection of a submerged and elite culture of an underground piracy scene that for several decades has operated on a secretive and hierarchical basis of suppliers, couriers, release groups, and “topsites”. The true “warez scene” as it is known, is undetected by the general public, but well-acquainted with high-level law enforcement.

This book offers the first academic study of the gigabytes of digital material surfaced by “The Scene” in the form of ASCII .nfo files and DemoScene executables from the Defacto2 archive, charting the structure, organization, and history of the criminal underground networks that race to release material before their competitors with bleeding-edge technology and connections. Using a combination of traditional and digital reading methodologies, this book presents both the historical structures but also *aesthetic* strictures of the underground warez scene at the turn of the twenty-first century. As such, this book is also one of the first studies to construct a distant-ethnography from a digital archive, reading from the digital-material traces the contexts and after-images of an otherwise inaccessible digital-cultural sphere.

From the technologies of text that it examines, *Original Pirate Material* resurrects a secretive space that has had wide-ranging implications for law, media, and many other areas of contemporary cultural digital life.

Biography

Martin Paul Eve is Professor of Literature, Technology and Publishing at Birkbeck, University of London. Martin holds a Ph.D. from the University of Sussex and is the author of five other books: *Pynchon and Philosophy: Wittgenstein, Foucault and Adorno* (Palgrave, 2014); *Open Access and the Humanities: Contexts, Controversies and the Future* (Cambridge University Press, 2014); *Password* (Bloomsbury, 2016); *Literature Against Criticism: University English and Contemporary Fiction in Conflict* (Open Book Publishers, 2016); *Close Reading with Computers: Textual Scholarship, Computational Formalism, and David Mitchell’s Cloud Atlas* (Stanford University Press, 2019). He has held three large grants from the Andrew W. Mellon Foundation and before beginning his academic career, Martin worked as a computer programmer.

Past Endorsements

“[*Close Reading with Computers*] marks a major intervention in debates about digital humanities, literary criticism, and textual scholarship. It demonstrates commanding knowledge of work in these broad fields, and provides persuasive perspectives on several interrelated questions taken up singly by other scholars. […] I do not know of any other scholar who has such facility with computer reading and computational methods AND such a subtle literary critical sensibility. This is truly a unique work.” – Professor Paul Harris, Professor of English, Loyola Marymount University

“Martin Paul Eve is one of the most brilliant scholars of his generation. His ground-breaking *Literature Against Criticism* combines new and insightful readings of contemporary novelists (from Jennifer Egan to Tom McCarthy and from Sarah Waters to Percival Everett) who are in animated competition with university English. There are very few authors who can combine ethical, political and aesthetic readings of the contemporary novel with an encyclopaedic knowledge of the modern university.” – Bryan Cheyette, Professor of English, University of Reading

“Conjuring our passwords has become a daily act of our computer-saturated existence. By no means sequestered to our digital present, Martin Paul Eve's excellent account of the password covers its long and lively history. Weaving literary references with lucid technical explanations, Eve skillfully traces the evolution of password to probe its fundamental connections to issues of human identity, trust, and ownership.” – Gabriella Coleman, Wolfe Chair in Scientific and Technological Literacy, McGill University, Canada

"Martin Paul Eve's *Pynchon and Philosophy* is a work of consummate scholarship. Breaking new ground in Pynchon studies, Eve offers an immensely erudite, detailed and in-depth account of the ways in which the ideas of Wittgenstein, Foucault and Adorno help us to think about his texts. A first-rate book." - David Cowart, Louise Fry Scudder Professor of Humanities, University of South Carolina, USA

Proposed Table of Contents

Chapter 7: NFO Aesthetics

Chapter 8: Dissemination Networks

Chapter 9: Points of Emergence

Part III: Conclusion

Chapter 10: Conclusion

Bibliography

Glossary

Projected extent: 105,000 words including bibliography.

Chapter-by-Chapter Summary

Part I: Introduction

Chapter 1: Introduction [10,000 words]

This introductory chapter frames the volume by giving a contextual background to the high-level "warez" (piracy) scene; to the digital archive of information about this space; and to the reading methods that will be deployed throughout the book.

The first part of this chapter handles an overview of existing critical and historical thought on piracy. This ranges from long-standing approaches to intellectual property and copyright – such as Adrian Johns's *Piracy* and John Willinsky's recent *The Intellectual Properties of Learning* – through to contemporary legal wrangles about piracy in the digital age. Along with various thinking on non-rivalrous digital objects (Suber, Hui), this philosophical background provides the intellectual context within which the emergence of the contemporary piracy scene must, I argue, be viewed.

The second part of this chapter gives an overview of the digital archive on which this book draws, the Defacto2 archive of scene NFOs (iNFO files; see Figure 1, below). This archive consists of roughly 12gb of documents that include plain-text files of inter-group spats, release announcements, courier statistics, but also demoscene executables, graphics, and other communications. The initial approach to the archive here is to document the different types of output and to show the reader what can be gleaned from insights within each type of file; a taxonomy of the digital-cultural materialism that I here deploy.

The third and final part of this chapter gives an outline of the remainder of the book, but also specifies the textual approach to the source material. With over 35,000 unique digital artefacts, the methodology of this book employs both traditional/close and computational/distant readings of the space. Much of the structural analyses of the Scene and its make-up come from a detailed, in-depth approach to these polymorphic texts, while in Part Two of the book I turn to the ways in which network analysis, named-entity recognition, and other stylometric profilings can help us to understand this underground culture.

Figure 1: A typical Razor1911 NFO.

Part II: Structure

Chapter 2: The Scene [8,000 words]

This chapter provides a high-level overview of the organisation of the shady underground piracy network known as the “warez scene”. It provides the reader with an overview of technical terms, organisational principles, and a high-level understanding of the mechanics of the Scene’s setup.

Technical concepts here covered include FTP, FXP, topsites and dumpsites, release groups and suppliers, releases and “pres”, nukers and nuking, IRC, eggdrop bots, glftpd, racing, cracks, the demoscene, camming and telesyncs, couriers, FXP boards, NFO files, and P2P. Many of these areas are given fuller treatments in the chapters devoted to them hereafter but the intention of this section is to give the reader a broad orientation with respect to the interrelation of parts and the whole.

While, by necessity, this chapter has to handle a wide range of technical matters in detail, I attempt here to outline these concepts in ways that are comprehensible to those without a technical background. To assist with this, a comprehensive glossary is provided at the end of the volume that provides a cross-reference back to the more detailed treatment in this chapter.

By the close of this chapter, the reader should have a good idea of the organisational principles behind the Warez scene, the hierarchies and chains of command, idea of competitiveness and political economy within the space, and a high-level background to many of the technological aspects of its implementation.

Chapter 3: Release Groups [10,000 words]

At the heart of the Scene are release groups; well-organised collectives of individuals – often only known to each other by anonymous online handles – that specialise in the high-speed acquisition, processing, and dissemination of copyrighted material. The precise activities conducted by each group varies by the media form in which they specialise. Software and game pirates, for instance, will likely have a period of "cracking" in which they must race to strip off the digital rights management technologies on the item in question, battling for speed against other groups who likely have a similar supply chain.

In addition to chronicling some of the most prominent release groups – Razor 1911, FAiRLiGHT, Centropy, RELOADED, and DEViANCE – this chapter draws on the archive to demonstrate the systems of priority and prestige that run through the behavioural and linguistic codes of these groups. Noting that, often, the "losing" group can be just minutes behind their rivals in the supply → cracking → distribution model, the Scene nonetheless enforces a strict system of both speed checking and quality control in which subsequent or non-working releases are marked as "nuked", with a penalty levied on the group in question. The actual merit of such a system for the end-goal of releasing pirate material for consumption is, I would argue, negligible since the releases can often be mere minutes apart. Instead, these groups primarily compete within a symbolic economy of prestige and credit, as has often been charted among subcultural groups.

Chapter 4: Couriers [12,000 words]

If release groups specialise in the acquisition and packaging of content, a class of users further down the food-chain are primarily responsible for distribution: couriers. These users have access neither to gold disc supply chains nor to the intricate knowledge of disassembly that is necessary for cracking. Instead, they specialise in having access to a wide variety of topsites and in building software that will allow them to shift releases from site to site, faster than their rivals, thereby earning credits on those sites.

As with release groups, couriers are rigorously organised into hierarchies of competition based on speed and quality. Weekly charts are produced that award points to couriers based on the number of bytes they successfully upload to sites, with each site awarded a rank multiplier (see figures 2 and 3 below) for scoring purposes. Courier groups recruit individual couriers who score well in such charts in order to gain access permissions to topsites based on a reputational economy.

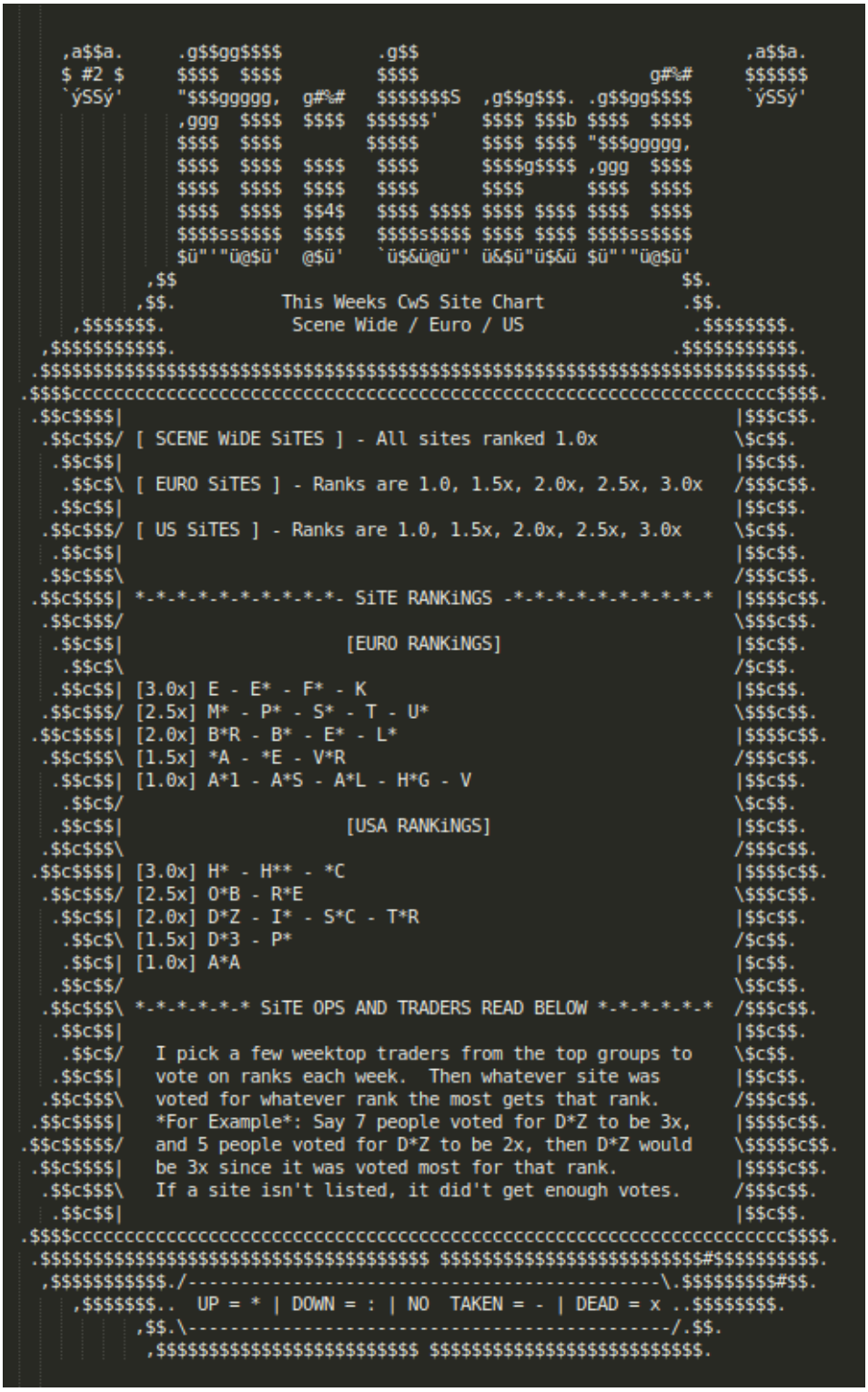
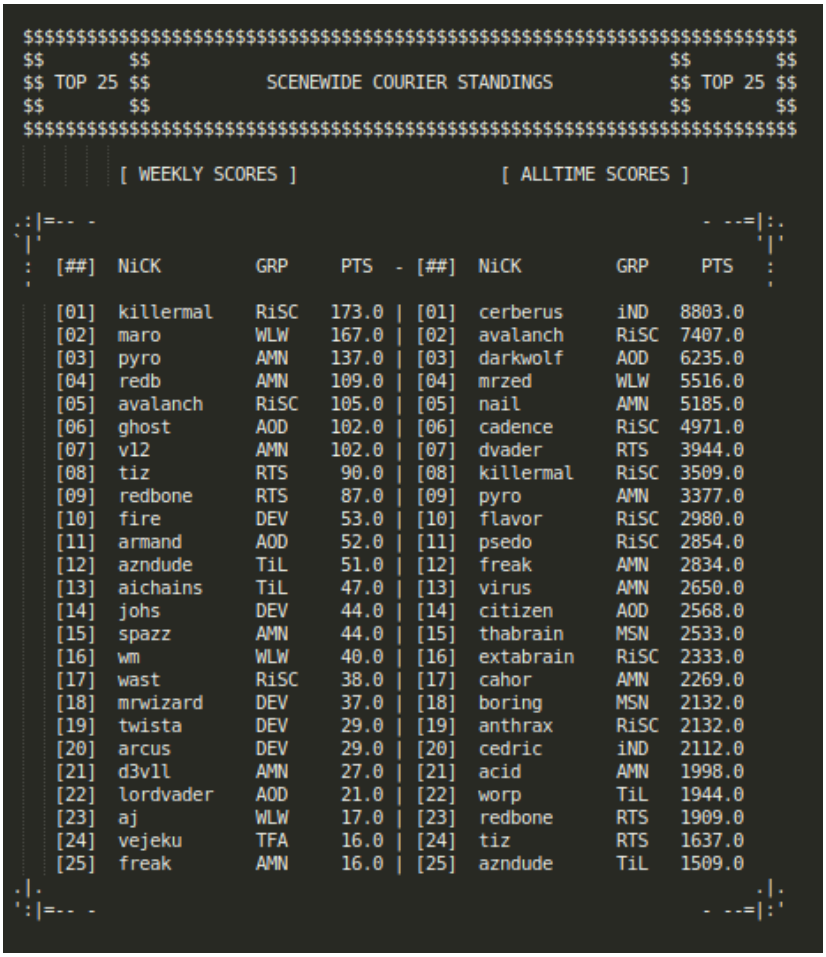


Figure 2: Site ranking scorecard #180 for 11/25/01 - 12/01/01



```
$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$
$$        $$                                                        $$        $$
$$ TOP 25 $$          SCENEWIDE COURIER STANDINGS                   $$ TOP 25 $$
$$        $$                                                        $$        $$
$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$

        [ WEEKLY SCORES ]                    [ ALLTIME SCORES ]

.:|=-- -                                                      - --=|:.
`|'                                                            '|'
 :  [##]  NiCK         GRP     PTS  - [##]  NiCK         GRP     PTS   :
 '                                                                    '
    [01]  killermal    RiSC   173.0 | [01]  cerberus     iND   8803.0
    [02]  maro         WLW    167.0 | [02]  avalanch     RiSC  7407.0
    [03]  pyro         AMN    137.0 | [03]  darkwolf     AOD   6235.0
    [04]  redb         AMN    109.0 | [04]  mrzed        WLW   5516.0
    [05]  avalanch     RiSC   105.0 | [05]  nail         AMN   5185.0
    [06]  ghost        AOD    102.0 | [06]  cadence      RiSC  4971.0
    [07]  v12          AMN    102.0 | [07]  dvader       RTS   3944.0
    [08]  tiz          RTS     90.0 | [08]  killermal    RiSC  3509.0
    [09]  redbone      RTS     87.0 | [09]  pyro         AMN   3377.0
    [10]  fire         DEV     53.0 | [10]  flavor       RiSC  2980.0
    [11]  armand       AOD     52.0 | [11]  psedo        RiSC  2854.0
    [12]  azndude      TiL     51.0 | [12]  freak        AMN   2834.0
    [13]  aichains     TiL     47.0 | [13]  virus        AMN   2650.0
    [14]  johs         DEV     44.0 | [14]  citizen      AOD   2568.0
    [15]  spazz        AMN     44.0 | [15]  thabrain     MSN   2533.0
    [16]  wm           WLW     40.0 | [16]  extabrain    RiSC  2333.0
    [17]  wast         RiSC    38.0 | [17]  cahor        AMN   2269.0
    [18]  mrwizard     DEV     37.0 | [18]  boring       MSN   2132.0
    [19]  twista       DEV     29.0 | [19]  anthrax      RiSC  2132.0
    [20]  arcus        DEV     29.0 | [20]  cedric       iND   2112.0
    [21]  d3v1l        AMN     27.0 | [21]  acid         AMN   1998.0
    [22]  lordvader    AOD     21.0 | [22]  worp         TiL   1944.0
    [23]  aj           WLW     17.0 | [23]  redbone      RTS   1909.0
    [24]  vejeku       TFA     16.0 | [24]  tiz          RTS   1637.0
    [25]  freak        AMN     16.0 | [25]  azndude      TiL   1509.0
.|.                                                            .|.
':|=-- -                                                  - --=|:'
```

Figure 3: Courier weektop scorecard #180 for 11/25/01 – 12/01/01

Courier groups are also responsible for the production of sophisticated trading software that leverages arbitrary IP endpoints in the FTP protocol's PASV and PORT commands to ensure maximum transfer speeds between servers. Such software uses a range of manual and automatic approaches to identify releases and their suitability to transfer according to the site rules on each topsite, primarily through regular expressions in the release titles and metadata displayed in each site's IRC channel. Such practices of "autotrading" are frowned upon and may even earn a user a ban if caught, since the procedures used for identifying releases are imperfect and may result in nukes.

There is, though, a more interesting parallel between courier autotrading and another area of contemporary life: stock trading. Both of these demographics seek the highest speed of transmission possible, situating bounce servers next to topsites and the central stock exchange respectively. Both have also developed automatic systems that operate at speeds greater than those possible for humans in order to maximise their personal return. The end of this chapter analyses this societal parallel.

Chapter 5: Topsites [10,000 words]

At the heart of the scene lies the physical and digital infrastructures that provide storage, network access, and archival functions: topsites. These sites, usually running on the GNU-Linux operating system and using an FTP daemon called glftpd (Greyline FTP Daemon), are hidden behind traffic bouncer proxies that cloak the IP address of the original site, thereby obfuscating the true location. Frequently located on high-speed lines in universities, however, there have also been a significant number of successful law enforcement raids against topsites, some of which are chronicled in this chapter.

The primary goal of this chapter is to orient the reader with respect to the hardware, software, and geographical positioning of topsites, the core endpoints of the warez scene. This includes considerations around intercontinental routing and the material underpinnings of geographic divisions of the scene. For instance, it is notable that courier charts are often divided into European, continental US, and Asian sections, but the rationale for this hinges upon the sub-oceanic cabling and the ability of the TCP/IP stack to efficiently route between these areas. In this way, this chapter not only details the technical infrastructures, but also examines the social considerations in the planning of the internet and the ways in which short-burst high-volume transfers remain geographically determined.

Chapter 6: Software [12,000 words]

While various software elements of the scene have been touched upon in preceding chapters, the aim in this final descriptive chapter is to outline the different categories of software used by these underground networks, including: IRC servers, IRC bouncers (psyBNC), IRC bots (eggdrop), traffic bouncers (ebBNC, ebBNC2, f-ftpbnc, simpleentry, SSL-lbnc, yatb), FTP/FXP clients (pftp, FTPRush, slftp), FTP servers (glftpd, Raiden), autotrading clients (mFtp), pre-bots, nukenets, and other software elements.

Although much of this chapter is devoted to outlining the core elements of the scene's software stack, the analytical portion of this chapter focuses upon the vast quantities of technical labour that are invested in the production of toolsets for the illicit dissemination of pirate material. Indeed, the argument of this chapter is that the labour time invested in the scene – in its software production and maintenance, in the work of system administration, and in the efforts invested in cracking, releasing, and couriering – substantially outweighs the economic benefit to any single actor in the acquisition of software, films, and music. While, then, well-known industry videos against piracy highlight the economic damage done to actors, the commensurate economic *gain* to those working in the scene, if they chose to invest their labour elsewhere, is tiny.

Part II: Signs

Chapter 7: NFO Aesthetics [12,000 words]

The DeFacto2 archive is not only a testifying record of the existence of an otherwise hidden subcultural community, but also an aesthetic log at the intersection of several digital art spaces,

including demoscene designers and ASCII art creators. In this chapter, I explore the aesthetics of this scene archive, noting the ways in which cultural kudos is rewarded in the piracy scene for those with design abilities. Thus, there is even an economy of aesthetics that operates within this space, rewarding those with specific design abilities with access to vast archives of pirated material.

The analysis in this chapter also touches upon gender and the scene. For much of the artwork presented in NFO files appears to cater for the male gaze, so aptly first identified in cinema by Laura Mulvey. While a full demographic breakdown of scene participants is near-impossible to assess, an analysis of sexualised content in NFO files and demoscene contents indicates that a primary group to whom this area caters is male.

Further, this chapter conducts a range of stylometric profiling exercises upon the archive, analysing the writing style, homogeneity or otherwise, and other stylistic features of the texts therein. Well-rehearsed profiling techniques, such as Burrows's delta method, will also be used to attempt to cross-correlate artefacts for shared authorship, thereby opening a path to understanding the ways in which groups overlap with one another in their memberships.

Finally, this chapter will broach the reasons for the emergence of specific scene aesthetics, including the importance of group distinctiveness, the establishment of aesthetic tradition, and the ability to demonstrate a "spending power" in the "hiring" of artists in the service of specific groups.

Chapter 8: Dissemination Networks [8,000 words]

This chapter explores the co-occurrences of specific terms within the archive, plotting a network map of actors and terms that allows for a cartography of the archive to emerge. In this chapter, I examine the ways in which different groups and individuals interact with one another, including positive and negative sentiment alignment in each case. This allows for an emergent cartography of this otherwise vast and unmanageable archival space to be presented.

Chapter 9: Points of Emergence [8,000 words]

This penultimate chapter examines the implications of the surfacing of this archive into the public space. Since many statements within the archive itself pertain to the ostracization of leakers and the need for information security – a frequent denunciation is that topsites are selling access, leading to the admission of insecure/unvetted individuals – it is notable that this material is now public.

In this chapter I examine the cultures of secrecy that surround such underground cultures and explore the politics of curating an archive of underground, illicit activities, even when cloaked by pseudonyms. It is here that I give the final, most thorough consideration to what it means to examine a digital "archive" and to think through the politics and ethics of such digital spaces.

Part III: Conclusion

Chapter 10: Conclusion [8,000 words]

This chapter concludes the volume.

Delivery Date

The manuscript can be ready by the 1st February 2023.

Figures and Tables

The manuscript will have approximately 20 figures that can be customized for black-and-white printing. There are also several tables.

Fit with Existing Literature

There are a number of existing journal articles on the underground piracy scene, but none that seek to document so thoroughly the archive and its material that I cover here. There is only one other book that works on this subject and it is not a rigorous academic study: Craig, Paul, *Software Piracy Exposed* (Rockland, MA: Syngress, 2005). Further, its sourcings are not apparent, even where the information contained within appears accurate. This book, therefore, affords a significant and important advance on existing work, both in terms of its expository content and its focus on a particular archive of the space.

Intended Audience

The intended audiences for this book are:

- Those interested in studying illicit underground online subcultures
- Historians of technology
- Digital archivists
- Digital humanities scholars
- Students of any of the above