

Understanding scanlation: how to read one million fan-translated manga pages

Dr. Jeremy Douglass, Post-doctoral Researcher, Software Studies Initiative.

William Huber, PhD student, Visual Arts Department, University of California, San Diego.

Dr. Lev Manovich, Professor, Visual Arts Department, University of California, San Diego (UCSD) and Director, Software Studies Initiative (softwarestudies.com) at California Institute for Telecommunication and Information (Calit2).

This article is the first in series of publications about our “One million manga pages” project. As new articles become available, they will be listed on the project homepage:

<http://tinyurl.com/one-million-manga>. All images in this paper are linked to high-resolution versions. The full image set is available online: <http://tinyurl.com/manga-articles>.



Exploring one million manga pages on the 287 megapixel HIPerSpace (The Highly Interactive Parallelized Display Space) at Calit2, San Diego. HIPerSpace offers 35,840 x 8,000 pixels resolution (287 megapixels) on a 31.8 feet wide and 7.5 feet tall display wall made from 70 monitors. This photo: Jeremy Douglass (Post-doctoral researcher, Software Studies Initiative) and Florian Wiencek (Jacobs-University, Bremen) zooming into a manga visualization.

INTRODUCTION

Digitization efforts by many museums, libraries and other institutions, and the massive growth of both user-generated and professional digital content open up new possibilities for the study of media and visual cultures. To explore these possibilities, in 2007 we set up the Software Studies Initiative at University of California, San Diego. The lab uses digital image analysis and interactive visualization to explore patterns in large sets of images, video, and interactive media.

Our work is closely aligned to the vision of digital humanities put forward by [Office of Digital Humanities](#) at the National Endowment of Humanities (the U.S. federal agency which funds humanities research). The joint NEH/NSF [Digging into Data competition](#) (2009) description opens with these questions: “How does the notion of scale affect humanities and social science research? Now that scholars have access to huge repositories of digitized data—far more than they could read in a lifetime—what does that mean for research?” The same questions guide our work.

Our most ambitious project to date is exploration of an image set of **one million manga pages**. These pages correspond to **883 manga series** that were available as “scanlations” (manga digitized and translated by fans) on OneManga.com in the fall 2009. At that time we downloaded all pages available on the site, along with the user-assigned tags indicating the genres and intended audiences for the series. We then analyzed visual features of these pages using custom digital image analysis software that we developed and then ran on supercomputers at the National Energy Research Scientific Computing Center (NERSC). (This part of the project was funded by a Humanities High-Performance Computing grant from the NEH Office of Digital Humanities.)

Humanities projects traditionally consider tens, hundreds, or at most thousands of images. Our manga set is several orders of magnitude larger. In order to explore an image set of this size, we use next-generation supervisualization systems developed at the California Institute for Telecommunications and Information Technology (Calit2) where our lab is located. The photo that opens this article shows a visualization of our complete one million pages set on 287 megapixel HIPerSpace video wall system. The manga project has pushed us to keep developing

visualization tools that can take advantage of such next generation displays, while providing a test case of their use.

This article is the first in a series of forthcoming publications which present our methodology, research findings and theoretical discussions of the larger issues raised by the computational analysis of large cultural data sets as exemplified by our one million manga pages set. The future texts will discuss the variations in graphical languages between manga aimed at different audiences; temporal changes in graphical languages of manga series during the duration of their publications; fans understanding of manga genres; and other topics.

While this and other papers that report on our manga research focus on manga images, the methods we describe are equally useful for researching many other types of visual media. In our lab, we have already applied them to various media related to manga - comics, webcomics, cartoons, and video games - as well as other media such as motion graphics, print magazines, paintings, and photographs. Examples of the data sets we analyzed include 100 hours of Kingdom Hearts video game play, 4535 Time magazine covers, and 20,000 pages of *Science* and *Popular Science* magazines. (For details, see the [Research section](#) of our lab blog softwarestudies.com; you can also find our visualizations on [Flickr](#) and [YouTube](#)).

There are a few reasons why we chose manga for our first large-scale study. First, manga is one of the most popular cultural forms around the world. Within manga, we find a range of graphical techniques – so whatever we learn from manga could also help us in researching other graphical forms, be they drawings, engravings, vector graphics, animations, etc. Second, manga is both inside and outside of the “culture industry.” Although there are plenty of independent artists publishing their own manga books or distributing their creations online, the mainstream manga read by millions of people around the world is packaged and distributed by large companies. However, while commercial movies and video games may be created by teams of thousands, manga series are typically created by single artists or artist-writer teams with the support of a few assistants. Finally, the last reason is practical: the distributed collaborations of manga fans resulted in the online availability of an extremely large digitized sample of manga culture.

Like most other cultural forms today, “manga” exists in multiple forms – original Japanese print publications (which themselves exist in several versions), official translations sold in other

countries, unofficial (but highly popular) translations produced by fans and distributed on the web, and others. The first step in a systematic computational study of manga is a clear understanding of the differences between these forms.

This article explains how we use digital image analysis and visualization to approach this question. While manual examination of samples from different manga forms certainly already reveals some of the differences between them, a computational analysis of a one million pages set allows us to understand these differences with much more precision; it also makes visible a number of new important patterns which would be impossible to notice otherwise.

Our discussion focuses on perhaps the most important version of manga - unofficial translations produced by fans which they make available online. These scanned and translated manga are called “[scanlations](#).” As detailed in the analysis below, the creative activity of scanlation groups is neither “authorship” nor “remix.” It also cannot be adequately described using a well-known distinction by Michel de Certeau between “strategies” and “tactics” (because in contrast to the unconscious tactics described by de Certeau, scanlation groups add new pages to manga series they publish quite consciously.) Similarly, scanlations are neither “remediations” (Jay David Bolter and Richard Grusin) nor “transmedia” (Henry Jenkins). In short, we currently lack proper terms to describe them - and this is already an important reason why we should study them.

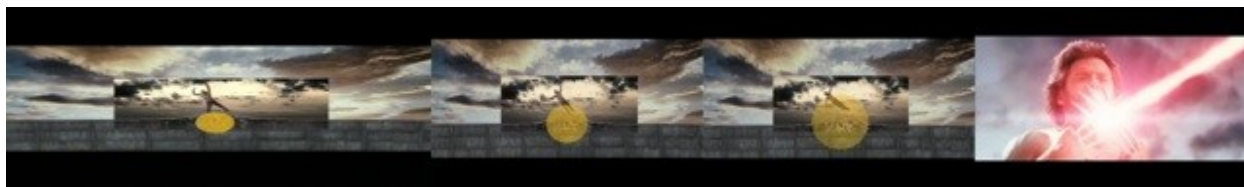
Another reason is their popularity. Given the statistics for visit numbers to the web sites that distribute scanlations around the world, it is likely that more people outside of Japan currently encounter manga in this form, rather than through the official translations. Given this popularity, an understanding of the specificity of scanlations - in particular, various creative additions the fans make to the manga series they republish - is quite important. To date, most writing has considered scanlation either as a subculture or as a market – that is, as a primarily sociological or economic phenomenon. Examples include scholarly articles, (Brienza, Lee, Leavitt, Nowlin, Rampant), newspaper and magazine features (Deppey, Macias, Yang), and online publications (Doria). This focus is vital to understanding scanlation, however it has often entailed a preoccupation with the legal and ethical discourses of market forces and intellectual property law: “Does scanlation help or hurt the manga publishing economy?” “When is scanlation legal/illegal?” “Should scanlation exist?” and so forth.

This focus on legal issues can preclude asking other, equally important questions: What *is* scanlation? More specifically, what is scanlation as a cultural form? What are the artifacts created by scanlators, how are they experienced by their readers, and how do they differ from the artifacts officially published manga? These questions may lead us towards a greater understanding of manga itself, not as a formally defined medium, or as an industry, but as an expansive set of cultural practices and artifacts, both official and unofficial.

However, the importance of such analysis goes beyond manga. Today, hundreds of millions of people encounter commercial media products in various unofficial versions, available via a variety of channels: bootleg recording, peer-to-peer sharing, YouTube reposting, [fansubbing](#), etc. (For example, in 2009 one of the most common peer-to-peer file sharing protocols used to distribute such unofficial versions, BitTorrent, accounted for [27 to 55 percent of all Internet traffic](#) in countries around the world.) If we are interested in studying contemporary media as it is actually experienced by audiences, we should take these unofficial versions as seriously as the original media products released by media companies. Recent scholarly considerations of bootleg audio (Marshall) and bootleg video (Hildebrand) have helped expand our understanding of music and film cultures, particularly the aesthetics and politics rooted in affordances, limitations and artifacts of the cassette tape and the video tape. In scanlation we find practices based originally in paper, but increasingly marked by complex networks of digital production, transformation, and distribution.

Real Proper by artist Artie Vierkant perfectly illustrates this circulation of media in the early 21st century. This two-hour long video consists of four different versions of the same commercial film. The artist explains:

Each version stems from a single release, a leaked unfinished copy of the film *X-Men Origins: Wolverine*, illegally uploaded to the Internet in 2009. The "Wolverine Workprint" spawned countless copies, re-encoded and re-distributed by many users to fit different formatting conventions, including the addition of subtitles, altered aspect ratios, altered running times, etc. 38 unique versions of the workprint still exist on various file-sharing platforms more than a year later, despite the availability of "true" copies released to file-sharing networks after the availability of the completed film and officially licensed DVDs. (Artie Vierkant, email to Lev Manovich, 12/10/2010.)



Artie Vierkant, [Real Proper](#). 2010. A frame from the HD video (1 hr 58 min) *Space* at Calit2, San Diego. HPerSpace offers 35,840 x 8,000 pixels resolution (287 megapixels) on a 31.8 feet wide and 7.5 feet tall display wall made from 70 monitors. This photo: Jeremy Douglass (Post-doctoral researcher, Software Studies Initiative) and Florian Wiencek (Jacobs-University, Bremen) zooming into a manga visualization.

The artist named his video “Real Proper” after the tag “proper” used to describe some of the movie files on BitTorrent. In the text accompanying the exhibition that included this video, he quotes the text by Abhishek Kunal, which describes this tag as follows:

Due to scene rules, whoever releases the first Telesync has won that race (for example). But if the quality of that release is fairly poor, if another group has another telesync (or the same source in higher quality) then the tag PROPER is added to the folder to avoid being duped. PROPER is the most subjective tag in the scene, and a lot of people will argue whether the PROPER is better than the original release.” (Abhishek Kunal, [All About Movie in torrent Tags](#),” uploaded to scribd.com 09/12/2010.)

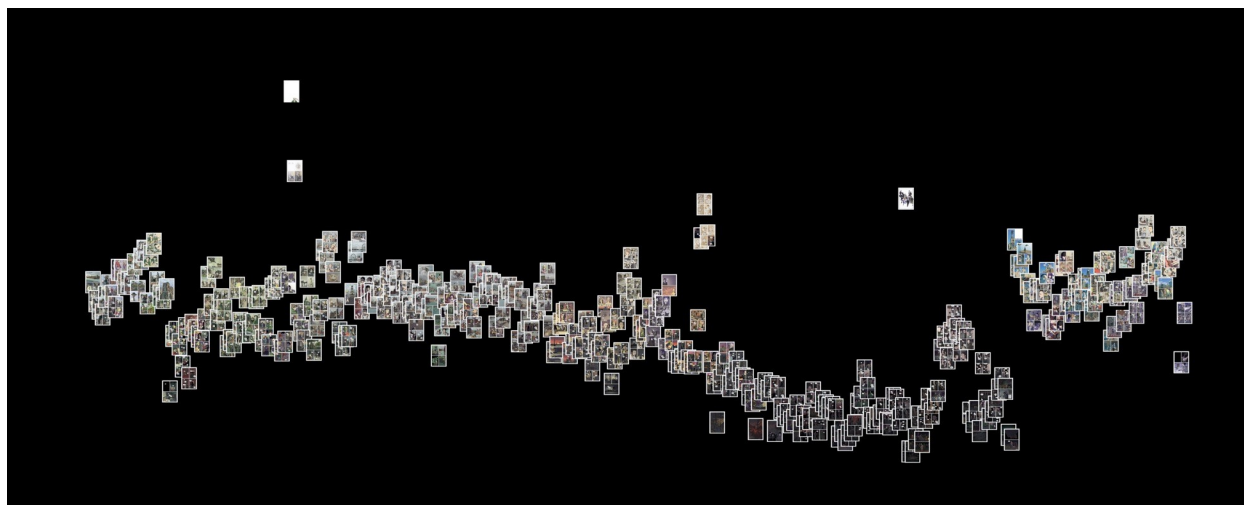
It is relevant to note that Kunal’s text defines a vocabulary of 30 different tags. This number alone hints at the richness of the BitTorrent scene for unofficial film versions. Here are few additional examples of these tags. *CAM* is “a theater rip usually done with a digital video camera.” *Telesync* indicates “the same spec as CAM except it uses an external audio source (most likely an audio jack in the chair for hard of hearing people)” *Workprint* is “a copy of the film that has not been finished.” The tag *Repack* is used when “a group releases a bad rip, they will release a Repack which will fix the problems.” These examples illustrate that cultures of unofficial versions are not a simple “copy and paste,” but instead have their own social dynamics, their own methods of production, and their own forms of creativity that - as this article will show using the example of scanlation - go beyond the simple competition to release copies of commercial media files before they become officially available.

METHOD: MEDIA VISUALIZATION AND DIGITAL IMAGE ANALYSIS

Scholarly work with images is not unique to manga. Most humanities fields use images. Art historians employ photographs of art and architecture. Film and television scholars deal with film and video stills. Historians work with digitized manuscripts, newspapers, and magazines. As more visual collections are digitized and made available online, the image becomes even more central for humanities work than it already was in the age of slides, photo prints, and photo copies.

Contemporary software applications for media viewing and cataloging allow users to browse through and search image collections, and display image sets in an automatic slide show or a PowerPoint-style presentation format. However, they are much less useful as serious research tools. Desktop applications such as iPhoto, Picasa, Adobe Bridge, Aperture, and image sharing sites such as Flickr can only show multiple images in a few fixed arrangements - typically a two-dimensional grid, a linear strip, or, in some cases, a map view. These views are quite limited - a user cannot cluster images by their visual properties or by semantic relationships, create animated timelines, compare collections of tens of thousands of images each, or use other techniques to find interesting patterns.

Since the fall of 2008, our lab has been developing new visualization tools that offer these possibilities. This software combines the functions of media viewing applications and graphing and visualization applications. Like media viewing applications, the software can display lists or grids of images. These may be arranged by a wide variety of metadata types, including information that came with the images (such as dates and authors), or added by researchers (such as annotations) or visual features that are mathematically extracted from the images, such as color, contrast, texture, and composition. (This process is called “digital image analysis.”) Like graphing applications, the software creates graphs to show relationships and patterns in a data set. However, if applications such as Excel and Google Charts can only display data as points, lines, spheres or other graphic primitives, our software displays all the images in the collection superimposed on the graph. We call such visualizations “media visualizations.”



Media visualization of the webcomic *Freakangels* by Warren Ellis and Paul Duffield, showing the first 57 episodes (342 pages) published from February 2008 to June 2009. X axis: page publication order. Y axis: overall brightness of each page (mean of all pixels' grayscale values in each page).

While similar techniques have been previously described in a number of computer science papers, they have not yet been offered in an application that allows individual researchers and research groups in the humanities to work with their own image collections. Accordingly, our goal has been to create tools that are easy for both ourselves and others to use because they work with common data formats: sets of image files stored in a directory on a computer or a network, and text files containing information about these images.

Our visualization tools run on several platforms. One version runs on the 70-monitor HIPerSpace tile display, and allows for interactive real-time visualization of large data sets at a resolution of 287-megapixels (a photo of this display appears at the beginning of the article). Another version is a cross-platform Adobe AIR viewing application that allows interactive exploration of smaller data sets on a personal computer. A third version, also for any personal computer, is a macro for an open-source application image analysis application ImageJ used for creating high-resolution still and animated visualizations. Most visualizations in this article were produced using this third tool.

In a typical project, we use digital image analysis to measure various image features of each image in a set; the results of this analysis are combined with existing metadata and annotations;

then we use our visualization tools to explore the image set. In order to work with very large sets of images that may come in a variety of formats, we developed a batch system that runs all images through a series of image analysis routines and records the results in text format. We implemented versions of this system have been implemented to runs on personal computers, and are variously based on Bash/ImageMagick, Python/Python Imaging Library, and Java/MATLAB. In addition, the MATLAB version of the workflow was developed into a version used to process very large collections of images on supercomputers at the National Energy Research Scientific Computing Center (NERSC). Most measurement data used in exploring our collection of one million images discussed here was produced at NERSC.

This article uses a few different types of media visualization. In one type, manga pages are sorted one-dimensionally by some metadata, but then wrapped into a 2D image, much like a line of text is wrapped into a paragraph on the page. We refer to these as “montages.” Another type compares and contrasts, in which two versions of a single page are overlaid, and only areas of differences are shown. We refer to these as “comparisons.” The final media visualization type uses different visual properties to position a set of images on the X-axis and Y-axis of a graph. We call these visualizations “image plots.” With all of these visualizations, the goal is create unique new views onto the subject matter, enabling researchers to understand and discuss manga in ways that would otherwise be difficult or impossible.

DATA: ONE MILLION SCANLATED MANGA PAGES

The word “manga” is Japanese in its origin, but has become international in its scope. The word, written as 漫画, can be translated as “whimsical pictures.” 画 (ga) does not connote in itself either gravity or frivolity, but refers to a painterly or illustrative depiction. “Man-” (漫) connotes things that are light-hearted, comic, humorous, and idle. Thus, manga is in its origins connotes triviality, the comic, the demotic — and thus seems like a straightforward counterpart to the “comic book” or “the comics,” “the funny pages,” etc.

Despite this similar etymology, manga in Japan surpass their Western counterparts in both breadth of matter and scope of cultural influence. Whereas comic books in the west are traditionally read by children (and generally boys), with variants such as “graphic novels” and “la bande dessinée” (BD) read by a devoted yet relatively small readership, manga in Japan

address a large portion of the general public. Examples of the breadth of manga publications include political tracts, such as those by Yoshinori Kobayashi, and economic history, such as *Nihon keizai nyumon* (translated as “Japan, Inc.”). However, it is those publications directed at young boys - shounen - that have captured the most attention from American and other western audiences.



Left image: sample page from *Vampire Knight*. Right image: sample page from *One Piece*. The relative popularity of titles may differ between original print markets, unofficial fan translations, and official translation markets. The OneManga.com June 2010 list of top 50 series titles lists *One Piece* as no. 2 and *Vampire Knight* as no. 13, while an ICv2 survey of English manga imports during 2010 Q3 lists *One Piece* as no. 2 and *Vampire Knight* as no. 4.

Manga is one of the three popular Japanese cultural forms that have developed a sizable global market. Unlike the other two forms, anime and videogames, manga had already been part of regional circulations in East Asia. These circulations took different forms. In Korea and China,

regional comic-making practices were understood as part of a general practice of manga-making. In the United States, manga was treated as a peculiarly Japanese variation of comic style: just as anime was termed “Japanese animation,” manga were called “Japanese comics.” Recently, however, some Western writers have come to use the term “manga” as a category that includes all comics.

Attempts to publish and sell translations of manga in the United States began in the 1980s (Schodt 1996); by the mid 1990s, most translations were produced and distributed by 3 major US firms working on coordination with Japanese publishers, which partially owned them. Although some of the first titles to achieve success in the United States were geki-ga (literary manga) such as *Lone Wolf and Cub*, written for an adult readership, it was publications for the young male market that would be most aggressively distributed. Today, only a small percentage of Japanese titles are officially translated into English, mostly shounen titles marketed to young boys. Even these titles may see delays of months or years between the original and translation release (although this delay has decreased recently).



Comic-Con Internal 2010 showroom floor. Manga is part of a transmedia consumer culture, with print versions created and sold alongside related anime, videogames, toys, and other products.

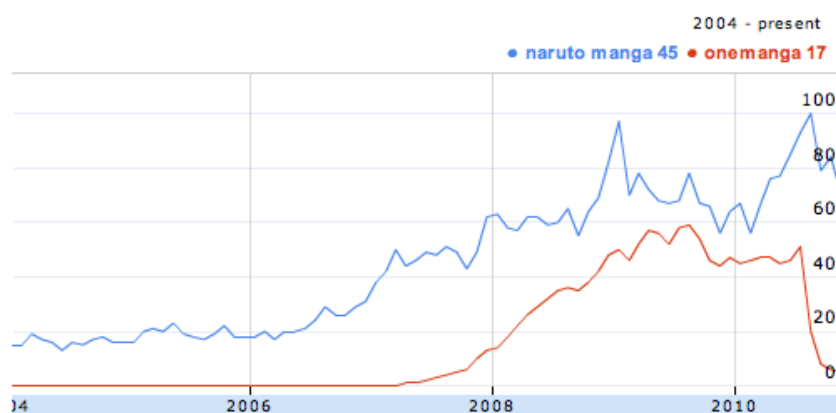
The slow pace of translation and the enthusiasm on the part of English language readers led to the emergence of fans and fan communities who translated and re-distributed the original works themselves, an activity the participants call “scanlation” (a portmanteau of “*scan*” and “*translation*”). Scanlation is the practice of scanning original Japanese editions of manga, translating the text into another language, then using image-editing software to replace the Japanese text and other textual image elements on the page with the translation text. (A related practice and subculture of “fansubs” exists for producing subtitled versions of anime videos as well as other TV programs and feature films.) Although we don’t have exact statistics on the number of active scanlation groups and their membership around the world, the bits of information available on many scanlation web sites and forums give indications of the scale of scanlation culture. (For example, the scanlator site mangahelpers.com lists 107,075 members).

The redistribution of scanlated manga takes place without explicit approval or material support from the original manga publishers, although much of it has been produced and distributed under a set of expectations and norms intended to minimize conflict, particularly intellectual property law litigation. In what has been called the “gentleman’s agreement,” most scanlation sites agree to promptly remove any material at the request of the original publisher, while most publishers have until recently limited the majority of their requests to occasions when official translations are to be released into non-Japanese markets.

Scanlated manga circulate in multiple ways online, with a core part of the subculture working, socializing, and distributing their work through IRC (Internet Relay Chat) channels. The public face of scanlation for a much larger online audience, however, is a group of websites that fans call “online viewers” and that publishers call “aggregators.” OneManga.com was, until July 2010, the largest and most active of these scanlation sites, distributing the work of thousands of scanlation groups for free download and online reading. (The site was rated no. 300 in the U.S. in terms of number of visitors; in Indonesia, Malaysia and Singapore it was in the top 20 sites.) However, recent trends in the political economics of Japanese popular culture have resulted in a change in the previously tolerant attitude from manga publishers. At the request of a newly formed anti-scanlation coalition of prominent Japanese and U.S. publishers, OneManga.com removed all pages from its site in the summer of 2010 (Cha, Eastman, Reid). These events were remarkable for the speed of the transition (from formation of consortium, to notice, to

takedown), and it remains to be seen whether they indicate a temporary or permanent shift in the cultural landscape. In either case, as with historical studies of online discussion groups, social networks, and file sharing practices that focus on key moments in time, a snapshot of OneManga.com from late 2009 captures a unique moment in global media culture.

Interest Over Time



Search volumes on Google for “naruto manga” and “onemanga” (2004-2010) plotted over time. The popularity of both search terms grows in parallel since 2006; however, after OneManga.com stops posting page images for online reading in reaction to copyright concerns, its search volume quickly drops. (Data and graph: Google Insights for Search, November 20, 2010.)

1. Jakarta (Indonesia)	100	1. Malaysia	100
2. Kuala Lumpur (Malaysia)	79	2. Indonesia	93
3. Singapore (Singapore)	41	3. Philippines	56
4. Toronto (Canada)	17	4. Singapore	40
5. Los Angeles (United States)	13	5. Trinidad and Tobago	29
6. Lima (Peru)	12	6. Jamaica	28
7. Mexico City (Mexico)	11	7. Myanmar	25
8. Sydney (Australia)	10	8. Mauritius	18
9. Montreal (Canada)	10	9. Canada	13
10. Paris (France)	10	10. Peru	12

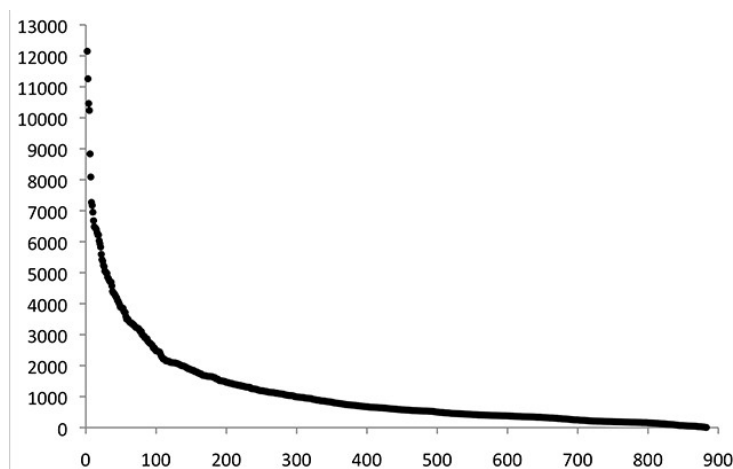
The global audiences for scanlations: cities (left) and countries (right) with the largest Google search volumes for “one manga” (2004 - 2010). (Data and graph: Google Insights for Search, December 9, 2010.)

Scanlation is not simply the translation of printed manga: it is the creation of new digital objects that have their own unique affordances. Printed manga are released episodically, most often serialized in weekly publications and later collected into small paperbacks. Scanlated titles are aggregated and are often available in their entirety: one can read a long serialized narrative continuously by viewing an entire title's volumes at once. Most importantly, scanlations are digital images: they can be emailed, attached, modified, and re-shared, all with little difficulty. They can also be processed and studied with tools designed for the analysis of digital images.

In the fall of 2009, we used web spiders to crawl OneManga.com, downloading 883 Manga series containing 1,074,790 unique scanlated pages. We then used our custom software system to analyze visual features of every page (range of gray tones, presence of texture and fine details, etc.) It is important to point out that this sample of mostly Japanese manga digitized and translated by fans has a different distribution of types than the proportions found in the market in Japan, where manga publishers categorize their offerings demographically by age and gender: the four main types are titles aimed at boys, girls, young men, and young women (shounen, shoujo, seinen, and josei manga). Instead, similar to the availability of officially translated manga, the OneManga.com data set has proportionally fewer titles for girls; the relative number of titles for the older groups is even smaller. This difference between the Japanese national and international English-language readership is also reflected in OneManga.com visitor demographics: the majority of readers are male in the 18-24 age group. (It is possible that this group actually has a younger lower bound – Alexa.com, which we used for this statistic, does not provide public demographic data for younger Internet users).

Audience segment	shounen (boys)	shoujo (girls)	seinen (young men)	josei (young women)
Number of series	375	235	183	18

Number of series for different audiences in the OneManga.com data set. Identification of a series as belonging to a particular audience segment comes from the tags assigned by OneManga.com editors. (Note that the tagging system used by the site only allows for a single primary audience per title.)



Graph plotting the series in the OneManga.com data set according to number of pages per title. Y axis: number of pages in a series. X axis: number of series with particular page count. Out of 883 series, 297 series contain over 1000 pages. The series with greatest number of pages in our data set: *Hajime no Ippo* by Morikawa Jyoji (15979 pages - not shown in the graph); *Case Closed* by Aoyama Gosho (12148 pages); *Jojo's Bizarre Adventure* by Araki Hirohiko (11259 pages).

MANGA SERIES		AUDIENCE	COUNTS		STATISTICS		TAG METADATA																
title	audience market	page count	book count	entropy	standard deviation	action	adventure	anime	comedy	drama	fantasy	historical	horror	live_action	martial_arts	mystery	psychologica	school_life	sci_fi	slice_of_life	sports	supernatural	tragedy
Hajime no Ippo	shounen	15978	871	5.27	90.0	x	x	x													x		
Case Closed	shounen	12147	711	5.91	88.5		x	x	x					x		x							
Jojo's Bizarre Adventure	shounen	11258	593	6.11	87.3		x	x		x	x											x	
One Piece	shounen	10461	563	5.20	92.1		x	x	x	x	x												
Inuyasha	shounen	10236	558	5.10	85.8	x	x	x		x	x											x	
Naruto	shounen	8834	594	4.90	90.6	x	x	x	x	x	x												
Bleach	shounen	8088	403	4.49	91.6	x	x	x	x	x	x											x	
Initial D	seinen	7269	551	5.60	87.4			x													x		
Dragon Quest Dai no Daiboken	shounen	7172	349	4.51	84.6	x	x	x			x												
The Ravages of Time	seinen	6950	297	6.23	84.8	x						x											
Prince of Tennis	shounen	6680	380	5.09	86.1			x	x					x							x		
Beck	shounen	6484	103	5.31	88.3					x													
H2	shounen	6454	337	4.40	83.9									x							x		
Boys Over Flowers	shoujo	6435	242	4.97	75.2			x	x	x				x									
Violinist of Hameln	shounen	6388	142	4.88	85.4	x	x		x	x	x											x	
Gash Bell, Konjiki no Gashi	shounen	6292	323	4.78	89.8	x	x	x	x														
Rave Master	shounen	6220	296	5.08	87.9	x	x	x	x		x												
Urusei Yatsura	shounen	6218	366	5.18	88.8			x	x									x	x			x	
Shaman King	shounen	6030	302	4.59	89.6		x	x	x													x	
Flame of Recca	shounen	5933	329	5.34	90.6	x	x	x			x												
Vagabond	seinen	5827	278	6.04	86.2	x	x			x		x											
New Grappler Baki	shounen	5594	276	6.31	82.4	x									x								
Kekkaishi	shounen	5421	281	5.03	92.2	x	x	x														x	
Reborn!	shounen	5371	270	5.18	90.2	x		x	x														
Captain Tsubasa	shounen	5237	115	5.21	82.4			x													x		
Shamo	seinen	5192	257	5.06	90.4	x				x				x	x		x						
Rurouni Kenshin	shounen	5052	256	4.94	92.2	x	x	x				x											
Hunter X Hunter	shounen	5032	290	4.33	86.2	x	x	x		x													
Negima! Magister Negi Magi	shounen	4990	268	6.11	83.1	x	x	x	x		x												
Air Gear	shounen	4979	258	5.23	88.7	x	x	x	x												x		
Tsubasa Reservoir Chronicles	shounen	4852	233	4.79	90.1	x	x	x		x													x
Touch	shounen	4818	257	4.32	90.0			x						x							x		
Yakitate Japan	shounen	4761	224	4.51	80.9			x	x														
GS Mikami	shounen	4724	249	5.12	93.8	x			x													x	
Anatolia Story	shoujo	4698	97	5.23	76.3		x			x	x	x											
20th Century Boys	seinen	4694	249	5.05	85.6					x				x		x							
Skip Beat	shoujo	4572	148	4.82	77.4				x	x	x												
Nodame Cantabile	josei	4398	136	4.72	80.1				x	x				x									
Fruits Basket	shouno	4350	136	4.62	79.6				x	x	x											x	

Sample manga series from our data set sorted by total number of pages ("page count" column). "Book count" lists the number of books in each series; statistics such as "entropy" and "standard deviation" provide the average value of various visual characteristics of all pages in series as measured by our software; tag metadata shows how each series was originally classified by fans on the OneManga.com website. This table only shows a portion of the complete data about manga series, which includes additional bibliographic information, tags, and statistics.

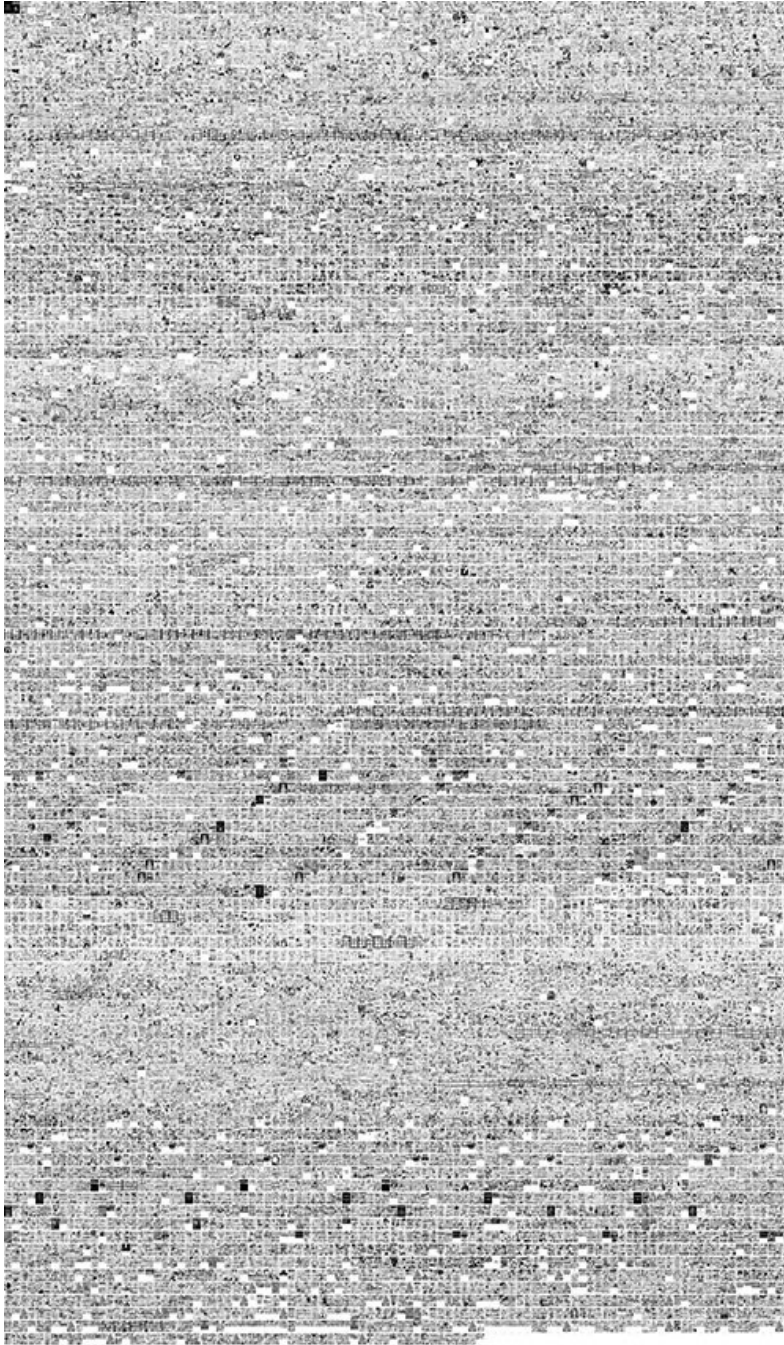
SCANLATION AS A MEDIUM

The use of computational techniques allowed us to gradually uncover various characteristics of our scanlation collection that reflect the specificity of the “scanlation medium.” While some of these characteristics are unique to the scanlation process, others are typical of our larger media universe, in which multiple versions of the same media content are created and made available both by the original producers and by fans. (For related analysis of fan inserted images in anime fansubs, see Hatcher, *Of Otaku and Fansubs*).

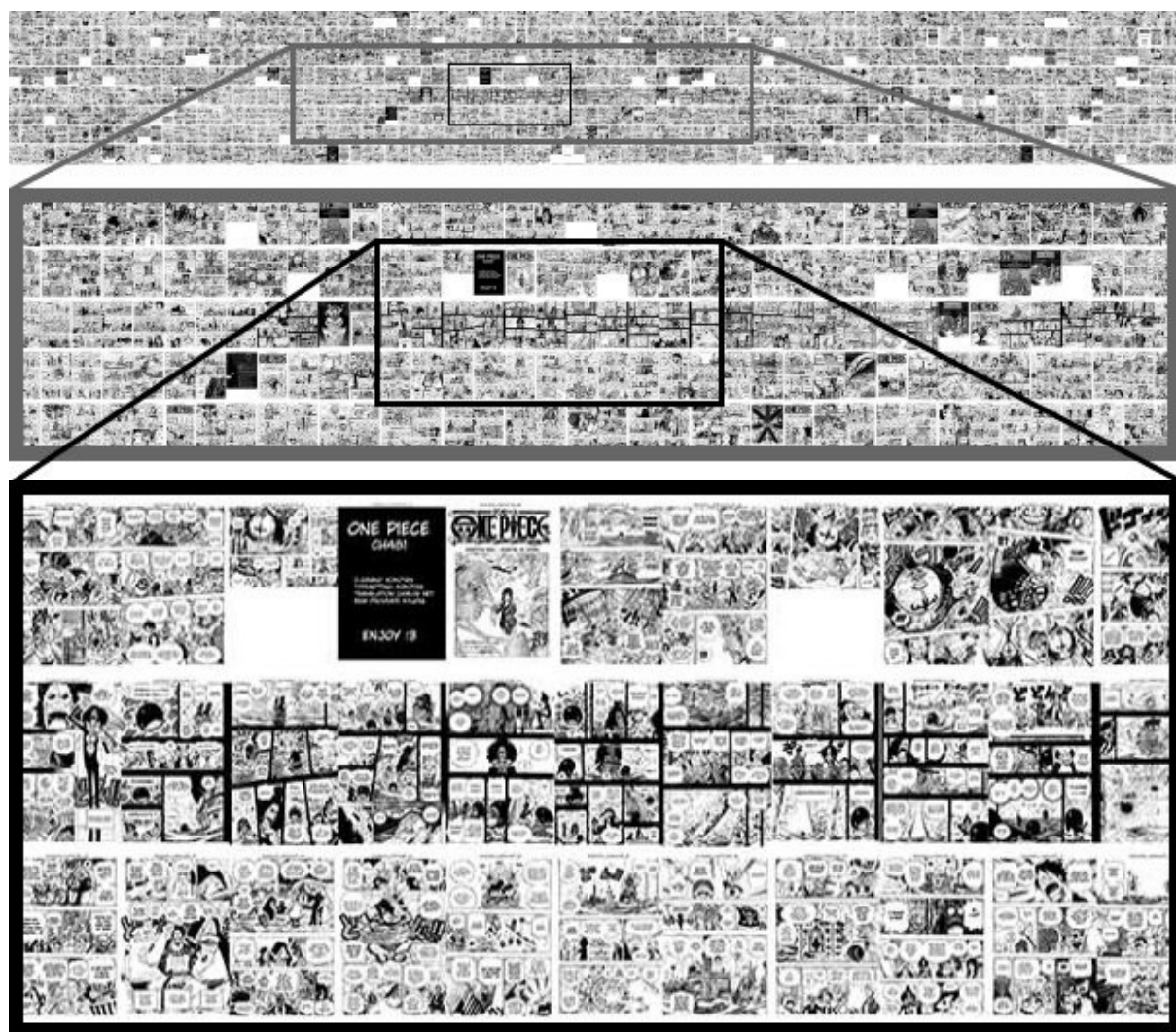
The OneManga.com data set does not always contain every book and every page for a particular title - especially in the case of very long titles running over a number of years. Which pages are scanned and translated also varies, with some groups including cover art, tables of contents, afterwords, and other paratextual material, while other groups only include the main story pages. Long-running series are more likely to vary over time in what pages are included. Such series are also more likely to contain visual artifacts that shift as a result of changes in the scanning equipment, settings, and software - resulting for example in different size, quality, or color space for the final images.



Over 5000 consecutive scanlation pages from *Anatolia Story*, as available on OneManga.com. This visualization illustrates the typical artifacts introduced as the pages are scanned, translated, and placed on the web by fans. Pages published over a number of years are often scanned by different people with different quality results: some pages may have lower resolution, lower contrast, higher compression artifacts, or entire chapters may be missing outright. Here, light bands indicate low-resolution originals.



10461 scanlation pages from *One Piece* as available on OneManga.com, organized by sequence of publication (left to right, top to bottom). This visualization includes special pages inserted by scanlation groups (some of them appear as dark black squares when you look at a small version of the visualization). To make them easier to see, the next image shows a close-up view. Note: To fit all pages in a tight rectangular grid, we used average dimensions among all pages; therefore, some of the pages shown in this visualization are cropped - in particular, rare two-page spread images.



Three progressively closer views of the *One Piece* visualization. Credit pages inserted by fans stand out from regular story pages (dark areas), gaps accompany dramatic two-page compositions (white areas), and narrative flashback sequences, indicated by black gutters between panels in the close-up (bottom) and still visible as darker bands in the distant view (top).

Along with these artifacts and omissions, we also found deliberately added pages. These additions are an important feature of “global digital manga.” Almost every series available on OneManga.com has special pages inserted by scanlation groups. These may include recruiting advertisements, announcements, or commentaries, but the most common addition is a scanlation credit page, which is typically inserted after every chapter. Credit pages usually list a small set of common production roles in the scanlation workflow, including translators, cleaners, and typesetters. They may also credit the original source of the “raws,” or original, unedited images.



Left: minimalist credit page added by the scanlation group “Deadbeat Scans” to the end of chapter 1 of the series *Shibatora*. Right: “Anime or Manga” group credit page with layered images, photo filters, logo, and text featuring a raw provider credit and a recruiting statement.

Although credit pages by scanlation groups are not marked in the OneManga.com web interface or site metadata, these special pages usually have very different visual characteristics from regular story pages - often either stark text on a black-and-white page or else a layered remix of many color elements, including sampled manga art, original fan art, photos, logos, and personal messages in ornate fonts. The following montage shows a variety of styles used for these pages using a selection from our data set:



31 scanlation credit pages selected to show the variety of identity styles used by different groups and projects. Pages may include simple text or complex layers of color graphics and photo elements. A black and white “scan-style” page on the upper left edge closely resembles a normal story page.

When viewed alongside the regular grayscale manga story pages that they accompany, most scanlation credit pages clearly stand out visually. This difference is probably intentional. By using bold visual designs, pages inserted by scanlation groups are able to differentiate themselves from the style of the main narrative and establish a separate voice from that of the (translated) artist/author. Interrupting the flow of much smaller visual differences between

subsequent story pages, the inserted pages call attention to themselves without making any apologies.



98 color pages that make up the first 7 chapters of *Digimon*. Recurring dark pages in the final column are scanlation group credits added to the end of each chapter.

In addition to their design, the majority of credit pages are visually distinctive simply because they are in color - in contrast to the typically gray story pages. In fact, along with the original series covers, color credit pages by scanlation groups make up a significant portion of all the color pages in our data set. However, even when color scanlation credits appear in all full-color chapters or series, there is usually a dramatic difference between story and fan credits, and use of color is an important aspect of that difference. Credit pages are born-digital artifacts, and they commonly have smooth gradients and full range of colors both in terms of hue and saturation. In contrast, when regular story pages use color, the result is typically more limited. While manga artists can certainly create pages in full color, the economics of manga industry prohibit this. For

this reason, color in most printed manga as it appears in magazines and weekly collections looks different from scanlator-created color, and that difference persists through the scanlation process and remains in the end product.



Examples of four different color printing styles found in the manga data set. Examples left-to-right are: full color (*Digimon*), limited palette watercolor/color-pencil style (*Chii's Sweet Home*), spot-color on grayscale (*Living Game*), and single-color ink (*Vampire Knight*).

Original manga color printing comes in a wide variety of styles, including full color, limited palette “watercolor” or “colored pencil” styles, spot-color on grayscale, and even single-color, in which all elements of the original gray scale art (including traditional always-black elements like text and panel borders) are printed in a single shade of ink such as magenta or cyan. This striking color effect arises in part out of publisher efficiency and economy, as it allows the artist to continue working in gray scale without an additional coloring step increasing production time and cost. Perhaps unsurprisingly, we found no examples of this single-color technique used in pages created by scanlators, who do not have these constraints.

Perhaps the best example of the complex relationship with color that is “negotiated” between manga scanlators and original publishers is the example of color-tinted paper. A majority of all “color” manga pages produced are actually black-ink manga art printed in weeklies on a variety of cheap newsprint. It is not the art that is color, in any sense we normally use, but rather the page, which is lightly tinted in a variety of pastel colors. Tinted paper can create an interesting visual effect, but the original motivation is primarily economic: lightly tinted recycled newsprint can be cheaper than recycled white newsprint, and manga publishers competing to produce huge

(but disposable) “phonebook” style weeklies at the lowest costs are driven to use the lowest cost materials. In older weeklies, bands of differently colored paper in the same “phonebook” proceeded seemingly at random, without respecting the boundaries between chapters, leading to reader complaints about the unpleasant aesthetic effect of paper color changing mid-story. Contemporary weeklies now generally align their recycled papers to chapter boundaries, but the surprising association of colored paper with cost-saving measures remains.

Manga scanning groups working with these early colored-paper releases may use scanning settings and digital editing to suppress these colors, transforming the art into a simple, consistent grayscale look. Depending on their technique, however, traces of the original page tint may remain in the digital image even after scanlation. Tens of thousands of pages in the manga database were initially identified by our software as “color pages,” even though that color was subtle or imperceptible to the human eye. There are several theoretically important points about this computer “misidentification” of color in apparently gray pages. One is that scanners and scanlators participate in digitally constructing an idea of the “original” manga while at the same time creating their own “translation.” Although print versions may initially hit the streets in a variety of pastel colors, the scans uploaded in uniform gray attempt to represent an idea of an earlier, truer authorial art as it is imagined to have been on the drafting table rather than hot from the printing press, and this is partly “restored” from its print version before being modified. More generally, this de-colorization practice reflects instability in the meaning of color for scanlators that can only be understood when we analyze very large collections. For a scanlation group, color may be seen either as signal or as noise, and shifts in color over time can signify boundary markers that are either undesirable or important. Scanlators may have two reasons to maintain a normative gray in most story pages. The first reason is to convey the idea of a consistent original artist and a continuous original art process. The second reason is to maintain a difference between the voice of that artist and the editorial voice of the scanlators.

The additions of pages in full color is perhaps the most extreme mechanism among many others used by scanlators identify their inserted pages. In general, pages added by scanlation groups create a “meta-narrator,” establishing the group as the ultimate authority over the original author/artist. Often this meta-narrator is introduced at the very beginning of a new chapter, and then disappears, reappearing again at the very end. To put this differently, the original manga

narrative becomes a story that is framed within a larger story told by the scanlation team - their work to bring the new chapters published in Japan to global English audiences as quickly as possible.

This interpretation is supported by our observation that only a very few credit pages actually resemble typical story pages in both lack of color and design. These “scan-style” credit pages are usually produced using the same scanlation workflow as the main story - an existing page story page is cleaned of original text, and the word bubbles have been filled in with a combination of credits and humorous commentary text.



Example of an original story page being remixed into a credit page that is then added by the scanlation group to the end of each release. Left: Scanlation story page from *Nagasarete Airantou* (chapter 4, page 10). Right: Scan-style credits / promo page for the “AT-Translations” scanlation group, introduced shortly after, at the end of chapter 6. Later versions of the credit page only altered the two lines in the upper-left corner.

Unlike most credit pages, many of which are quite easy for a human reader to immediately detect while going through a chapter, “scan-style” credits can be difficult to distinguish from story pages without additional viewing time. They are, however, fairly easy to identify with computers using algorithms for detecting duplicate and near-duplicate images. The first instance of a scan-style credit page is usually a minor alteration of a recent story page, and, like other credit page images, it is often reused multiple times thereafter with only minor alterations to text elements. So while scan-style credits are visually unusual (not because they do stand out from the story, but because they don’t!), their use follows the same general strategy as other credit page.

To generalize, the visual characteristics of scanlation pages clearly reflect the traces of their production - which in the case of credit pages is almost always free-form creation of digital images in general purpose layer-based image editing software such Photoshop, often with heavy use of fonts (as opposed to line art) or extensive use of photo layers, gradients, filters, and effects plugins. By contrast, almost all story pages in scanlated versions reflect more conservative and well-defined multi-step editing workflows of scanlation, a process intended to translate page scans through minimal alternation of the original printed artwork.

The question “how can we differentiate between two types of images in a collection of one million images?” raises both a practical and a theoretical problem. In order to find not just an answer, but a scalable solution, we are challenged to think this process of differentiation through in ways we never would have if we had focused on a traditional digital humanities hand-tagging approach. First we considered a wide variety of examples and counter-examples, then explored how their characteristics arose out of processes, then investigated the traces that those processes left in digital artifacts. Importantly, this was only possible because we are able to write new computer programs as new questions arise during the research process, as opposed to always relying on existing software.

Using a combination of digital image analysis and file metadata analysis, we were able to identify most pages inserted by scanlation teams. In our estimate, credit pages make up a little over 2% of the entire data set (twenty-one thousand images out of one million). Work to identify them originally began as a simple attempt to filter our data - increasing the “signal” of story pages by eliminating the “noise” of covers, credits, and other pages extraneous to our initial research questions. Once we began to understand these “excess” pages as a set, we realized that

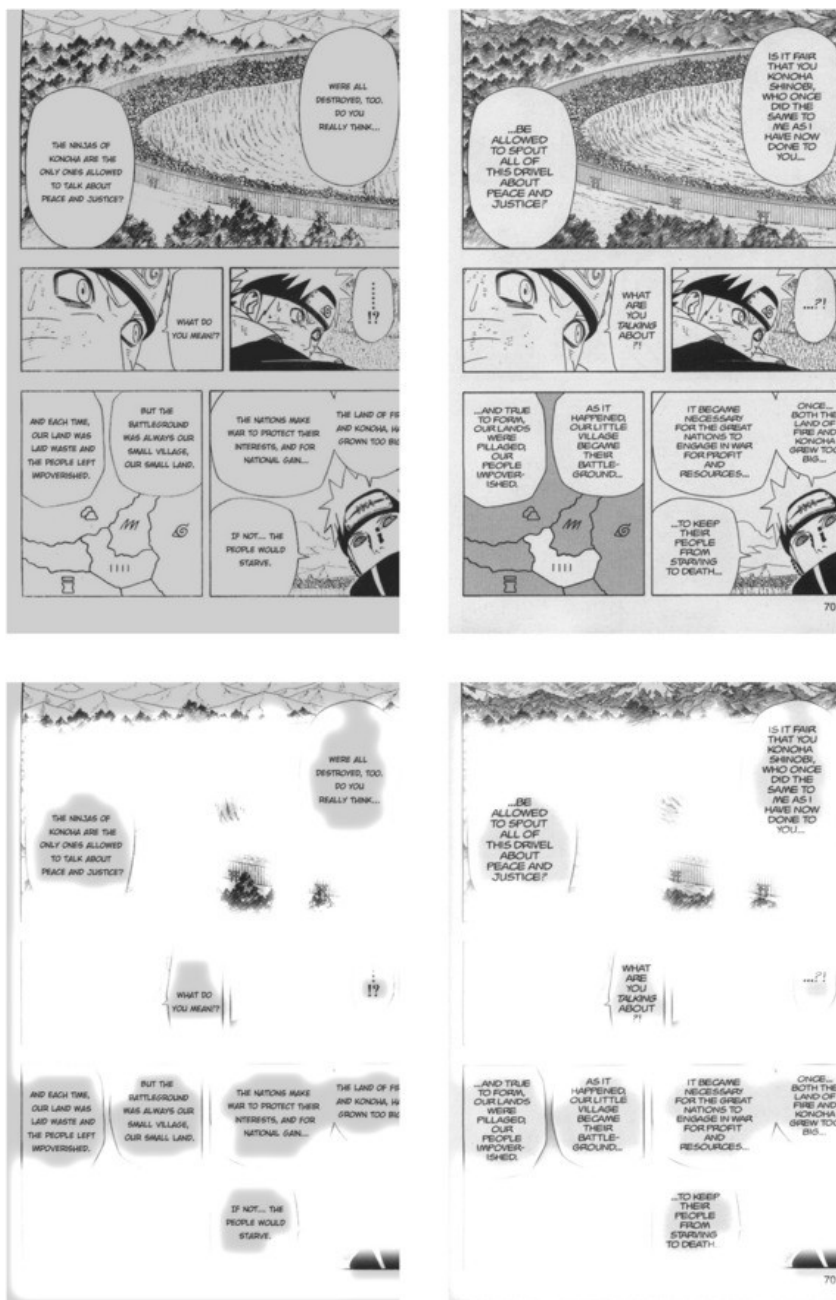
we could systematically study their characteristics and their relationships to story pages and regular covers. In short, what began as a computational analysis of manga visual language then generated a new project of its own: the analysis of visual language used by scanlation groups to help them earn bragging rights and prestige from their work.

FROM MANGA TO “MANGAS”

So far our discussion has focused primarily on just 2% of the pages in our data set and their difference from the other 98% - the original pages created by scanlation groups vs. the story pages. We will now turn to another set of differences - between the original story pages as they first appear in print, their translated versions generated by scanlators, and official translated versions which typically appear much later.

Shortly after a new chapter of a manga series appears in Japan, its translated, scanlated version becomes available online. In the case of popular titles, the Japanese publishers eventually bring out official translations - often with a significant delay. As with any translation or adaptation, fans encounter many differences between the original Japanese versions, officially translated versions, and scanlated versions.

Below are a few comparisons between the pages from scanlations and the same pages as they appeared later in the official translations licensed for foreign distribution. Using a custom image processing algorithm, we automatically aligned (“registered”) the corresponding pages from scanlations and official translations against each other. Next, our software compared these pages and generated new versions showing only the parts that were different. The results are fascinating images of only what changes from context to context, from audience to audience - the text of the translation, the short text expressions indicating sound effects, and various retouches and substitutions in the graphics themselves.



Comparison of two different English versions of the same manga page. On the left is the unofficial scanlation, produced by fans from page scans of the initial low-quality print weekly (*Naruto* issue 436, page 8). On the right is the official publisher translation of the later *tankōbon* collected release (*Naruto* volume 47, page 70). Upper images are original pages; lower images were generated with our custom digital image processing algorithm to only show page areas where lines and texture differ between the two. Additional comparison images are available on Flickr: <http://tinyurl.com/manga-compare>.

Through this process, most of the original art is screened out, leaving the text. Much of what we see is what we might expect—a collection of word balloons, consistent with the idea that scanlation in particular and manga translation in general is primarily about replacing text. However, if we look at large collections of these “difference images,” we also see several other common types of changes. Some are differences in texture between low-resolution scanlations (typically scans of cheap newsprint print runs) and the higher print and scan resolution of foreign trade paperbacks. The kinds of texture that trigger a difference are often consistent in a given work or a given artist’s style—for example, only the fine lines used in rushing action backgrounds, the textures in fine cloth, and the details in close-ups of eyes.

In addition to these subtle changes and artifacts, we also see elements in manga that are generally untouched in the scanlation process, but often reworked in official translation. Some of these are improvements in the overall “production values” of the original images that were added once the pressure of initial deadlines was gone - adding additional shading, retouching lines, or increasing the level of detail in backgrounds. For example, in the pages from *Naruto* in the illustration above, the sketched mountains in the initial weekly and the scanlation were fully textured with snowy caps in the official translation. The software also picks out a pair of small image sections in the middle of each page - both places where the tree line was edited back in the later release in order to more clearly reveal gates along the city wall. Examples like this second one may reflect the differences in software used by scanlators and by the artists and editors working on translated version. When original artists or official translators have the opportunity to work with original digital art assets in layer-based image editing software, they can easily make changes to specific layers of the artwork. Such layer-based changes would be difficult and time-consuming for scanlators working with the final flattened output as it was printed.

When publishers release official translations of the original Japanese editions, they also commonly rework the written/graphic elements indicating sounds and emotions, replacing them with English equivalents - “BOOM,” “WOW,” etc. In contrast, fans doing scanlations generally leave these short elements alone. This may be for several reasons - in part because they are one element of the original artist’s graphic design, which must be preserved, in part because they are integrated into the images (and thus difficult to change), and in part because they reflect a “foreignizing” aesthetic, popular with hardcore fans and Japanophiles, as opposed to a

“domesticating” aesthetic, preferred by importers in order to target the widest possible market. (For more on the aesthetics and politics of foreignizing vs. domesticating translation, see Venuti). The example below illustrates this; it also shows how the same text was translated differently by the scanlation group and by the publisher.



Comparison of two different English versions of the same manga page. Only the elements that differ between the two versions are shown. Left: scanlation from the original Japanese publication of *Naruto* (issue 433, page 14). Right: the later official translation into English by the publisher.

Perhaps the most theoretically interesting thing about studying scanlation using software-generated “difference images” is how those images complicate our initial concept of scanlation as a simple process that preserves pictorial elements while replacing Japanese text with translations. Instead, we discover a wide variety of changes in the images, as well as a

breakdown in the distinction between text and image: elements that are both text *and* image (as with sound effects), and even plain text being treated as image. For example, both of the previous illustrations contain text bubbles with the punctuation marks “!?” Because these marks exist in both the source and the target language, scanlators often leave the original alone, following a policy of minimal intervention. As a result, the undisturbed image of the original “!?” may stand out visually in the later scanlation: in the full-page illustration we see that the font is different from other English text on the page, and even the ellipsis is in a Japanese vertical style rather than horizontal. Quickly browsing thousands of *Naruto* pages, we find that, whether the punctuation is retyped or not, the *Naruto* scanlators consistently treat expressions such as “?!?” like a visual unit, or an image, and leave them unchanged. *Naruto*’s official translators, on the other hand, consistently treat them like language, and reverse them to “!?” This different treatment of the same symbols illustrates the subtle ways that one manga series is transformed as it moves through different software production workflows under the guidance of the unofficial and official editors with different aesthetic sensibilities and different audiences in mind.

The global “manga universe” includes a number of different “mangas” recognized as legitimate artistic forms of their own. The term “original English-language manga” or OLM is commonly used to describe “international manga” which appears outside of Japan in other languages. (The term “international manga” is used by Japanese Ministry of Foreign Affairs.) *Dōjinshi* (or *doujinshi*) are manga created and self published by Japanese non-professional artists.

These terms describe different cultural objects created by different individuals or teams. But they do not capture the rich and complex life of a single media title or series as it travels between countries, languages, official publishers and multiple fan cultures, changing its identity in subtle but important ways. A single manga series may exist in at least half a dozen forms, including publications in Japanese magazines, separate black and white volumes, color editions, multiple scanlated versions, and official translations.

So far, media scholars paid more attention to other kinds of content transformation described by concepts such as “adaptation,” “remix,” “database consumption” (Hiroku Azuma, *Otaku: Japan’s Database Animals*) and “transmedia.” This later concept describes how media assets which form a single “story world” in which characters, words, and stories are transformed into products in multiple media such as manga series, anime, movies, video games, digital wallpaper,

toys, and so forth. But less obvious differences between the many versions of an object that exist in the same medium have received less critical attention.

As the examples presented and analyzed above make clear, “global digital manga” (manga scanned and translated by fans and made available online) is different from “Japanese print manga” (new manga chapters published in weekly magazines in Japan and later collected in volumes), and is also different from “global print manga” as experienced by global readers who buy officially translated series. Rather than thinking of this fan-produced global digital manga as merely an inferior and distorted copy of the original Japanese manga publications, we should recognize it as equally valuable for research - because this is how many millions of readers around the world encounter manga. A clear picture of the differences between these manga universes only emerged because we applied digital image analysis to our data set and selected samples of other “mangas.” This involved writing special algorithms to address particular questions that arose in our research (such as identifying differences between scanlation pages and corresponding officially translated pages).

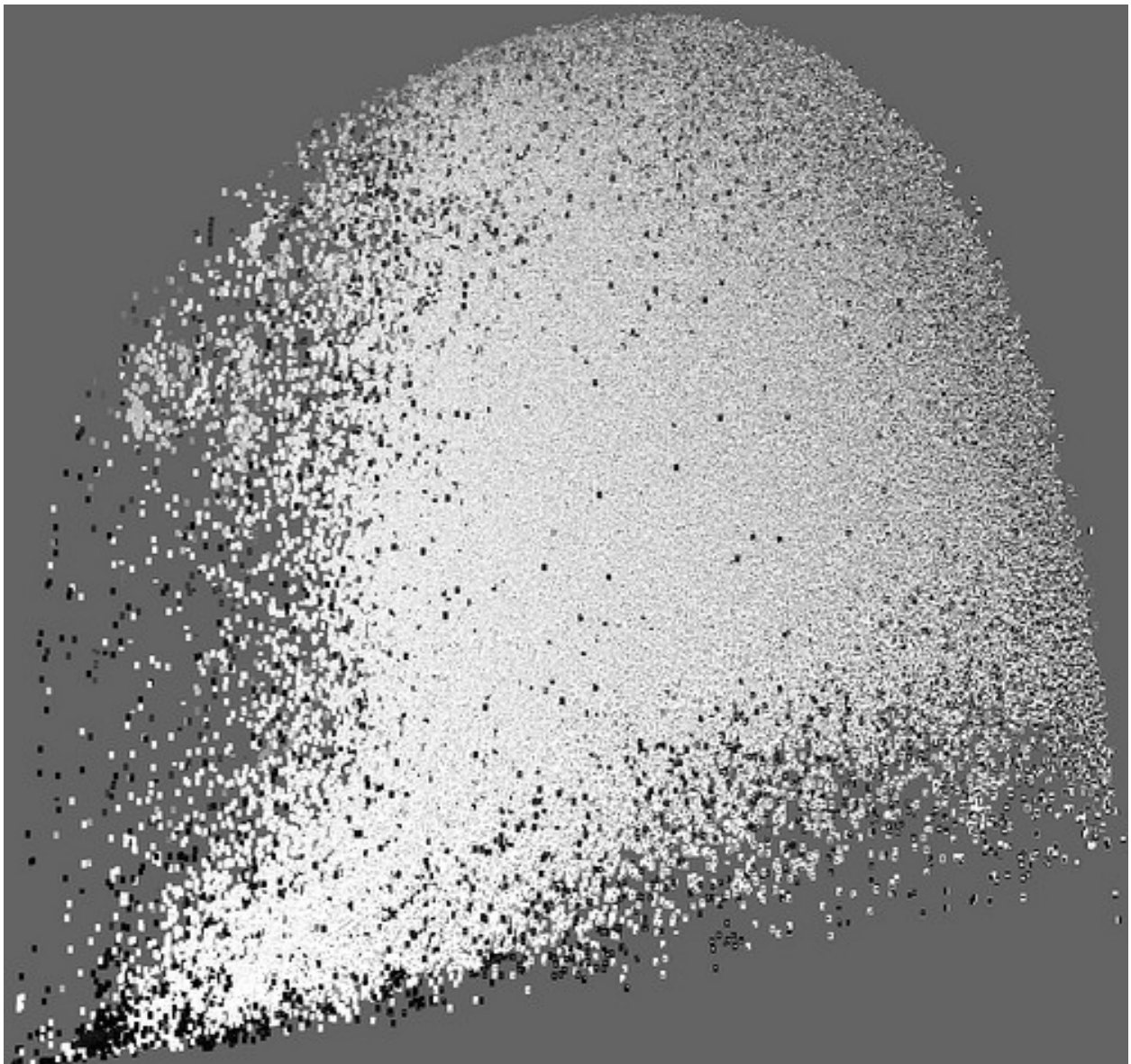
The legitimacy of scanlations as being no less valuable than other mangas is also connected to the fact that what we may think as the true original publication of a series is just one of the “renderings” of the manga “data.” Because weekly chapters are published on a cheap newsprint paper (which may change one day if Japanese manga goes all digital), fine details drawn by manga artists disappear. Book-size *tankōbon* volumes that collect a number of previously chapters and use better paper are more true to the actual artist’s drawings. Popular titles may eventually also be turned into expensive deluxe volumes which feature many color drawings created by artists just for these editions. So which of these generations of a title is the “original”?

VISUAL CONCLUSION: ZOOMING OUT

As a way of closing this first article in a series of reports on our one million manga pages project, let us zoom out - moving from the details of individual pages to an overview of our complete data set. The following visualization organizes all pages by their visual characteristics. Specifically, the positions of the pages along vertical and horizontal axis are determined by their values on separate visual dimensions calculated by digital image analysis software. (Remember

that since interactive visualization allows us to look at the visual data sets in multiple ways, this particular visualization represents just one of these possibilities.)

The vertical axis (Y) represents a dimension which can be described as presence / absence of texture and fine details. Pages that are more graphic and contain little texture and few details appear in the lower part of the visualization; pages with fine textures and lots of details appear in the upper part. The horizontal axis (X) represents another visual dimension: the range of gray tones in an image. Pages that only contain black and white appear on the extreme right; images where all grey tones fit within a smaller range appear on the left.



popular titles such as *Naruto* and *One Piece* (i.e. where do they appear in this space)? Which manga series out of the 883 we have in our data set have the most unique visual languages? And, perhaps most importantly: does the very concept of a “visual language” - along with other widely used critical terms such as “genre” and “style” - actually help us to understand really large cultural collections?

ACKNOWLEDGMENTS

Software Studies Initiative research reported in this article was made possible by the generous support provided by California Institute for Telecommunication and Information (Calit2), Center for Research in Computing and the Arts at UCSD (CRCA), and Chancellor Office of University of California, San Diego (UCSD). The development of custom digital image analysis software and the processing of manga images on supercomputers at the National Energy Research Scientific Computing Center (NERSC) was funded by a Humanities High-Performance Computing grant from the NEH Office of Digital Humanities.

REFERENCES

- Azuma, Hiroki. *Otaku: Japan's Database Animals*. University of Minnesota Press, 2009. 47-58. Print.
- Brienza, Casey. "Paratexts in Translation: Reinterpreting 'Manga' for the United States." *The International Journal of the Book*, 6(2), 2009.
 <http://cambridge.academia.edu/CaseyBrienza/Papers/145365/Paratexts_in_Translation_Reinterpreting_Manga_for_the_United_States>
- Brienza, Casey. "Producing comics culture: a sociological approach to the study of comics." *Journal of Graphic Novels & Comics*, v1(2), 2010. p. 105—119.
 <http://pdfserve.informaworld.com/760063__931232617.pdf>
- Bolter, Jay David and Richard Grusin. *Remediation: Understanding New Media*. The MIT Press, 2000.
- Cha, Kai-Ming. "Manga Publishers Face Uphill Battle Against Scanlations." *Publisher's Weekly*. 29 June 2010. Accessed 25 Nov. 2010. <<http://www.publishersweekly.com/pw/by->

topic/book-news/comics/article/43672-manga-publishers-face-uphill-battle-against-scanlations.html>

de Certeau, Michel. *The Practice of Everyday Life*. University of California Press, 2002.

Deppey, D. "Scanlation nation: amateur manga translators tell their stories." *The Comics Journal*. 2005.

Doria, Shawn [a.k.a. gum]. "Inside Scanlation: Inside the world of scanlation." Accessed 25 Nov. 2010. <<http://insidescanlation.com/>>.

Eastman, Angela. "American & Japanese Manga Companies Form Anti-Piracy Coalition." *Suite101.com*. 8 Jun 2010. Accessed 25 Nov. 2010. <<http://www.suite101.com/content/american--japanese-manga-companies-form-anti-piracy-coalition-a246625>>

Eastman, Angela. "Manga Scan Site Will Remove Unauthorized Content by End of July." *Suite101.com*. 22 July 2010. Accessed 25 Nov. 2010. <<http://www.suite101.com/content/manga-scan-site-will-remove-unauthorized-content-by-end-of-july-a264657>>.

Hatcher, Jordan S. "Of Otaku and Fansubs. Appendix - Fansub Samples" (PDF). Script-ed. Vol. 2, No. 4, 2005. Accessed December 11, 2010. <http://www.law.ed.ac.uk/ahrc/script-ed/vol2-4/otaku_appendix.pdf>

Hilderbrand, Lucas. *Inherent Vice: Bootleg Histories of Videotape and Copyright*. Duke University Press. 2009.

Jenkins, Henry. "Transmedia Storytelling." *Technology Review*. 15 Jan. 2003. <<http://www.technologyreview.com/biomedicine/13052/>>.

Leavitt, Alex. "Inside Scanlation [Web site review]". *Transformative Works and Cultures*, no. 5. 2010. <<http://journal.transformativeworks.org/index.php/twc/article/viewArticle/215/172>>

Lee, Hye-Kyung. "Between fan culture and copyright infringement: manga scanlation." *Media, Culture & Society*, v31, November 2009. p. 1011-1022. <<http://mcs.sagepub.com/content/31/6/1011.citation>>.

Lee, Hye-Kyung. "Manga Fan Culture and Scanlation: Implications for Cultural Marketing." *Marketing the Arts: A Fresh Approach*. O'Reilly and Kerrigan, eds. Routledge, 2010.

- Lessig, Lawrence. *Free Culture: how big media uses technology and the law to lock down culture and control creativity*. Penguin Press, 2004. Ch. 4. <<http://www.free-culture.cc/freecontent/>>
- Macias, Patrick. "Fans Lift J-Culture over Language Barrier." *The Japan Times*. 7 Sep 2006. Accessed 25 Nov. 2010. <<http://search.japantimes.co.jp/cgi-bin/fa20060907a1.html>>
- Marshall, Lee. "For and against the record industry: an introduction to bootleg collectors and tape traders." *Popular Music*, v22, 2003. p. 57-72.
- Nowlin, N. Scanlations: Copyright Infringement for Literature and Art Fans Brought to You by the Internet. Series in *Interface on the Internet* 10(1,8,9). 2010. <<http://bcis.pacificu.edu/journal/article.php?id=633>>.
- Office of Digital Humanities. "Announcing the Digging into Data Challenge," <<http://www.diggingintodata.org>>. Accessed 9 Dec. 2010.
- OneManga.com web site. <<http://www.onemanga.com/>>.
- "Original English-language manga." *Wikipedia: The Free Encyclopedia*. Wikimedia Foundation, Inc. Accessed 20 Nov. 2010. <http://en.wikipedia.org/wiki/Original_English-language_manga>.
- Pink, Daniel H. "Japan, Ink: Inside the Manga Industrial Complex." *Wired*. 22 Oct. 2007. <http://www.wired.com/techbiz/media/magazine/15-11/ff_manga>.
- Rampant, J. "The Manga Polysystem: What Fans Want, Fans Get." *Manga: An Anthology of Global and Cultural Perspectives*. Toni Johnson-Woods ed. 2009. Continuum Intl. Pub Group.
- Reid, Calvin. "Japanese, U.S. Manga Publishers Unite To Fight Scanlations." *Publisher's Weekly*. 8 June 2010. Accessed 25 Nov. 2010. <<http://www.publishersweekly.com/pw/by-topic/digital/copyright/article/43437-japanese-u-s-manga-publishers-unite-to-fight-scanlations.html>>.
- Schodt, Frederik L. *Dreamland Japan: Writings on Modern Manga*. Stone Bridge Press, 1996.
- Schulze, Hendrik, and Klaus Mochalski. *Internet Study 2008/2009*. Rep. Ipoque, 18 Feb. 2009. Web. Accessed 10 Dec. 2010. <<http://www.ipoque.com/resources/internet-studies>>.
- Taneska, Biljana Kochoska. "Otaku – the Living Force of the Social Media Network." Third International Communication Conference -- Glocal: Inside Social Media. New York

University Skopje, Macedonia. 15 Oct 2009.

<[http://glocalconference.com/downloads/Kochovska B_Otaku_pdf](http://glocalconference.com/downloads/Kochovska_B_Otaku_pdf)>

"Top 25 Manga Properties, Q3 2010." *ICv2: Inside Pop Culture*. 6 Oct. 2010. Web. Accessed 10 Dec. 2010. <<http://www.icv2.com/articles/news/18501.html>>.

Venuti, Lawrence. "Genealogies of Translation Theory: Schleiermacher." *TTR : traduction, terminologie, redaction* v4(2), 1991. p. 125-150.

Venuti, Lawrence. "Translation and the Formation of Cultural Identities." *Current Issues in Language and Society*, v1(3), 1994. p. 201-217.

Yang, Jeff. "Manga Nation: no longer an obscure cult art form, Japanese comics are becoming as American as apu pai." *SF Gate*. 14 Jun 2004. Accessed 25 Nov. 2010.

<<http://www.sfgate.com/cgi-bin/article.cgi?f=/g/a/2004/06/14/manganation.DTL>>

Manga series

Chie, Shinohara. [Anatolia Story](#). (also known as *Red River*).

Eiichiro, Oda. [One Piece](#).

Hiroshi, Izawa and Yabuno Tenya. [Digimon Adventure V-Tamer 01](#).

Masashi, Kishimoto. [Naruto](#).

Matsuri, Hino. [Vampire Knight](#).

Software Studies Initiative

Lab blog. <<http://www.softwarestudies.com>>.

Still visualizations. <<http://www.flickr.com/photos/culturevis/collections/>>.

Animated visualizations. <<http://www.youtube.com/user/softwarestudies>>.

Software tools. <<http://code.google.com/p/softwarestudies/>>.