

Foucault@Wiki

First Steps Towards a Conceptual Framework for the Analysis of Wiki Discourses

Christian Pentzold, Sebastian Seidenglanz

University of Technology Chemnitz

Media Communication

Thüringer Weg 11

D-09126 Chemnitz, Germany

+49 (0) 371 531 34529

{pech | sese} @

hrz.tu-chemnitz.de

ABSTRACT

In this paper, we examine the discursive situation of Wikipedia. The primary goal is to explore principle ways of analyzing and characterizing the various forms of communicative user interaction using Foucault's discourse theory. First, the communicative situation of Wikipedia is addressed and a list of possible forms of communication is compiled. Second, the current research on the linguistic features of Wikis, especially Wikipedia, is reviewed. Third, some key issues of Foucault's theory are explored: the notion of 'discourse', the discursive formation, and the methods of archaeology and genealogy, respectively. Finally, first steps towards a qualitative discourse analysis of the English Wikipedia are elaborated. The paper argues, that Wikipedia can be understood as a discursive formation that regulates and structures the production of statements. Most of the discursive regularities named by Foucault are established in the collaborative writing processes of Wikipedia, too. Moreover, the editing processes can be described in Foucault's terms as discursive knowledge production.

Categories and Subject Descriptors

H.5.2 [User Interfaces]: Natural language. H.5.3 [Group and Organization Interfaces]: Computer-supported cooperative work, Theory and models, Web-based interaction, Asynchronous interaction. H.5.4 [Hypertext, Hypermedia]: Theory.

General Terms

Design, Human Factors, Languages, Theory

Keywords

Wiki, Wikipedia, Computer-mediated communication, Online collaboration, Foucault, Discourse theory

1. INTRODUCTION

"Slouching toward the ordinary." [1] – that was the commonly accepted vision of computer-mediated communication (CMC) at the beginning of the new millennium. Regarding the ubiquity and

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

WikiSym '06, August 21–23, 2006, Odense, Denmark.
Copyright 2006 ACM 1-59593-413-8/06/0008...\$5.00.

indispensability of CMC this statement seems to be uncontroversial. But with regard to Wikis as one of the most popular online tools in the recent history of the Web it can be argued that there is a rather galloping change of the discursive horizon [2]. Even though their primary purpose is not to support communication and their "discursive features have not attained full maturity" [2] they cannot be neglected as a new form of online communication.

Wikis and especially Wikipedia have already generated a significant amount of scientific writing. The linguistic features however have only recently become a field of research. Therefore, the goal of this paper is to explore whether or not a detailed discourse analysis of the communicative situation of Wikipedia is a fruitful approach towards the understanding of the collaborative writing processes as well as the social and technical structure of this Wiki system. Up to now, the attempt to characterize the activities of the authors meets considerable difficulties.

Thus, the paper aims to study the usefulness of discourse theory as a qualitative approach of dealing with Wikipedia as a socio-technical, highly prolific online collaboration platform. Only few attempts have been made to apply Foucault's work to online material. This can partly be explained by the lack of almost any form of coherent methodological framework; his style of argumentation is far from straightforward. Despite these inherent theoretical obstacles it is argued that an exploration of discourse theory seems to be rewarding in terms of a more complete analysis of the textual material in connection to the social interaction and technical framework.

The paper addresses these problems in the following manner: First, the communicative situation in Wikipedia is described with the help of a list of forms of possible communicative interactions. Second, the research on linguistic aspects of Wikipedia and Wikis in general is reviewed. In the third section, the theoretical groundwork for a discursive analysis is laid, describing some of the key issues of Michel Foucault's theory: 1. the notion of discourse as a practice that regulates and structures the production of statements, 2. the concept of discursive formations and 3. the methods of archaeology and genealogy. Five steps of dealing with Wikipedia as a discursive formation are outlined. In the final section, a tentative approach towards a Foucault inspired

discourse analysis of a corpus from the English Wikipedia will be explained.

2. THE COMMUNICATIVE SITUATION

Although the Wikipedia is first and foremost an encyclopedia and despite the declarative statement on the Meta-Wiki page [[What Wikipedia is not]] that Wikipedia is not a discussion forum, it also contains a plethora of different ways of communication. While looking at the Wikipedia and its related Meta-Wiki sites today, one will find a complex network of pages where users can communicate with each other, exchange arguments, dispute, request (un-)deletions, ask for help, report vandalism, talk on IRC channels, subscribe to mailing lists or conduct surveys. For this reason, it seems appropriate to describe the synchronous communicative situation of Wikipedia before engaging in an analysis of the material. Given the constant growth and diversification of Wikipedia this list must be regarded as being far from exhaustive.

2.1 The talk pages

The talk pages are the most common form of Wikipedia communication. Every article is accompanied by a talk page whose purpose cannot be reduced to one single function. On the contrary, they are, for example, used to discuss the general direction of an article, its structure, scope and connection to other topics and fields of knowledge. In addition, they offer a place where authors can debate contributions or information quality issues.

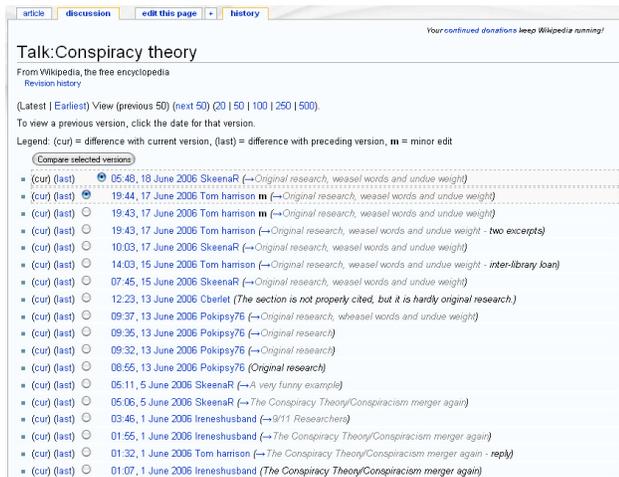


Figure 1. The talk page of the article [[Conspiracy theory]]

The talk page presents itself as a comparatively unstructured forum without a predetermined topical framework. The authors develop its structure the moment they start a new line of argumentation. Their initial statement is followed by responses stating the author and the exact point of time. Sometimes, the talk on these pages outweighs the actual content with respect to its volume. For example, the talk pages of the article [[Conspiracy theory]] are approximately ten times as long as the associated article. Unsurprisingly, this effect seems to correlate with the importance or controversy of a topic. Apart from the articles, the user pages and almost all other pages of the Wikipedia feature a discussion page. Voss [3] showed that the percentage of used sites varies between different language versions. Whereas in the

Danish version only 6.8% of all pages were talk pages, the German one contained 11.3%, and the Japanese 9.8%.

2.2 The page history

A second prominent form of user communication is the page history, a site where all edited versions of the article are listed with the date, the author and usually a comment.

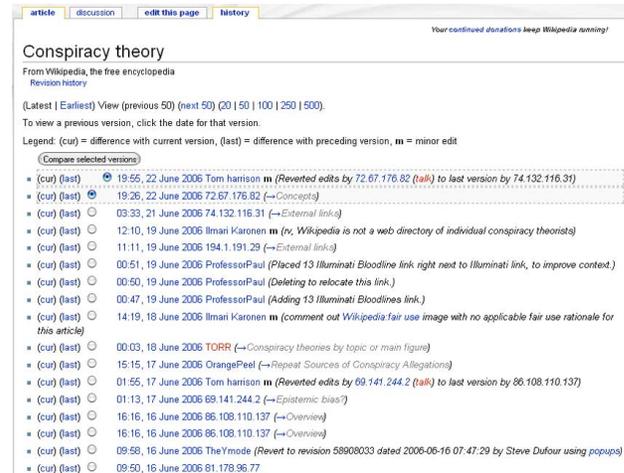


Figure 2. The page history of the article [[Conspiracy theory]]

At first glance, these pages do not seem to provide the necessary setting for user communication. But all editors are asked to comment on their contributions, i.e. to describe and justify their alterations. Viégas, Wattenberg and Dave [7] noticed that during edit wars these normally short messages could grow into mature discussions, which are closely connected to the topics debated on the discussion pages.

2.3 The village pump

Apart from these two rather article centered pages many other forms of user communication have almost mushroomed in the Wikipedia structure. The [[Village pump]] presents the most general platform consisting of a set of different topic related pages divided into six subsections: News, Policy, Technical, Proposals, Assistance and Miscellaneous.

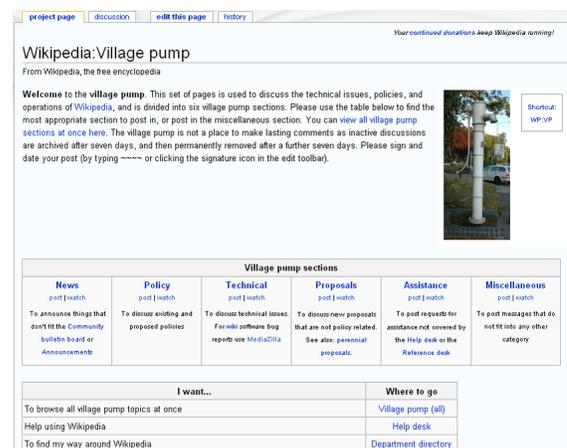


Figure 3. The entry page of the Wikipedia [[Village pump]]

It is a dynamic discussion board following the same principles as the talk pages. Every author has to sign her post, which will be archived after seven days, and removed after a further seven days. However, they will still be accessible via the page history of the [[Village pump]].

2.4 The bulletin board

The bulletin board of the Wikipedia community portal offers a site where the users can post news and announcements. It supports the organization of the Wikipedians as a central place where one can find information, for instance, about current projects and upcoming tasks, new initiatives as well as collaborations. By contrast to the village pump and the talk pages this site does not generate much textual material. Its principal function is limited to the publication of short notes.

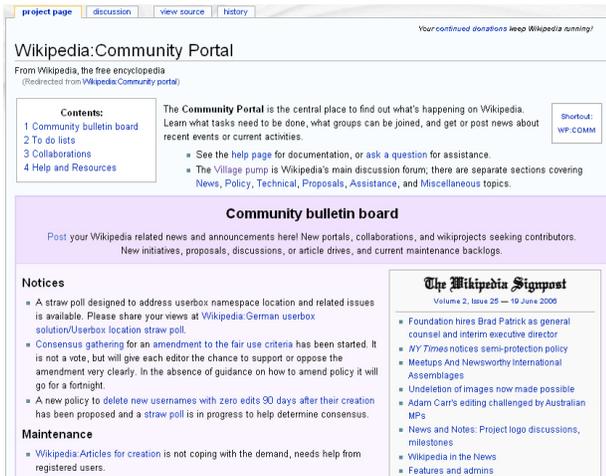


Figure 4. The bulletin board of the [[Community portal]]

2.5 The comment pages

This section contains all different forms of ‘request a comment’ pages, which are designed for resolving topical, personal and policy disputes. In order to request comments or to conduct a survey on an issue, for example, the behavior of a second user or the appropriateness of a topic, one has to state the name or title, an explanatory statement, the date and the link to the related talk page. The requests are separated into 24 issue-related sections including Economy and Trade, Politics, User names and the Misuse of Admin privileges. The underlying principle of this practice is the concept of consensus as the most desirable form of decision-making. Therefore, a lot of the Wikipedia processes are debated by the community.



Figure 5. The main page of [[Requests for comment]]

Beside these general requests there are also more specialized forms like the requests for deletion and undeletion or the requests for arbitration and mediation.

2.6 The mailing lists

The several (ca. 50) Wikipedia mailing lists are either open to public or (in a few cases) only open to selected users. Most of them are moderated and the posts are manually approved. They range from research and announcement lists to Wikimedia project lists (Wikispecies, Wikibooks, Wikinews etc.) and local lists (French, Italian, Sanskrit etc.). Mailing lists are available in a number of formats: by email, via the Wikipedia web archive, or by NNTP using the mail-to-news gateway Gmane. Moreover, the posts are indexed by search engines and can also be found at Gmane and Gossamer Threads (a mailing list archive).

2.7 The IRC channels

Whereas the talk pages, bulletin boards, comment pages and mailing lists offer a more or less asynchronous form of communication the Wikipedia IRCs are designed as synchronous chat forums. Most of these channels (almost 80 are currently listed) have been created for the purpose of a particular project, a language version, or a technical issue.

3. ANALYZING WIKI DISCOURSES

Up to now, social research on Wikis has mainly focused either on content quality [4], [5] or user motivations [6]. However, the Wiki discourses described above have only recently become a field of scholarly engagement.

In their influential study of the dynamic editing processes Viégas, Wattenberg and Dave [7] described the application of the so-called *history flow*-visualization. This software allowed them to analyze and display the complex structure of the evolution of Wikipedia articles by visualizing the textual contributions of different authors at different points of time. As a result, they presented a system of patterns of vandalism, including mass deletion (90% of an article deleted), offensive copy (insertion of vulgarities), phony copy (insertion of unrelated material), phony redirect (redirect to unrelated article) and idiosyncratic copy (insertion of biased, one-sided material). Moreover, the study showed that during edit wars the contributions on the discussion pages escalated. These periods of conflict seemed to generate the

need for a broader discussion of certain arguments, which took place at the article related discussion page. Yet Viégas, Wattenberg and Dave concluded that further research is needed to shed light on the systematic correlations between articles and discussion pages.

Another point of departure for analyzing the linguistic structure has been established by Emigh and Herring [8] who used Anthony Giddens' structuration approach [9] to formulate a genre analysis of Wikipedia as a repository of encyclopedic knowledge. To this end, they used corpus linguistic methods and factor analysis of word counts to review different levels of formality and informality in both the articles and the discussion pages. Formality was measured by counting non-formative suffixes (e.g., -ment, -(t)ion, -ity) and informality by counting contractions (e.g., I'm, don't, he's) and personal pronouns (I, we, you etc.). They showed that the articles of Wikipedia resemble traditional print versions in terms of language formality and standardization. In contrast, the discussion pages were proved to be consistently informal reflecting the stylistic features of web-boards and asynchronous discussion forums. The dual structure of Wikipedia consisting of both an article and a discussion page seems to promote, on the one hand, a reproduction of traditional print norms in the articles and, on the other, the use of a more informal language in the discussions. Emigh and Herring postulated two tentative reasons to explain their findings. First, the Wikipedia authors apply learnt norms and expectations about lexical writing style. Second, those norms are constantly enforced through the permanent editing processes and the agency of socially-approved members (e.g., the *sysops*) of the Wikipedia community.

Resnick et al. [10] highlighted the Wiki structure because of its advantages in relation to other forms of asynchronous online communication that are characterized by the following elements: topics, threads, permanence, homogenous view and the absence of summarization. In opposition to that, the Wikis with their "new fundamental entity: the editable node" seem to establish a new form of editing pushing the boundaries of conventional online communication.

Moreover, other approaches stress the productive power of Wiki discussions in the collaborative knowledge creation [11], [12]. Under this perspective, the ongoing processes of contribution and discussion as well as the surveillance of the articles help to maintain a form of soft security that protects the data better than any form of restrictive rules.

The crucial factor for this positive function of conflict in the Wikipedia work processes seems to be the integrative potential of the community, i.e. the ability to facilitate effective negotiations and disputes. Therefore, Reagle [13] regards the discussion features of Wikipedia as means of transforming divergent controversy into convergent controversy so that the project can constantly attract new authors and turn into the prophesized "hypertext on steroids" [14: p. 27] instead of crumbling under a gargantuan accumulation of edit wars, arbitration requests, bannings, deletions and discussions.

More ambitious perspectives understand these open forms of authoring and editing as the realization of Enzensberger's demand for an emancipatory media usage [15] and of Habermas' concept of an autonomous, non-hierarchical [*herrschaftsfrei*] discourse, respectively [16].

Apart from that, topic related discussions have furthermore been used to examine the information quality processes of Wikipedia articles. Stvilia et al. [17] for example made use of the article discussion pages to compile a list of ten information quality problems named by the authors: accessibility, accuracy, authority, completeness, complexity, consistency, informativeness, relevance, verifiability and volatility. On the basis of 60 randomly chosen articles they were able to show that accuracy was the pivotal point of information quality assurance discussions (57 entries), followed by completeness (49), verifiability, (19) and relevance (18).

4. A FOUCAULDIAN PERSPECTIVE

4.1 Theoretical Background

At first sight, the discourse theory of Michel Foucault does not seem to offer a useful position concerning the analysis of Wikis. The main point of criticism is the almost complete absence of any form of a coherent methodological framework. Neither is his method easy to follow nor did he provide an accessible guide of how to engage in discourse analysis. Despite these theoretical hurdles a closer look at his work, especially *The Archaeology of Knowledge* (L'Archéologie du Savoir) [18] and *The Discourse on Language* (L'ordre du discours) [19], can lead to the formulation of first steps towards a fruitful approach of examining Wikis. Yet a full-scale exploration of his work would obviously exceed the scope of this paper. Nevertheless, it is feasible to contemplate some key issues and apply them to an example from the English Wikipedia.

4.1.1 Discourse

In the scientific literature the name Foucault is closely connected to the term discourse. In contrast to concepts that view discourse as every form of conversation or as a form of academic disputation he used it to refer to "the general domain of all statements, sometimes as an individualizable group of statements, and sometimes as a regulated practice that accounts for a number of statements". [18: p. 80] Firstly, discourse is a general domain insofar as it can contain all meaningful, oral or written utterances with some effect [see 20]. Secondly, discourse understood as individualizable group refers to a certain discursive grouping, e.g. that of punishment or sexuality. Thirdly, discourses are practices that regulate and structure the production of statements, practices that "systematically form the objects of which they speak." [18: p. 49]

Thus, discursive practices are delimiting the field of objects, defining a legitimate perspective and fixing the norms for the elaboration of concepts [see 21]. In order to produce a statement in a discourse one has to adapt to the constraints, focus on a subject and claim authority for herself: "one would only be in the true if one obeyed the rules of some discursive 'policy'". [19: p. 224] In the *Discourse on Language* Foucault emphasizes this regulatory function. The "great proliferation of discourse" [ibid.: p. 229], its power and danger, is at once controlled, selected, organized and redistributed.

It is incorrect though to think of discourses as human centered. In fact, he considers them to be an impersonal system, "a nameless voice" [ibid.: p. 215].

With his emphasis on constraints, norms and authority Foucault opens the field for a detailed discussion of discourse as “the great game of contradiction” [18: p. 153] – a topic highly related to the collaborative Wikipedia processes. In the course of his argumentation he separates different **levels of dissension**: the inadequation of the objects spoken about, the *divergence of the modes of argumentation*, the *incompatibility of concepts*, and the *exclusion of options*. The **functions of conflict** are various, too. Foucault draws the attention to the *additional development* of the discourse, the *reorganization* of the discursive field and the *critical role* played by contradictions in the definition of acceptability.

Another essential issue in the light of the paper’s topic is Foucault’s view on knowledge. “Knowledge is that of which one can speak in a discursive practice” [ibid.: p. 182]. It is not identical with information. By contrast, knowledge is not thinkable without a discourse as the site of its dynamic constitution. It is a cognitive as well as social phenomenon, which is closely connected to power. ‘Power/knowledge’, as this compound is termed, describes knowledge as a conjunction of power relations and information-seeking [see 20]. Knowledge is always produced in the struggles and debates of power/knowledge. Therefore, power is not understood as a centralized force but rather as a regime of mechanisms in their “capillary form of existence” [23: p. 39] that reach into the production of every single statement. It is not the domination of one group over others. By contrast, power is “employed and exercised through a net-like organization” [24: p. 98] in which the subjects are simultaneously in the position of undergoing and exercising power. This exercise of power itself leads to the emergence of new objects of knowledge; power and knowledge are integrated with one another.

It is noteworthy that discourse is not only about language. Foucault rather prefers to define it as the “group of statements that belong to a single system of formation” [18: p. 107]. This system goes beyond the idea of discourses as particularly linguistic systems and encompasses, among others, the material settings of discourses, their institutional sites.

4.1.2 Systems of Discursive Formation

With the notion of the discursive formation Foucault is able to describe the system of regularities and dispersion that rules a number of statements. He aims to answer the question: “How is it that one particular statement appeared rather than another?” [ibid.: p. 27] Therefore, it is necessary to “grasp the statement in the exact specificity of its occurrence, determine its conditions of existence, fix at least its limits, establish its correlations with other statements that may connect with it, and show what other forms of statements it excludes.” [ibid.: p. 28] Foucault’s analysis does not stop at the level of mere words or sentences as other linguistic approaches do. On the contrary, his argumentation stresses the importance of a holistic perspective which would, for instance, include an examination of (a) the different **types of discursive organizations**, (b) the **strategies** developed in a discursive formation, and (c) the essence of the notion ‘**statement**’ itself.

- (a) Following Foucault, the field of statements can be organized in various ways that regulate the production of the discourse.

The most common forms of organization are the forms of *succession*, i.e. the orderings of the statements (generalization, specification, spatial distribution, linear succession etc.), their dependence (hypothesis/verification, assertion/critique, general law/particular application etc.) and the rhetorical schemata they belong to (description, deduction, definition etc.).

The field also involves forms of *coexistence*: the field of presence (statements formulated elsewhere that are acclaimed, affirmed, negated, criticized, discussed or judged), the field of concomitance (statements formulated elsewhere that are used as analogical confirmation or general principle), the field of memory (statements formulated elsewhere that are no longer accepted).

Finally, the discursive statements are organized in a number of *procedures of intervention*. These include techniques of rewriting, methods of transcribing, modes of translating and means to increase the approximation of statements, to delimit their validity or to systematize them.

- (b) An examination of the strategies of discourses can for example look at *points of diffraction*, i.e. where two concepts or enunciations seem incompatible, where two incompatible statements appear in the form ‘either ... or’, where incompatible statements are systematized.
- (c) The often uncritically used notion of the ‘*statement*’ is described by Foucault as the “atom of discourse”. [18: p. 80] The statements or “*énoncés*” are not definable as propositions, sentences or speech acts. In a somewhat peculiar passage he rather thinks of them as oral or written formulations with a place and a date that enable the discursive rules and forms described above to become manifest. Statements make sense according to the rules they follow. The crucial point is, that Foucault turns away from a grammatically or lexically centered analysis to a more context sensible form of engaging with the discursive material.

A practical way of dealing with discourses from this Foucauldian perspective is proposed by [22] who outlined the following five steps:

1. recognizing a discourse as a corpus of regularly and systematically organized statements
2. identifying the rules of production
3. identifying the rules delimiting the sayable
4. identifying the rules creating spaces for new statements
5. identifying the rules ensuring that a practice is material and discursive at the same time (i.e. that the discourse is always connected to the settings and places where it is produced)

4.1.3 Archaeology/Genealogy

The term ‘archaeology’ has been coined by Foucault to name the procedure or method that describes the discourses as practices in their systems of formation and transformation. Archaeology explores the networks of statements in a material/institutional arrangement. Thereby, it focuses on the forms of organization, the strategies and regulations without being interpretive (archaeology

as ‘pure’ description) or anthropological (archaeology as concentration on statements, not authors). In relation to the analytical steps listed above [22] defined additional goals of an archaeological research à la Foucault:

1. chart the relation between the sayable and the visible – the relation between the statements and their contexts (e.g., the discourse of schooling includes the utterances in school, the written material, the pedagogical literature, the rules, the buildings, the blackboards etc.)
2. analyze the relation between the statements
3. formulate rules for the use of statements
4. analyze the positions between the subjects (subjects as subjective positions, e.g., the principal, the teacher, the pupil etc.)
5. describe ‘surfaces of emergence’ (the domain of a discourse, e.g., the school, the family, the university etc.)
6. describe ‘institutions’, which acquire authority and provide limits (e.g., the school as discursive institution can delimit the sayable or the range of possible activities)
7. describe the ‘forms of specification’ as systematic ways of understanding discursive objects (e.g., the concepts and terms of educational psychology that are used to refer to the behavior of schoolchildren)

The concept of ‘genealogy’ has often been described as being a counterpart to archaeology. Nevertheless, it resembles the former one in its essential methods, including the exploration of the discursive formations. But whereas archaeology provides us with a “snapshot, a slice through the discursive nexus” [18: p. 30] genealogy is more interested in the processual character of the discourse, its mobility that makes it “move to the rhythm of events”. [ibid.: p. 168]

4.2 First Steps Towards a Discourse Analysis

In relation to Foucault’s theory outlined above Wikipedia can be viewed as a discourse in the sense of a formation that regulates and structures the production of statements. The discursive practices, i.e. the collaborative editing processes, constitute knowledge that is permanently under scrutiny. These regular practices can be analyzed with the help of the five steps proposed by [22]: 1. defining a corpus, 2. identifying the rules of production, 3. identifying the rules delimiting the sayable, 4. identifying the rules creating the space for new statements, and 5. identifying the rules ensuring the material and discursive practice. As a start, the first three of them should be addressed in the present case acting on Foucault’s advice to regard discourse analysis as a “breaking-down of the material according to a number of assignable features whose correlations are then studied”. [18: p. 11] Hence, the level of analysis limits to the structure of the statements and the active types of connection that unite them. Moreover, the analysis attempts to be non-interpretive and non-anthropological. It is avoiding judgments and rejects investigations about the authors. Foucault argued that it is a futile attempt to analyze the “anonymous field” [ibid.: p. 122] of authors – an advantageous position in the case of Wikipedia with its considerable number of anonymous users.

4.2.1 Definition of a corpus

The definition of a coherent corpus starts with limiting the field of research to a reasonable amount of data. A section has to be selected from the enormous material of Wikipedia. For this reason, the discourse analysis will only be applied to the article [[Conspiracy theory]] and its page history over a period of four months (November 2005-February 2006). All the following examples are taken from the page history. They include the date, the author (underlined) and the comment. The last edit is cited first.

4.2.2 Rules of production

In order to shed light on the rules of production it is sensible to start with looking at one fundamental aspect of the Wiki technology: A Wikipedia article is almost completely open. Anyone, even anonymous users, can edit pages. These edits range from minor alterations to a complete rewriting of the material. The positions between the single authors (subjects in Foucault’s term) seem to be equal.¹ The organization of the discourse unfolds according to most of the regularities named by Foucault.

a) Succession:

All forms of succession can be detected in the writing processes of a Wikipedia article. There are forms of **generalization**, for instance, when too detailed descriptions or commentaries are removed

Example 1 – user ‘Tom harrison’ removes (rm) detailed description of former user

14:53, 9 March 2006 Tom harrison (rm Kennedy assassination commentary)

09:19, 9 March 2006 El magnifico (→Real conspiracies)

forms of **specification** when concepts are elaborated or material is added to existing sections

Example 2 – the user added material to the article

19:12, 30 November 2005 RK (→Types of conspiracy theories - Adding some of the most popular and enduring conspiracies.)

19:10, 30 November 2005 RK (Types of conspiracy theories. Bible conspiracy theories)

and also spatial **redistributions**, for example, when the textual material, sections, paragraphs or the word order of the article are rearranged.

Example 3

20:37, 15 December 2005 Zen-master (move one sentence back to pre-intro)

b) Coexistence:

In writing Wikipedia articles authors usually rely on statements made elsewhere. The **field of presence** consists therefore, on the one hand, of the alterations made by other authors, their arguments and justifications and, on the other hand, of all the

¹ A closer look at the Wikipedia reveals a power structure of nominated sysops, arbitration committees, a benevolent dictator etc. However, this can be neglected in the present analysis.

possible statements posted on talk pages, surveys, bulletin boards etc.

Example 4

02:24, 20 December 2005 Carbonite (this is the version I believe is best suited as the "working" version; there seems to be some agreement on the talk page that this is the case)

The rules and guidelines of the Wikipedia can be regarded as one **field of concomitance**. These are often brought into play during edit controversies. A second field is established by other Wikipedia language versions.

Example 5 – NPOV = Neutral Point of View

00:24, 14 January 2006 Nrcprm2026 (well, I tried; am kind of glad I didn't have to defend the idea; adding NPOV tag per Wikipedia talk:Consensus#http://en.wikipedia.org/wiki/Wikipedia_talk:Consensus)

Example 6

17:05, 11 December 2005 Zen-master (restore NPOV intro)

Example 7

05:59, 11 February 2006 SkeenaR (please assume good faith)

Example 8

19:58, 30 November 2005 Harald88 (replaced biased and overly restrictive definition by a neutral one that reflects the German example)

There is also a **field of memory** for every article. It is formed by the statements that once entered the discussion but were declared as wrong, inappropriate or not suitable. Authors sometimes make use of this field of memory in their argumentation.

Example 9 – rv = revert

20:39, 20 December 2005 Tom harrison (rv - see talk archive for reasons not to)

20:35, 20 December 2005 Zen-master (see talk for need to include all details)

c) Procedures of intervention

The examination of the edit history of the [[Conspiracy theory]] article shows that most of the procedures of intervention are used by the authors in the writing process, too. Apart from adding and deleting content they often **rewrite** passages.

Example 10

04:44, 1 December 2005 Peter McConaughy (Rewrote the preamble to be a super-combination of all aspects from all editors, as well as other dictionaries and encyclopedias.)

In addition, there are forms of **transcribing** and means to increase the **approximation** of statements.

Example 11

07:44, 9 December 2005 Zen-master ("often" --> "potentially", add clarity to last intro sentence)

Example 12

18:18, 11 December 2005 Drknexus m (Fixed Spelling & Merged First two paragraphs, swapping order so that the more accurate and clear definition was first, with the slightly unclear clarifying statement second rather than the other way.)

4.2.3 Rules of limitation

A second way of looking at the discursive processes in Wikipedia is the identification of rules, which delimit the sayable, define legitimate perspectives and fix the norms for the elaboration of concepts. The production of statements is always restricted by constraints that emerge in the discursive practice itself. To this end, an investigation of the levels and functions of conflict leads to the description of the norms prevalent in the writing process of a Wikipedia article.

According to Foucault, different **levels of dissension** can be determined. The examination of the page history of [[Conspiracy theory]] recorded:

a) Inadequation of objects:

In this case, one object – the assassination of J.F. Kennedy – is declared as not being part of the discursive field of the article. Therefore, the deletion is justified with regard to the boundaries of the accepted use of the concept ‘Conspiracy theory’.

Example 13

23:26, 30 November 2005 Harald88 (description became too restrictive again: JFK isn't commonly considered a straightforward event!)

b) Divergence of the modes of argumentation:

The users revert each other's contributions because of their different understandings of the form and content of a so-called ‘neutral’ statement. Moreover, they prefer two different modes of argumentation. Whereas ‘Zen-master’ uses the NPOV-tag to signal controversy, ‘Jayjg’ seems to prefer the discussion on the talk page.

Example 14

21:50, 30 November 2005 Jayjg (we've been through this before, Zen-master, you can't hold this article hostage forever with an NPOV tag. Your editing has passed the line into disruptive at this point.)

21:33, 30 November 2005 Zen-master (add {npov}, we disagree as to what is neutral here)

Example 15

18:47, 2 December 2005 Harald88 (rv: Tom, it's a logical fallacy to think that pejorative use can't be frequent because non-pejorative use is common!)

c) Incompatibility of concepts:

User ‘Jayjg’ criticizes the rather inconsequent word choice of the former editor. The alteration is reverted because the terms ‘theory’ and ‘allegation’ seem to be incompatible.

Example 16

22:17, 12 December 2005 Jayjg (well, if nothing else, a theory cannot be an "allegation", it must at least be a "theory", and various other attempts to fix the grammar)

d) Exclusion of options:

Both examples stress the importance of the barriers of the discourse. The sayable and writeable is structured by the definition of thresholds. Statements that violate these inherent rules form a group of dangerous elements that need to be deleted.

Example 17

08:22, 15 February 2006 [85.185.76.132](#) (WHY mysticism in general should be here as a conspirat theory?)

Example 18

19:50, 24 January 2006 [Tom harrison](#) (rv - rm spam link)

4.2.4 The discursive regime

The analysis of the edits during this four-month period shows that the production of the statements takes place in a field of intense debate. The article, even though it presents itself on the surface as a robust part of an encyclopedia, is never the “ideal, continuous, smooth text” [18: p. 155]. Rather it is a space of multiple dissensions. Almost no alteration of the text passes without comments. Every change generates following changes. However, despite its complexity the collaboration still seems to be productive. As Foucault noted, power/knowledge does not hinder the proliferation of statements but is instead the process itself. The discursive regimes, the effects of power are an essential part of the play of statements [see 25].

The way the speakable and writeable is structured can be tentatively examined by looking at a small number of edits made on November 27, 2005. With their help the regulatory power of the discursive regime, the working of the power/knowledge processes can be described. *Example 19* taken from the history of [[Conspiracy theory]] lists all the alterations made during that day (starting with the latest edit):

18:54, 25 November 2005 [Carbonite](#) (“commonly considered”; zen-master, stop reverting please)

18:43, 25 November 2005 [Zen-master](#) (if you have a problem with "ostensibly" please discuss on the talk page)

18:18, 25 November 2005 [SlimVirgin](#) (rv; Zen-master has violated 3RR with "dubious narrative genres" and "controversially", and has violated WP:NPA)

17:38, 25 November 2005 [Zen-master](#) (compromise, new plus old cleanups, watch out for the highly suspicious and coordinated POV pushing/bullying)

17:32, 25 November 2005 [Carbonite](#) (rv; looks like it's about time for a 3RR report)

17:29, 25 November 2005 [Zen-master](#) (old plus cleanups, see responses on talk, very interesting that the POV pushing bot-esque gang is monitoring this article so closely and so coordinatedly)

17:25, 25 November 2005 [Carbonite](#) (rv; this version is clearer and more NPOV)

17:24, 25 November 2005 [Zen-master](#) (removed "controversially" for now)

17:18, 25 November 2005 [SlimVirgin](#) (rv; zen, have you violated 3RR?)

17:15, 25 November 2005 [Zen-master](#) (rv, indeed I am alleging controversy)

17:09, 25 November 2005 [Jayjg](#) m (please stop injection your POV into the intro, zm. the only "controversy" here is the one you keep alleging.)

16:56, 25 November 2005 [Zen-master](#) (changed "believe and portrayed" to "ostensibly", i presume we agree to disagree otherwise)

16:32, 25 November 2005 [SlimVirgin](#) (rv to Tom harrison: zen-master's additions are POV and inappropriate)

04:01, 25 November 2005 [Zen-master](#) (the saga continues)

02:44, 25 November 2005 [Tom harrison](#) (words fail me - ameliorating intro? benignifying text? enhancing lucidity?)

02:24, 25 November 2005 [Zen-master](#) (fix "term")

02:22, 25 November 2005 [Tom harrison](#) (→See also - added link to 'September 11, 2001 researchers')

02:21, 25 November 2005 [Zen-master](#) (old plus new together, early october version was better in a lot of ways)

00:42, 25 November 2005 [Jayjg](#) m (no, let's keep it simple)

00:37, 25 November 2005 [Zen-master](#) (big intro re-org and clean ups)

These altogether 20 edits are made by five users who discuss the proper definition of the term ‘Conspiracy theory’. The controversy starts because of an edit by ‘Zen-master’ who reorganizes the introductory sequence by inserting older material from previous, deleted versions. In Foucault’s notions, this single event generates a series of statements, which produce regularities that reflect on the possible conditions of existence of future statements. Five minutes after the initial event the user ‘Jayjg’ reverts the edit and reinstalls the former introduction. At first, it seems as if ‘Zen-master’ accepts this edit as a proposal to put “old plus new together” and therefore to approximate the versions.

However, the minor change of the word ‘term’ into ‘phrase’ (edit 02:24) is contested by ‘Tom harrison’ who criticizes the alteration as not contributing to clarity and lucidity. Thus, the user introduces a rule from the Wikipedia as field of concomitance to which the authors should adhere. It functions as a barrier or threshold to master the growing complexity of the article. ‘Zen-master’ is asked to recognize the rule as part of the doctrine of the ensemble of discourse. Nevertheless, the user does not accept the authority of ‘Tom harrison’ and comments: ‘the saga continues’ (edit 04:01). Moreover, the user introduces the highly critical term ‘controversial’ to the definition of ‘Conspiracy theory’. As a consequence, the users ‘SlimVirgin’ and ‘Jayjg’ revert the edits as inappropriate and biased with reference to the ‘point of view-rule’ (POV). A second time the user is asked to accept the rules of conformity. The structuring force of the power/knowledge relation is working: The statements that supposedly do not fit into the argumentation of the article are deleted. They cannot be part of the text because they oppose the discursive boundaries of the accepted use of the concept ‘Conspiracy theory’. The discursive string of the article seems to have produced the constraint or

governing statement respectively that a ‘Conspiracy theory’ cannot be called ‘controversial’.

Additionally, the adherence to the discursive doctrine does not only involve the written or spoken statements but the speaker/writer herself. The subject is also part of the power/knowledge relation; it forms a nod in the net-like organization of power. Thus, the rules of exclusion and the rejection mechanisms do affect the subject, too. “Doctrine effects a dual subjection, that of speaking subjects to discourse, and that of discourse to the group, at least virtually, of speakers” [19: p. 226]. In the given example, the user ‘Zen master’ is blocked from further activities because of violating the 3RR rule. The first reference to this quasi-official policy of Wikipedia that authors must not perform more than three reversions on a single article within one day is made by the user ‘SlimVirgin’ (edit 17:18). The rule seems to have a strong regulatory power because ‘Zen-master’ deletes the term ‘controversial’. This can be regarded as an act of self-policing – the user offers a consensual version without the critical expression. However, the edit war cannot be stopped. A second time the 3RR rule is brought into play and finally the user is blocked. The notice published on the talk page of ‘Zen-master’ says: ‘You have been temporarily blocked for violation of the three-revert rule. Please feel free to return after the block expires, but also please make an effort to discuss your changes further in the future.’ The discursive regime punishes the user for violating the established boundaries of the article: the statements are deleted and the user is (at least temporarily) excluded from the discursive community.

This brief analysis of a small part of the editing processes of the article [[Conspiracy theory]] is almost like a window to the discursive universe of Wikipedia. However, what becomes obvious is, that the collaboration is not as chaotic as it may seem at first sight. In contrast, the effects of power/knowledge do follow ordered procedures for the production, regulation, distribution and circulation of statements. The example shows the structured emergence of single events which in turn can form series that generate regularities which form possible conditions of existence to delimit the sayable.

5. SUMMARY AND CONCLUSION

Wikipedia is primarily characterized as being one of the largest if not the largest encyclopedia in the history of reference books. However, a closer examination of the communicative situation of Wikipedia produced some evidence that Wikis foster a growing number of different forms of computer-mediated communication, too.

The principle goal of the paper has been to explore ways of analyzing this material and in doing so to shed light on the structure of the Wikipedia, its social organization as well as the complex processes of writing and editing articles.

First, the paper differentiated between the various ways of communicating in Wikipedia. The compiled list includes 1. the talk pages as most common forms of communication, 2. the page history with the comments of the editors, 3. the village pump as central platform for non-article related discussions, 4. the bulletin board of the community where news and announcements are published, and 5. the numerous forms of comment pages where authors can post a request or conduct a survey. In addition, there are 6. several mailing lists (ca. 50) and 7. about 80 IRC channels.

In a second step, the paper reviewed the current research on Wiki discourses. The first who intensely engaged with the textual material were Viégas, Wattenberg and Dave [7] whose *history flow*-software visualized the contributions of the authors. On that basis, they were able to systematize patterns of vandalism. Another view on Wikis has been established by Emigh and Herring [8] who showed that Wiki articles resemble traditional print versions in terms of their language formality. Apart from that, most studies focused either on the communicative structure [10], the positive function of conflict [11], or the processes of quality assurance [17].

Third, the paper made use of the discourse theory of Michel Foucault to examine not only the grammar or vocabulary but to develop a holistic approach towards the analysis of Wikipedia. To this end, five steps [see 22] of dealing with Foucault’s notion of discourse as a practice that regulates and structures the production of statements have been outlined: 1. defining a corpus, 2. identifying the rules of production, 3. identifying the rules delimiting the sayable, 4. identifying the rules creating the space for new statements, and 5. identifying the rules ensuring the material and discursive practice.

Finally, the paper presented an exemplary analysis of a small section of the discursive material that addressed the first three steps. Thus, a corpus consisting of the article [[Conspiracy theory]] and its page history was created. Then, this corpus was examined concerning its rules of production and limitation.

As a result, the example showed that the discourse unfolds according to most of the regularities listed by Foucault. This lends considerable support to the assumption that a Foucauldian discourse analysis can be fruitfully applied to the Wikipedia as a whole. Not only does his concept of discourse address the various forms of Wiki communication it also encompasses the technical and social structure of the Wikipedia. A full-scale analysis would therefore resemble one of Foucault’s impressive studies, for example on sexuality or punishment, where he described an abundance of statements in their complex relations to other statements and contexts.

In order to conduct such an analysis, several steps remain to be done. First, the complete Wikipedia needs to be conceptualized in terms of Foucault’s theory. The question is, if it is possible to build-up coherent corpora of documents. Second, the methods of discourse analysis must be adapted to the context. This includes the establishment of principles of choice and the definition of levels of analysis. Third, it is necessary to look at other parts of the Wikipedia in order to enlarge the material basis for a qualitative discourse analysis and therefore to verify the findings of the present initial and tentative examination.

6. REFERENCES

- [1] Mejías, U.A. *Online Discourse: Past, Present, Future*. Presentation of the 16th Annual Instructional Technology Institute, Utah State University, September 01, 2004. Online: http://ideant.typepad.com/ideant/files/mejias_online_discourse.pdf (last retrieval April 01, 2006).
- [2] Herring, S.C. Slouching toward the ordinary: current trends in computer-mediated communication. In *New Media & Society*, 6(1), 26-36, 2004.

- [3] Voss, J. Measuring Wikipedia. In *Proceedings of the ISSI 2005* (Stockholm, Sweden, July 24-28, 2005). Online: <http://eprints.rclis.org/archive/00003610/01/MeasuringWikipedia2005.pdf> (last retrieval April 01, 2006).
- [4] Lih, A. Wikipedia as Participatory Journalism: Reliable Sources, Matrix for Evaluating Collaborative Media as a News Resource. In *Proceedings of the 5th International Symposium on Online Journalism* (Austin, TX, April 16-17, 2004) Online: <http://jmssc.hku.hk/faculty/alih/publications/utaustin-2004-wikipedia-rc2.pdf> (last retrieval April 01, 2006).
- [5] Brändle, A. Too Many Cooks Don't Spoil the Broth. In *Proceedings of the Wikimania '05* (Frankfurt am Main, Germany, August 4-7, 2005).
- [6] Rafaeli, S., Hayat, T. and Ariel, Y. Wikipedians' Sense of Community, Motivations, and Knowledge Building. A Cross-Cultural Study. In *Proceedings of the Wikimania '05* (Frankfurt am Main, Germany, August 4-7, 2005).
- [7] Viégas, F., Wattenberg, M. and Dave, K. Studying Cooperation and Conflict Between Authors with History Flow Visualization. In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI'04)* (Vienna, Austria, April 24-29, 2004) Online: http://alumni.media.mit.edu/~fviegas/papers/history_flow.pdf (last retrieval April 01, 2006).
- [8] Emigh, W. and Herring, S. C. Collaborative Authoring on the Web. A Genre Analysis of Online Encyclopedias. In *Proceedings of the 38th Hawaii International Conference on System Sciences (HICSS'05)* (Waikoloa, Hawaii, January 3-6, 2005) Online: <http://csdl2.computer.org/comp/proceedings/hicss/2005/2268/04/22680099a.pdf> (last retrieval April 01, 2006).
- [9] Giddens, A. *The Constitution of Society. Outline of a Theory of Structuration*. Cambridge University Press, Cambridge, UK, 1984.
- [10] Resnick P. et al. Beyond Threaded Conversation. In *Proceedings of the SIGCHI conference on Human factors in computing systems (HCI'05)* (Portland, OR, April 2-7, 2005) Online: http://www-personal.si.umich.edu/~shakmatt/CHI2005/beyond_threaded_abstract.pdf (last retrieval April 01, 2006).
- [11] Shah, S. Productive Controversy. In *Proceedings of the Wikimania '05* (Frankfurt am Main, Germany, August 4-7, 2005).
- [12] Lawler, C. Wikipedia as a Learning Community. Master Thesis, University of Manchester, Manchester, UK, 2005.
- [13] Reagle, J. M. *A Case of Mutual Aid: Wikipedia, Politeness, and Perspective Taking*. Online: <http://reagle.org/joseph/2004/agree/wikip-agree.html> (last retrieval April 01, 2006).
- [14] Aronsson, L. Operation of a Large Scale, General Purpose Wiki Website. In Carvalho, J.A., Hübler, A.C. and Baptista, A.A. (eds.) *Elpub 2002. Technology Interactions. Proceedings of the 6th International ICCS/IFIP Conference on Electronic Publishing* (Karlovy Vary, Czech Republic, November 6-8, 2002). Verlag für Wissenschaft und Forschung, Berlin, Germany, 2002, 27-37.
- [15] Ebersbach, A. and Glaser, M. Towards Emancipatory Use of a Medium. The Wiki. In *International Journal of Information Ethics*, 2, 11, 1-9.
- [16] Froomkin, A.M. Habermas@Discourse.net: Toward a Critical Theory of Cyberspace. In *Harvard Law Review*, 116, 3, 751-871.
- [17] Stvilia, B. et al. *Information Quality Discussion in Wikipedia*. Technical Report ISRN UIUCLIS--2005/2+CSCW. Online: <http://www.isrl.uiuc.edu/~stvilia/papers/qualWiki.pdf> (last retrieval April 01, 2006).
- [18] Foucault, M. *The Archaeology of Knowledge*. Transl. by A. M. Sheridan Smith. Pantheon Books, New York, NY, 1972.
- [19] Foucault, M. *The Discourse on Language*. Transl. by A. M. Sheridan Smith. Pantheon Books, New York, NY, 1972a (incl. in *The Archeology of Knowledge*).
- [20] Mills, S. *Michel Foucault*. Routledge, London, UK and New York, NY, 2003.
- [21] Foucault, M. Interview. In Bouchard, D. (ed.) *Michel Foucault: Language, Counter-memory, Practice. Selected Essays and Interviews*. Blackwell, Oxford, UK, 1977.
- [22] Kendall, G. and Wickham, G. *Using Foucault's Methods*. Sage, London, UK, Thousand Oaks, CA and New Delhi, India, 1999.
- [23] Foucault, M. Prison Talk. In Gordon, C. (ed.) *Michel Foucault. Power/Knowledge. Selected Interviews and Other Writings 1972-1977*. Prentice Hall, New York, NY, et al., 1980.
- [24] Foucault, M. Two Lectures. In Gordon, C. (ed.) *Michel Foucault. Power/Knowledge. Selected Interviews and Other Writings. 1972-1977*. Prentice Hall, New York, NY, et al., 1980.