ABSTRACT: The public computer network Internet is used intensively in research and scholarly communication. An important tool for the immediate exchange of research ideas and scholarly arguments is e-mail based discussion lists. E-mail messages sent to such a discussion list are distributed automatically to hundreds or even thousands of recipients worldwide.

In this paper we discuss the role of e-mail based discussion lists in scholarly communication and analyze the structure and the development of CERRO-L, a Regional Economics oriented discussion list with focus on regional economic development in Central Europe. We will analyze the structure of subscribers and contributors to CERRO-L, and the content of their messages, the major topics, themes, and places (countries and regions) discussed. With this analysis we want to be able to answer questions like the following: "To what extent can e-mail discussion lists tie together research communities?" "How stable are such communities?" "Can e-mail discussion lists help overcome some of the limitations (spatial, cultural, language) of more traditional means of scholarly discussion?" "How focused are discussions on e-mail based discussion lists?" "What does electronic communication portend about scholarly exchanges and communities?" "What are the impacts of such communication on the formation, structure, and dissemination of knowledge itself?"
1. Introduction

Although the growth of the global computer-network Internet is usually measured in technical terms, like the number of hosts it connects, this growth is actually driven by the imagination and innovativeness of the Internet-users. The times are long gone when an individual could have an overview over "what is available on the Internet". The estimated 50 - 100 million users worldwide use "the net" as infrastructure in any way that is technically feasible. They present their personal interests in thousands of homepages on the World-Wide-Web, communicate via e-mail, discuss business and hobbies on newsgroups and discussion lists, share their views and their electronic collectibles, search for advice, download and upload software, and communicate in many other ways on a daily basis and on a global scale.

By subscribing to discussion lists, reading newsgroups, tuning into IRC-channels, etc. the Internet users join together with others in forming "virtual communities" (Rheingold, 1993). Such virtual communities are the fibres in the social tissue that forms the Internet culture. "People in virtual communities use words on screens to exchange pleasantries and argue, engage in intellectual discourse, conduct commerce, exchange knowledge, share emotional support, make plans, brainstorm, gossip, feud, fall in love, find friends and lose them, play games, flirt, create a little high art and a lot of idle talk. People in virtual communities do just about everything people do in real life" (Rheingold, 1993, p. 3), but they do this in the virtual world of cyberspace, and are not constrained by the need to overcome physical distance.

Typically, the members of a virtual community come from different parts of the world, speak different languages, and have never met face to face. In joining a virtual community, they bring in their own personality, their social and cultural background, their views and believes, many of which have been influenced by their membership in other communities, be they virtual or real. In this way each individual links together communities in the real and in the virtual world, and transfers ideas – in the broadest sense of the word – between different communities. Because of this transfer of ideas and because individuals drop in and out of virtual communities, these communities evolve over time. With discussion moving from one topic to another, the emphasis of the virtual community may gradually shift, some people may gain importance while others may lose, communication that was once considered inappropriate may become acceptable and vice versa.

Although the Internet today is a commercial rather than an academic network, it has high and still growing popularity in the academic world. This is small wonder when we take
into account that on the one hand the Internet's ultimate function is communication, and that on the other hand “communication is the only general scientific behavior” (Griffith 1989, p. 600). Also, the Internet developed in the academic world for about two decades before it was allowed to spread out to businesses. For these reasons, many of the virtual communities on the Internet today are academic in nature.

In this paper we will deal with one academic virtual community (AVC) on the Internet. This community has an explicit regional economics orientation (“regional economic development in Central Europe”) and is centered around the electronic discussion list “CERRO-L” (see also Graute, 1995). We will look at the structure of this academic virtual community and its development over a period of 3½ years. We will analyse the membership of the electronic discussion list (EDL) and the content of the discussion that was distributed over it.

Before moving to the empirical analysis, in the following section (section 2) we will take a closer look at (academic) virtual communities and the spatial implications they may have. We will start the empirical analysis (section 3.1) by describing the original intentions of the virtual community under investigation, the main administrative steps that were taken as well as the development of the focus of the community, the regional economic development of Central Europe. All this is necessary in order to understand the evolution of the academic virtual community over time. Section 3.2 will give a detailed analysis of the structure of the AVC and of the content of its discussion. The paper will end with a concluding section.

2. Academic virtual communities, scholarly communication, and the geography of science

According to some definitions, "community" is defined by spatial proximity. Johnston (1981, p. 46), for example, defines "community" as a "spatially delimited set of interacting face-to-face groups". And Johnston, Gregory and Smith (1994, p. 80) call it "a social network of interacting individuals, usually concentrated into a defined territory".

These definitions not only show that the term "virtual community" – with the understanding of it being a group of people who are connected via a computer network – is somewhat contradictory, they also illustrate the radical chance that resulted from the use of this technology.

Traditional forms of scholarly communication can tie together a small group of people
living and working in close proximity. However, they extend only poorly over space and over a larger number of people (Maier, Wildberger, 1993; Wildberger, 1992). Either highly formalized forms of communication (journal articles, monographs, etc.) or forms that involve major administrative overhead (conferences) are used in order to reach larger numbers of scholars beyond the regional scale. The standards that have been developed over time in order to make this form of communication work, result in major time lags as well as severe restrictions of communication. Parallel sessions at conferences, for example, break down the number of participants into groups of operational size, but force the participants to decide for one and against all other sessions that run parallel. Moreover, scholarly discussion at conferences needs to be squeezed into short sequences of statements that allow only few participants to contribute. When participants do not obey to these rules and standards, e.g., when two or three talk at the same time, the whole communication process is in jeopardy.

We have become so used to such standards and rules, that we sometimes see them as integral part of scholarly communication. However, Internet-based forms of scholarly communication show that many of them are only necessary because of the limitations of the traditional forms of communication. With more and more academic work becoming available over the World-Wide-Web, for example, the linear text of a printed article appears quite old-fashioned. Standard references turn out to be weak substitutes for the hyperlinks of a web-based text. With more and more scholarly material appearing on the World-Wide-Web, the medium will gain reputation and may eventually bypass the whole academic publishing industry. "It is ironic that by subscribing to journals, libraries in effect buy back the scholarship that university faculty members have created and given away" (Bennett and Matheson, 1992, p.B2). Rapidly rising prices of scientific books and journals, and long publishing delays will add to this transition.

One of the main advantages of Internet-based scholarly communication is the weak distance friction of the underlying infrastructure. Once a document is put on the web, it can be accessed from anywhere on the network at the same cost. By use of hyperlinks to other material on the web, the document can be woven into and become part of a truly global knowledge base. Similarly, e-mail-based discussion lists typically have subscribers from different parts of the world. One’s arguments are therefore carried to other parts of the world. They may enter a traditional face to face conversations thousands of miles away, be transferred to other discussion lists and thus spread beyond control of the original sender.

These media have the potential of overcoming distance friction in scholarly discussion. They may move us a big step toward truly global academic communities.
3. CERRO: an academic virtual community in regional science

In this section we will take a closer look at one of the many hundred academically oriented electronic discussion lists that are available on the Internet. CERRO-L, the list under consideration, operates out of Vienna, Austria, and is devoted to the discussion of regional economic development in Central Europe.

The electronic discussion list CERRO-L is just one of the electronic tools available to the academic virtual community under consideration. Connected to CERRO-L is a Gopher-based archive of electronic documents and data concerning Central Europe (CERRO archive). This archive provides access to the log-files of the EDL, so that users can browse through earlier exchange of ideas and arguments, as well as to a WAIS-index for full-text search of those logs. In our terminology, these two elements (“CERRO-L”, “CERRO-archive”) constitute “CERRO”\(^1\). Loosely coupled to it is information concerning european sections of the Regional Science Association (most importantly the electronic member directory) and links to other electronic information systems in the discipline. This RSA-related information system and CERRO together form the “Vienna Information Service in Regional Science” (VISRS; see figure 1).

The option to link services together is an attractive feature of Internet-based information systems and one of the reasons for their success. Such links allow for transfer from one virtual community to another and actually blur the borders between them. This is attractive in “real life”, but complicates the analysis of virtual communities. In this paper we take a pragmatic approach and will will focus the analysis in the following subsection on the EDL “CERRO-L”. This is justified because “CERRO-L” clearly constitutes the core of the AVC. Also, it provides the information that is necessary for the analysis.

3.1. CERRO and CERRO-L: motivation, management, and a dynamic environment

CERRO is the result of two important developments in the late 1980s and early 1990: the collapse of the communist block in Central Europe (Maier, Mašek, 1992) and the introduction of Internet in European Universities (Maier, Wildberger, 1993). These developments have stimulated discussions between three people from three different parts of the world: (1) Gunther Maier, Vienna University of Economics and Business Administration, (2) Edward M. Bergman, The University of North Carolina at Chapel Hill, and (3) Martin Mašek, Slovak Academy of Sciences, Bratislava. Among these three individuals the idea was
born to use the capabilities of the Internet for discussing the regional-economic development of the reemerging region of Central Europe. The initial element of this endeavor was the electronic discussion list CERRO-L.

The basic idea of an EDL is relatively simple: e-mail based information should be distributed to a group of e-mail addresses instead of just one. This mechanism can be implemented in many different ways. The simplest version is provided by the nickname-function of many e-mail programs. It allows one user to send e-mails to a group of
recipients. However, the functionality is available only for the one participant who defines the nickname and cannot be made available to others. A more sophisticated version is based upon the alias-function of the UNIX operating system. It allows one to define a single e-mail address as an alias for a group of addresses so that when a message is sent to the alias it is forwarded to each recipient in the group. Therefore, they are often called "mail-exploder". Such mail-exploders are easy to set up and very useful for small and stable groups. Since recipients’ e-mail addresses need to be entered and removed to and from the alias by hand, they are difficult to manage for large or volatile groups of participants. The most sophisticated way of managing an EDL is via a specialized EDL-administration program. Such programs allow users to subscribe to and unsubscribe from the discussion list without the intervention of a list-administrator. This reduces the administrative overhead dramatically. EDL-administration programs allow for different parameter settings both for the EDL as a whole and for individual subscribers. Moreover, many of them provide statistics and archival functions.

CERRO-L started as a UNIX-based mail-exploder in the fall of 1991. By mid April the group of participants to CERRO-L has grown so large that administering it in this way had turned into a tedious task. Therefore, on April 30, 1992 CERRO-L was moved to an installation of the program "listserv" on one of the computers of the University of Vienna. In addition to distributing messages this program also archives them in log-files. This function provided the information for most of the analysis in this paper.

The evolution of an EDL like CERRO-L is influenced by a number of factors. The most important ones are:

1. The software used for managing the discussion list. Some programs or program versions are more sophisticated in detecting anomalies in messages. This may prevent unnecessary traffic and reduce the noise on the list. With most programs the management of the list can set parameters that influence the operation of the software. Some of these parameters have major implications for the way how the list operates.

2. The development of the topic of the discussion list. Certain economic and political events in Central Europe (e.g., the breakup of Czechoslovakia) may have stimulated discussion of these events on CERRO-L.

3. The policy of the list manager. Discussion lists differ tremendously in how much the list manager gets involved. This and the attitudes of the subscribers determines what is considered to be acceptable on the list and what is not.

For most list-administrators managing an EDL is a process of learning-by-doing. The
management of CERRO-L is no exception. The installation of "listserv" that manages CERRO-L is quite sophisticated in detecting messages that are generated by other mail programs as reaction to faulty or non-existing e-mails addresses. Such messages may cause a discussion list to explode. Despite the sophistication of the software, CERRO-L exploded in April 1995. In order to avoid this in the future, the list management changed the parameters of the list such that messages were accepted only from registered e-mail addresses, i.e., the e-mail address under which a subscriber is registered. This policy made it more difficult to distribute messages through CERRO-L and most likely caused a drop in the activity. Another problem CERRO-L experienced was that many subscribers accidentally sent replies to the whole list rather than individual other subscribers. Not only did this cause unnecessary network traffic, but also led to messages from other subscribers who expressed anger with such waste of bandwidth. In spring 1995 the list management changed the parameters of the software in such a way that a reply to a message that is distributed through CERRO-L by default is addressed to the original sender rather than the list. This removed the problem but may have further reduced the number of distributed messages.

In December 1993 the “listserv” at the University of Vienna was upgraded from version 1.7 to version 1.8. This new version of the program was able to detect “spams”, messages that are sent to a number of discussion lists at the same time. When the program detects a “spam” it does not send it out, but mails it to the list manager for inspection. In the past such “spams” have generated heated discussions on CERRO-L about misuse of the discussion list and whether a certain message is of potential value to subscribers or not. Since the new version of the software has detected an estimated 20-25 “spams” to CERRO-L since its installation. It has saved the list quite a number of unproductive messages, but may have led to a further drop in the measured activity.

3.2. Empirical analysis

The development of an EDL can be analyzed in many different ways. Information can be derived through (1) interviewing subscribers of the EDL, and (2) observing the activities on the EDL over time. While the first approach has predominantly been used by Brunn et al. (1996) in their analysis of GEOGRAPH, we will use the second approach.

Again, the activities of an EDL can be looked at from different perspectives. We can ask questions like:

* who is subscribed to the EDL?
* who sends messages?
* with what content?

We will turn to all three questions in this section.

### 3.2.1. Subscription

In the first step of our empirical analysis we will concentrate on the subscribers to CERRO-L. Unfortunately, LISTSERV keeps only a current list of subscribers and does not archive information about when participants subscribed to or unsubscribed from the list. Therefore, we have to base our analysis of subscription on the lists of subscribers that we have requested from LISTSERV in irregular intervals and stored. With this information we can cover a period of 34 months between September 1992 and June 1995. No information is available for March to May and July to October 1993, and May, June 1994. For all other months we have at least one subscriber list available. When we have more than one list, we use the first one in the month.

Two problems arise with subscriber lists: (1) e-mail addresses change over time so that in some cases one and the same person is registered under two or more e-mail addresses over the observation period; (2) many users are subscribed from bitnet-nodes that do not provide information about the country where the host is located. We did not try to solve the first problem systematically. Only in a few cases, where we were aware of the change we made the appropriate corrections. As far as the second problem is concerned, we tried to locate those hosts in Bitnet’s nodelist-file and added the correct country-code\(^6\). This worked in most cases. Only 202 of 10,371 e-mail addresses could not be assigned to a country. For those we used the "UND" (=undefined) category.

The number of subscribers to CERRO-L grew from 133 in September 1992 to 670 in June 1995\(^7\) (see figure 2). At the same time the number of countries (categories in the US) of subscribers increased from 21 to 45. Throughout the whole observation period, by far the largest group of subscribers belong to the "EDU"-category; i.e. US academic institutions.

In order to get a clearer picture about the spatial origin of subscribers, we grouped them into the mutually exclusive categories shown in table 1. Austria and Canada are treated separately because they are the only non-US countries that reach average shares above 5%.

When we look at the shares of these categories (figure 3) we see a surprising stability in the composition of the EDL. Despite the rapid growth in the number of subscribers to CERRO-L, the shares change only in a narrow range. Table 2 lists the minimum and
maximum shares and their weighted average.

Throughout the observation period the EDL is clearly dominated by US subscribers. This is no surprise for the beginning of the observation period, when the Internet was still heavily dominated by the US. However, despite the diffusion of Internet to other countries the share of US subscribers to CERRO-L tends to increase through the observation period rather than decrease⁸. The second largest group of subscribers are those from Western Europe, whose share is increasing slightly.

Only on third place are subscribers from an area that is the focus of the EDL;
subscribers from East-central Europe (ECE). The most developed post-socialist countries contribute around 10 percent of subscribers to CERRO-L. This is much higher than the share of subscribers from Central and Eastern Europe (CEE) whose weighted average is only 1.13 percent. Austria, the country where the EDL is located, reaches an average share of 7.81 percent. This is surprisingly high, particularly in comparison with a country like Germany, whose average share is below 3%. It would be interesting to know whether this high share is related to the location of the EDL. In any case, the shares of the subscribers from the focus area of CERRO-L (AT, CEE, ECE) all tend to decrease over the observation period. Despite the fact that CERRO-L is not US-oriented and despite the diffusion of Internet technology, the subscription of CERRO-L is becoming more US-oriented during the observation period.

Another important indicator concerning the subscription to CERRO-L is the turnover of subscribers. This is directly related to the average duration of subscription. The same total

**Fig. 3:** Subscribers by month: share of regions

**Table 2:** Minimum, maximum and average shares

<table>
<thead>
<tr>
<th>Category</th>
<th>Min.</th>
<th>Max.</th>
<th>Av.</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>49.43</td>
<td>58.70</td>
<td>56.15</td>
</tr>
<tr>
<td>WE</td>
<td>13.93</td>
<td>17.98</td>
<td>15.91</td>
</tr>
<tr>
<td>ECE</td>
<td>7.03</td>
<td>12.70</td>
<td>8.05</td>
</tr>
<tr>
<td>CEE</td>
<td>0.73</td>
<td>2.26</td>
<td>1.13</td>
</tr>
<tr>
<td>AT</td>
<td>6.52</td>
<td>9.27</td>
<td>7.81</td>
</tr>
<tr>
<td>CA</td>
<td>4.59</td>
<td>6.74</td>
<td>5.45</td>
</tr>
<tr>
<td>OTH</td>
<td>2.46</td>
<td>4.35</td>
<td>3.56</td>
</tr>
<tr>
<td>UND</td>
<td>0.45</td>
<td>4.91</td>
<td>1.95</td>
</tr>
</tbody>
</table>
numbers of subscribers can result from many individuals subscribing and unsubscribing, or from fewer individuals staying with the EDL for longer time periods. This results in different levels of stability of the EDL and its topics.

We can find 1,242 different e-mail addresses throughout the observation period. When we assume that each address corresponds to one individual, we see that 1,242 individuals have been in contact with CERRO-L during this 34 month period, 670 of whom (54%) are still subscibed at the end of the period. Of the 133 individuals in the first month of observation, 55 remain subscibed throughout the whole period. According to our dataset, the average length of subscibption is 9.09 months. This figure is not corrected for the censoring both at the beginning and at the end of the observation period. So, it underestimates the true average subscibption period.

3.2.2. Activity on the list

In total 1,893 messages were distributed through CERRO-L in the period between May 1, 1992 and October 31, 1995. The analysis of the sender and the content of the messages can be based upon the log-files that LISTSERV generates automatically. Because of this archiving feature, we have the full set of messages available.

Before we analyze who contributed what to the discussion list, let us look at the activity over time. In average, 1.48 messages per day (45.07 per month) were sent out via CERRO-L. But, as can be seen from figure 4, these messages appeared very unevenly during the observation period. The lowest level of activity with just 9 messages was in July 1993. Seven months later, in February 1994, the number of messages passed the 100 mark for the first time. The highest level of activity could be observed between October 1994 and January 1995 (between 117 and 127 messages/month). Thus far, January 1995 was the month with the highest level of activity. After January 1995 the level of activity drops of markedly. It hardly reaches 1/3 of the activity that we found before. Most likely, this has to do with the above mentioned change of the parameters of the list.

Before this structural break, the number of messages increased almost with the same rate as the number of subscribers. While the subscribers increase by 5.02% per month, the number of messages increases in average by 4.97% in this period. Beginning with February 1995, the number of messages decreased by about 1% per month, leading to an overall growth rate of only 2.20%.
3.2.3. Contributors to discussion

The 1893 messages that were distributed via CERRO-L in the observation period originate from 659 different e-mail addresses. The most active contributors to CERRO-L can be seen in figure 5. By far the most messages (120) came from the list-owner’s e-mail address. The second largest number of contributions (90) originates from the address that was used in spring 1993 for distributing the "Austrian News Service". On the other side of the spectrum, 397 subscribers sent only one message throughout the observation period, and 107 subscribers two messages.

Although the two most active contributors are from Austria, the largest number of contributions to the list come from the US. When we use the groups of countries shown in table 1, more than half the contributions originate from the USA (see table 3). Of course, this is mainly the result of the high share of US subscribers to CERRO-L as we have discussed it in the previous section. However, the groups also differ in terms of their activity. When we divide the share of messages in table 3 by the corresponding share of subscribers, we get an indicator for the relative chance that a subscriber from this group of countries will send a
message to the list (table 3, "activity"). As we see, subscribers from Central and Eastern Europe are not only rare among the subscribers, those who are subscribed are also much less active than other subscribers.

In this raw form, this indicator is biased by the fact that administrative messages originate from Austria as well as by the messages from Austrian News Service. Consequently, the indicator for the activity out of Austria is by far the largest. We correct for this by subtracting the messages from "maier@wu-wien.ac.at" and from the account we used for Austrian News Service. This gives the indicator "activity corrected" in table 3. With this

Table 3: Sent messages and activity by group of countries

<table>
<thead>
<tr>
<th>Country</th>
<th>number of messages</th>
<th>share of messages</th>
<th>activity</th>
<th>activity corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>989</td>
<td>52.25</td>
<td>0.93</td>
<td>1.05</td>
</tr>
<tr>
<td>WE</td>
<td>226</td>
<td>11.94</td>
<td>0.75</td>
<td>0.84</td>
</tr>
<tr>
<td>ECE</td>
<td>94</td>
<td>4.97</td>
<td>0.62</td>
<td>0.69</td>
</tr>
<tr>
<td>CEE</td>
<td>8</td>
<td>0.42</td>
<td>0.37</td>
<td>0.42</td>
</tr>
<tr>
<td>AT</td>
<td>376</td>
<td>19.86</td>
<td>2.54</td>
<td>1.26</td>
</tr>
<tr>
<td>CA</td>
<td>136</td>
<td>7.18</td>
<td>1.32</td>
<td>1.48</td>
</tr>
<tr>
<td>OTH</td>
<td>32</td>
<td>1.69</td>
<td>0.48</td>
<td>0.53</td>
</tr>
<tr>
<td>UND</td>
<td>32</td>
<td>1.69</td>
<td>0.87</td>
<td>0.98</td>
</tr>
</tbody>
</table>
correction the figures for all countries but Austria increase, such that the US-subscribers are now slightly more active in sending messages than the average. However, the figure for Austria is still the second highest (after Canada). It seems that regular subscribers from Austria contribute more actively to the discussion than those from other countries, especially those from their neighboring countries to the east.

3.2.4. Content of discussion

To get a handle on the type of discussion on CERRO-L, we categorized the content of each message according to two criteria: (1) what topics were discussed, and (2) what countries or groups of countries were mentioned. Each message was assigned at least one topic-code (these codes can be found in the appendix). Country-codes were assigned only where appropriate.

All together, the 1893 messages were assigned 2,701 topic codes and 3,074 country codes. This implies an average number of 1.62 topic codes and 1.43 country codes per message.

Topics

As far as topic is concerned, "information, communication and media" (f-) is the most important category. More than half the messages (50.08%) are assigned at least one topic code from this category. Responsible for this result is just one topic, "request for information" (f-10), to which almost 47% of messages are assigned. This shows clearly the importance of CERRO-L as a source of information – at least in the eyes of those who request information. Obviously, they see the EDL as an information pool that can be queried whenever necessary.

This result is underlined by the fact that the second largest category is "scholarly materials and resources" (e-). More than 20 percent of the messages contain such type of information. Of particular importance in this category are messages about "new books, publications, newsletters" (e-6, 4.81%), "meetings, conferences & seminars" (e-7, 5.71%) and "new listservs" (e-10, 2.69%). This type of messages corresponds nicely with the scholarly orientation of CERRO-L.

The third most important category is "economics" (a-), with 18.86% of the messages being assigned to it. This again corresponds with the aim and orientation of the list. Within
this category "economics general" (a-1, 2.69%) and "foreign investment & assistance" (a-16, 2.54%) are the most important topics. They are followed by "restructuring and deindustrialization" (a-17, 1.69%) and "employment" (a-19, 1.0%).

The above analysis seems to indicate that CERRO-L is a perfectly focused discussion list. That this is not necessarily the case shows the fact that the content of 14.58% of the messages falls into the category "miscellaneous" (z-). This is mainly the result of 6.23% "inappropriate messages" (z-8), which triggered 0.58% "complaints about inappropriate messages" (z-9). Viewed over the whole observation period, this level of noise is relatively moderate. However, its share increases steadily over time: from 0.56% in 1992 to 13.99% in 1995. As a matter of fact, this development has not been undetected by the administrators of the list. They started to send messages to subscribers who generated such inappropriate messages informing them about the purpose and orientation of the list and asking them to refrain from sending inappropriate messages. For few subscribers who did not refrain from sending inappropriate messages the parameters of their subscription were reset in such a way that the software rejects their messages. Also, the software’s new ability to detect "spams" helped to considerably reduce the noise level in recent months. However, all these measures tend to reduce the level of activity on the list.

The last category to which more than ten percent of messages are assigned, is "political" (b-). Within this category the distribution is more spread than in those discussed before. The two topics that stick out a little bit are "elections and public opinion polls" (b-2, 4.23%) and "government changes" (b-5, 2.01%).

With the exception of "education" (j-) to which 8% of the messages are assigned, none of the remaining categories shows up in more than 3% of the messages.

Table 4: Change in topics

<table>
<thead>
<tr>
<th>category</th>
<th>1992/93</th>
<th>1994/95</th>
</tr>
</thead>
<tbody>
<tr>
<td>economics</td>
<td>30.26</td>
<td>21.34</td>
</tr>
<tr>
<td>political</td>
<td>5.39</td>
<td>14.54</td>
</tr>
<tr>
<td>scholarly material</td>
<td>22.26</td>
<td>18.69</td>
</tr>
<tr>
<td>information</td>
<td>36.87</td>
<td>54.81</td>
</tr>
<tr>
<td>miscellaneous</td>
<td>11.65</td>
<td>11.49</td>
</tr>
<tr>
<td>other</td>
<td>19.83</td>
<td>17.28</td>
</tr>
</tbody>
</table>

Some interesting developments can be observed when we analyze the distribution by content over time. To get a clearer picture, we aggregate the first two years and the last two years
and combine the small categories into "other". The results can be seen in table 4. On the one hand, the table shows a shift from economic topics to political topics – although economics is the more important category in both years. This reflects the shift from general economic problems to problems related with policy and political institutions in Central Europe in general. The second trend we can see in table 4 is the growth of the category "information, communication and media", to some extent at the expense of the category "scholarly material and resources". This reflects the growth of the EDL and its transition from a small discussion circle to a large information resource.

Countries

In addition to its topical orientation, the discussion list CERRO-L also has a geographical focus: Central Europe. Therefore, in the second step of our content analysis we will discuss which countries or groups of countries were addressed in the messages distributed over the discussion list.

Table 5: Countries discussed (10 most important)

<table>
<thead>
<tr>
<th>country/region</th>
<th>percent of messages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eastern Europe</td>
<td>13.84</td>
</tr>
<tr>
<td>Hungary</td>
<td>11.73</td>
</tr>
<tr>
<td>Slovakia</td>
<td>11.62</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>11.46</td>
</tr>
<tr>
<td>Russia</td>
<td>10.62</td>
</tr>
<tr>
<td>Central Eastern Europe</td>
<td>10.41</td>
</tr>
<tr>
<td>Poland</td>
<td>8.77</td>
</tr>
<tr>
<td>Austria</td>
<td>7.92</td>
</tr>
<tr>
<td>Romania</td>
<td>5.12</td>
</tr>
<tr>
<td>Ukraine</td>
<td>5.02</td>
</tr>
</tbody>
</table>

Although the majority of subscribers to CERRO-L comes from North America and Western Europe, the discussion clearly focuses on Central Europe (see table 5). The seven most important countries in the list are post-socialist countries in Eastern Europe or aggregates therefrom. "Eastern Europe" is mentioned in 13.84% of all messages in the observation period. The most important individual country is Hungary which shows up in 11.73% of the messages. It is closely followed by Slovakia and the Czech Republic (11.62% and 11.46% respectively). When we aggregate these two countries and add the score for Czechoslovakia (which made it to 16th place in the list of countries) we find that more than one quarter of
the messages (25.67%) deal with "former Czechoslovakia". Similarly, the countries of the former Soviet Union are mentioned in almost 1/5 of the messages (18.49%).

The most mentioned country outside this group is Austria which appears in 7.92% of all messages. No other western country makes the top-ten list. The United States, which contribute more than half the subscribers to CERRO-L is mentioned only in 3.75% of its messages. This contrast between the spatial origin of the subscribers and which spatial areas they discuss is quite striking. It shows that even in a large EDL these aspects can be kept separate.

**Threads**

Not all messages that are distributed through an EDL are "original" messages. Many are reactions (replies) to earlier messages and need to be seen in context with these earlier messages. E-mail programs support replying to messages through a special "reply-function". When this function is used the program usually uses the subject-line of the original message and puts the string "Re: " in front of it. We use this feature for identifying replies in our dataset.

Of the 1893 messages that were distributed through CERRO-L in the observation period, 718 were replies to earlier messages. This corresponds to a reply-rate of about 38%.

The reply-rate is a measure for the intensity of discussion on the EDL. However, it is also influenced by the type of messages that are sent out and by the parameters set by the administrators of the discussion list. In many cases replies are complaints of subscribers to inappropriate messages (e.g., unsolicited commercial information). So, a list that is frequently the target of such inappropriate messages, is likely to show a higher response-rate. Also, when the EDL-software enters the e-mail address of the discussion list into the field "return-address" of the mail header, the chances are higher that replies that are intended for the original sender of the message will go to the list. Because of this and the chance for following complaints by other subscribers, the reply-rate will be higher for such a list as well.

As we have mentioned above, CERRO-L was to some extent plagued by such "noise" which has resulted in the appropriate change of parameters of the EDL. Therefore, we would expect the reply-rate to increase over the observation period and to drop off again at the end. In order to check this hypothesis we split the set of messages – ordered by date – into four quarters and compute the reply-rate for each subset separately. The result clearly confirms the
hypothesis. The reply-rate increases from 32% in the first quarter to 46% percent in the second and 48% in the third. In the fourth quarter this indicator drops off to 26%. This shows that this change in the parameters of the EDL worked the way it was expected to.

While many messages don’t generate any public replies at all, others trigger a long series of discussion. In netnews, where messages and replies are more closely tied together, such a cascade of arguments is called a "thread". We tried to identify threads in our dataset as well. We did this again by use of the subject-line of the messages. For each reply – identified by the leading "Re:" in the subject line – the program searched through earlier messages for one whose subject-line started with the string in the subject-line – less the leading "Re:", of course – of the message under investigation. It should be noted that this method cannot assign all replies to their original message. When participants in the discussion alter the subject-line of their reply-message, as they sometimes do, the link is broken and the message cannot be identified as part of a thread. Because of this, the figures we get for number and length of threads actually underestimate the true ones.

With this method we could identify 255 different threads. The majority of them (142) is just one reply long (i.e., 2 messages including the original one). The longest thread consists of 26 replies. In average, the replies that we could identify are 3.2 messages long, including the original message.

4. Conclusions

In this paper we have analyzed the development of CERRO-L, an electronic discussion list (EDL) with a Regional Science orientation. The EDL deals with regional economic development in Central Europe. The discussion in this paper is based on some theoretical arguments concerning academic virtual communities, scholarly communication and the geography of science that were presented in sections 1 and 2.

The empirical analysis is based on information concerning subscription to the discussion list and on the messages that were distributed over the list. We looked at the period 1992 - 1995. A number of important observations could be made. They will be summarized in this concluding section.

As it turns out, the EDL has grown continuously over the observation period. Subscription was highly stable in the sense that subscribers tended to stay on the list for a fairly long period of time. This demonstrates once again that also non computing related topics like those in Regional Science are well suited for this type of electronic discussion
The analysis showed a clear difference between the subscriber structure and the spatial focus of the discussion. While more than half the subscribers are located in North America, the discussion is strongly focused upon Central Europe. Only a small percentage of messages mentioned the US. Also in terms of topics the discussion on CERRO-L is well focused and centered around the aim of the EDL.

This discussion of Central European issues happens almost exclusively without Central European participation. Although the EDL clearly deals with this region, relatively few subscribers are located there. Moreover, those Central Europeans who are subscribed, are significantly less active in sending messages to CERRO-L. We can only speculate about the reasons for this. There may be cultural or historical factors that influence the willingness to become actively involved in a public discussion. This would be an interesting question for further analysis. However, it cannot be done with the material we have used in this paper. It would require a questionnaire-based survey of the subscribers.

We could also observe some changes in the EDL over time. With the growth of CERRO-L it became increasingly important as a general source of information about development in Central Europe. This could be seen in the growing share of requests for information. This development was supported by actions taken by the list management attempting to keep noise off the list. As we could see, small changes in the parameters of the list or in the policy of the list management, significantly influenced the level of activity as well as the content of the discussion. This shows the importance of monitoring the activities on a discussion list and the key role of the list-owner.

It is still too early for a final conclusion concerning the role of electronic discussion lists for the development of scholarly disciplines and academic communities. We did not find any evidence, however, that would raise serious doubts about the usefulness and potential value of e-mail based discussion lists for scholarly communication. The strong geographical concentration of the subscribers to this EDL in North America in relation with the fact that it is operated and managed out of Vienna, Austria, demonstrates the ability of such instruments to overcome space. The implications of this are difficult to predict, yet. But, we should keep them under close observation and collect more evidence in the years to come.
REFERENCES:


Rheingold, H., 1993. The Virtual Community, Reading: Addison Wesley.

APPENDIX: Categories and Topics

A. Economics
1. general
2. agriculture
3. industry
4. finance
5. trade
6. retailing
7. privatization
8. rural
9. urban
10. services
11. minerals
12. energy
13. forestry and forest industry
14. military conversion
15. high tech
16. foreign investment & assistance
17. restructuring and deindustrialization
18. marketing
19. employment
20. training, administration & consulting
21. transport
22. technology transfer
23. women in the labor force
24. consumer expenditures
25. housing

B. Political
1. general
2. elections and public opinion polls
3. parties and ideologies
4. history
5. government changes
6. human rights
7. wars and conflicts
8. treaties
9. unions (EU etc.)
10. current events
11. labeling regions
12. laws
13. intellectual property
14. nuclear: security, disarmament, smuggling
15. telecommunications policy
16. govt. revenues & expenditures
17. nationalism & ideology

C. Cultural and Social
1. general
2. population
3. ethnicity
4. migration
5. conflicts (non war)
6. rural
7. urban
8. recreation & leisure
9. sports
10. music
11. theater, films, books & exhibits
12. religion
13. welfare & social services
14. social issues & problems

D. History
1. general
2. economic
3. social
4. political

E. Scholarly Materials and Resources
1. general
2. library
3. maps
4. financial aid
5. govt. documents
6. new books, publications, newsletters
7. meetings, conferences & seminars
8. archives, data, theses
9. journal contents
10. new listservs
11. computer software
12. computer hardware
13. files
14. GIS (geographic information systems)
15. new societies
16. glossaries and directories
17. organizations

F. Information, Communication and Media
1. general
2. discussion group announcements
3. Fax, e-mail, WWW, etc.
4. computer usage
5. telecommunications
6. computer networks (including internet)
7. technical questions
8. computer policy & legislation
9. information newsletter & groups
10. requests for information (of all kinds)

G. Environmental
1. general
2. problems: pollution, projects, etc.
3. monitoring
4. environmental industries
5. global warming
6. water
7. sustainable development

H. Health
1. general

I. Planning and Policy
1. general
2. rural
3. urban
4. regional planning & policy

J. Education
1. general
2. exchanges, including funding
3. universities
4. new programs & institutions
5. courses and teaching methods
6. contacts with others
7. job opportunities
8. elementary & secondary schools
9. education policies of governments
10. research projects

X. Moderator's Comments
1. general

2. administration of CERRO
3. messages to subscribers
4. questionnaire
5. purposes of CERRO

Z. Miscellaneous
1. general
2. subscribe
3. unsubscribe
4. questionnaire
5. purposes of CERRO
6. job announcements
7. personal requests
8. inappropriate messages
9. complaints about inappropriate messages
10. personal information about subscribers
11. comments about what others submitted
12. incomplete messages
13. comments about LISTSERV
14. lists of subscribers - requests

NOTES:

1. CERRO is an acronym for “Central European Regional Research Organization”. This name is slightly misleading since there is no “organization” behind CERRO. It only exists in the virtual space. The administration and management of CERRO is done out of the Vienna University of Economics and Business Administration (Department for Urban and Regional Studies with the help of Central Computing Services).

2. Depending on the e-mail program there are different names for this function.

3. Common names are "listserv", "listproc", and "majordomo".

4. This is the case when such a message is interpreted as a regular message to the discussion list. It is then sent out to all subscribers, including the one who has generated the error message. So, another error message is triggered that is again sent out to all subscribers. The message bounces back and forth, in each round generating messages to all subscribers. Some subscribers may react and send messages to the list that complain or ask for an explanation. In doing so, they set off another thread of exploding messages.

5. The exact date cannot be derived any more.

6. Outside the US we used the ISO country codes. Since Bitnet is an academic network, we assigned all US-hosts to the "EDU"-category.

7. Today, the number of subscribers is larger than 800.

8. A linear regression of the US share against months shows a significantly positive slope.

9. Number of replies / total number of messages

10. Of course, we don’t know about responses via private e-mail. Such responses do not show up in the log-files of the EDL.