This paper addresses a set of issues related to language documentation that are not often explicitly dealt with in academic publications, yet are highly important for the development and success of this new discipline. These issues include embedding language documentation in the socio-political context not only at the community level but also at the national level, the ethical and technical challenges of digital language archives, and the importance of regional and international cooperation among documentation activities. These issues play a major role in the initiative to set up a network of regional language archives in three South American countries, which this paper reports on. Local archives for data on endangered languages have recently been set up in Iquitos (Peru), Buenos Aires (Argentina), and in various locations in Brazil. An important feature of these is that they provide fast and secure access to linguistic and cultural data for local researchers and the language communities. They also make data safer by allowing for regular update procedures within the network.

1. INTRODUCTION. There is a long and rich tradition of documenting and describing indigenous languages in South America. Since 2001, South American linguists have also

1 A preliminary version of this paper was given at the Fifth Annual Meeting of the Digital Endangered Languages and Musics Archive Network (DELAMAN): “Global-local archive relationships”, in November 2007 at the Instituto Nacional de Lenguas Indígenas (INALI) and Centro de Investigaciones y Estudios Superiores en Antropología Social (CIESAS) México, D.F. The authors would like to thank the Volkswagen Foundation for financial support and the technical group of the Max Planck Institute for Psycholinguistics (in particular Peter Wittenburg and Paul Trilsbeek) for manifold support.
been creating digital archives of primary language data in collaboration with European linguists within the Documentation of Endangered Languages program (DOBES) of the Volkswagen Foundation. For these activities, intellectual property rights and control of access to data published on the World Wide Web are important topics. With good reason, indigenous groups in South America are usually quite sensitive about research on and use of their traditional knowledge, and here each national administration is developing its own legal and practical strategies to protect local knowledge conceived of as a part of the national patrimony from bio-piracy and other kinds of misuses (see Carneiro da Cunha 2005 on the Brazilian policy). The “repatriation” of research results on material and intangible cultural heritage is therefore an important issue in South America. It is in this context that the emergent regional network of local and national language archive centers has to be seen.

In what follows, we will discuss the circumstances under which local archives were set up in Peru, Argentina, and Brazil (sections 2–4) before addressing some important shared aspects of this emerging network and its integration into an international network of language archives (section 5).

2. A REGIONAL LANGUAGE ARCHIVE IN THE PERUVIAN JUNGLE. In Peru, a DOBES project has been documenting the languages of the so-called “People of the Centre” since 2004. The People of the Centre are a multilingual cultural complex (Echeverri 1997), whose traditional territory was between the Caquetá and Putumayo rivers in what is now Colombian territory. However, in the early 1930s, some of the population was deported to Peruvian territory to the south of the Putumayo river by former employees of the rubber company known as “Casa Arana.” Thus, speakers of these languages are now separated by considerable geographical distance and by an international border. The documentation project focuses on Bora, Huitoto, Ocaina, and Resígaro, which are taken to be a representative subgroup of the People of the Centre. An international team is collecting cultural and linguistic data on these groups in Peru and to some extent also in Colombia: the German linguist Frank Seifart, the Colombian linguist Doris Fagua, and the Swiss-Peruvian anthropologist Jürg Gasché, who has also contributed his data collected since the early 1970s in Colombia and Peru. These data are augmented by a salvage documentation of the by now practically extinct Nonuya language, collected by the Colombian anthropologist Juan Álvaro Echeverri and the Franco-Basque ethnolinguist Jon Landaburu in the 1990s in Colombia.

The project was housed until recently at the University of Bochum and has now moved to the University of Regensburg. The Peruvian host institution is the Instituto de Investigaciones de la Amazonía Peruana (IIAP) in Iquitos, in the heart of the Peruvian jungle. The IIAP is a major research center for the scientific study of various aspects of the Amazonian environment. The technical group of the IIAP already has a strong tradition in building and maintaining databases with a focus on the biodiversity of the Peruvian Amazon, but increasingly including data from social sciences as well. In September 2006, a digital archive for linguistic data was set up at the IIAP and a formal agreement was signed between the

Thanks also to Nicholas Thieberger, José Antonio Flores Farfán, and two anonymous reviewers for their valuable comments.
IIAP and the Max Planck Institute for Psycholinguistics as the institution representing the central archive of the DOBES program. In this archive, a complete copy of the data collected by the People of the Centre DOBES project has been deposited.

Bringing together data on the People of the Centre in one central archive will greatly facilitate comparative studies on this cultural complex. For instance, the genetic relationship between Huitoto, Ocaina, and Nonuya may be studied in more detail using the extensive lexical as well as grammatical data centrally archived in Iquitos. Additionally, the centralized documentation of elaborate ritual discourses and repertoires of thousands of songs (for which the People of the Centre are renowned) facilitates comparative ethnological research as well as research on the indigenous rhetoric cultures, a new research area rarely pursued so far for Amazonian languages.

The local archive in Iquitos also plays an important role for the Peruvian speech communities, since they can physically access the data in Iquitos, which they visit regularly. Once reliable Internet connections reach the indigenous communities, which we expect to happen in a few years, the digitally stored data in the Iquitos archive will also more easily bridge gaps between the indigenous speech communities that were forcefully separated. For instance, Colombian Boras can access recordings of Bora songs made in Peru, and vice versa. This is important in the context of language loss—in particular, the obsolescence of culturally important genres such as songs and good spells. So far, Colombian and Peruvian Boras and Ocainas have been sending tapes with recordings of songs and good spells to one another with the help of researchers.

The digital archive at the IIAP is not intended to be just a self-contained linguistic documentation of the People of the Centre. On the contrary, the linguistic data stored in the archive at the IIAP will be linked to existing databases at the IIAP on its website (see http://www.siamazonia.org.pe), which comprises a variety of data on different aspects of the Peruvian Amazon region, among them geological, botanical, and zoological. Of particular interest will be the link between linguistic documentation and a set of socio-cultural data that describe the current state of the societies and cultures of these groups. These were collected in the context of a collaborative project on the social diversity of “forest people,” carried out by the IIAP and the Universidad Nacional de Colombia. Another set of data that will be stored in the archive in Iquitos is linguistic and cultural data on Cashinawa, an indigenous language spoken in the Peruvian-Brazilian border region, which is being documented by another DOBES team, led by Eliane Camargo and Sabine Reiter. Depositing this data at a Peruvian institution responds to an explicit request by the Peruvian Cashinawas.

Within the specific Peruvian context, the archive in Iquitos has two major impacts. First, it stimulates collecting data from a broader variety of Peruvian languages and cultures and making them available in Peru. Unfortunately, Peruvian national policy pays relatively little attention to documenting the indigenous cultures that are alive today, giving preference instead to studying the more prestigious civilizations of the past, for instance the Incas. Another problem is that there are very few national institutions that could serve as archives for such data anyway. Therefore the best currently existing ethnographic, ethnomusicological, and ethnolinguistic collections of the indigenous cultures of Peru are abroad in Europe or North America. Thus, one of the aims of the archive in Iquitos is to contribute to changing the tendency of exporting objects and data towards a culture of sharing data with the national institutions of the countries where these data were collected.
Second, the archive in Iquitos will have a positive effect on distributing data and making them available within Peru, which has sometimes been difficult because of the geographic fragmentation of the country. It is difficult to gain access to the Amazon region, where many of the indigenous languages are spoken, from the large cities on the coast and in the Andes, where most of the universities and research centers are. There are no roads connecting Iquitos, which lies in the center of the Peruvian Amazon, with the rest of Peru. It can be reached only by expensive air travel or week-long boat trips. For this reason, it is generally difficult for Peruvian linguists or anthropologists to study Amazonian languages and cultures. The data of the People of the Centre in the archive at the IIAP offer Peruvian researchers—professional linguists, and both graduate and undergraduate students—access to a broad range of the manifold uses of these languages. It is hoped that it will motivate and help their research in various ways and stimulate more Peruvians to study more of the fifty-seven indigenous languages spoken in Peru.

The acceptance and success of the language archive and the DOBES data in Iquitos on a national and regional level depend on their usability for local linguists. Therefore, the team of the People of the Centre project, together with members of the Max Planck Institute for Psycholinguistics, held the first workshop to introduce the electronic tools of the DOBES program at the Universidad Nacional Mayor de San Marcos in Lima. Similar workshops have been held by project members at the Universidad Nacional de Colombia en Bogotá. Further workshops are planned in collaboration with the DOBES teams from Argentina and Brazil (see sections 3 and 4).

In summary, the first of the three local South American language archives was set up in a research institute with a tradition of digital archives in the Peruvian jungle. It will facilitate comparative research on languages spoken in the Peruvian Amazon and neighboring areas in Colombia and Brazil and will improve the access to data on these languages for national and international researchers, as well as for the speech communities themselves.

3. AN ENDEAVOR FOR THE ARGENTINE LANGUAGES, ESPECIALLY THE LANGUAGES OF THE CHACO. In Argentina, the initiative to set up a local archive came from a DOBES project focusing on the languages of the Chaco, a region in northern Argentina near the Paraguayan border. This project, entitled “Endangered Languages, Endangered Peoples in Argentina (South America): Mocoví (Guaycuruan), Tapíete (Tupí-Guaraní), Vilela (Lule-Vilela) and Wichí (Mataco-Mataguayan) in their Ethnographic Context,” originated in the Facultad de Filosofía y Letras, Universidad de Buenos Aires, in collaboration with the Linguistics Department of the Max Planck Institute for Evolutionary Anthropology in Leipzig. Within this framework, the project director Lucía Golluscio and her team have collected linguistic and ethnographic data for four endangered indigenous languages in the Chaco region. Part of the output of this project are publications for academic purposes (Golluscio and González, in press; Lozano 2006), eleven documentary videos, and a significant amount of written materials for dissemination, especially in the community schools in the context of language maintenance and revitalization. Among them are Citro 2006, Citro and Gualdieri 2006, Hirsch 2006, and a set of educational materials authored by members of the communities in collaboration with team members and teachers: Ciccone, González, Hirsch, E. López, Mercado, and Montes 2006, F. López 2006, M. López and Casal 2006, and Segundo and Riquelme 2006.
After detailed consideration of institutional and legal issues, the Max Planck Institute for Psycholinguistics and the Argentine Council for Scientific and Technical Research (CONICET, see http://www.conicet.gov.ar) signed an official agreement on scientific and technological cooperation. CONICET is the national Argentine institution in charge of designing scientific policies and the promoting scientific and technological research. This agreement established a first Argentine “Laboratory of Documentation and Research in Linguistics and Anthropology” (LabDILA) at the Centre for Scientific and Technological Information (CAICYT, see http://www.caicyt.gov.ar), a subdivision of the CONICET, which hosts a copy of the digital documentation that resulted from the above mentioned DOBES Chaco Languages Project. This agreement has considerable importance in the Argentine national context, raising public awareness and ensuring political support for endangered languages and language documentation. The participating German institutions, the president of CONICET, and the German ambassador in Argentina expressed strong support for establishing the archive in Buenos Aires during the ceremony of the signing of the agreement. The event has also been an opportunity to foster cooperation among the regional archives in Brazil, Peru, and Argentina.

The agreement underlines the more general commitment of CONICET to collaborate with the Max Planck Institute for Psycholinguistics in technological infrastructure and networks of archives of linguistic resources. In order to maintain the archive from an organizational as well as a technical perspective, two positions have been created at CONICET to take care of digitizing and updating to more durable backup formats of tapes, photos, and other endangered Argentine language materials, which have been created in the past and are now gradually deteriorating.

Documenting and archiving the indigenous languages of Argentina are particularly urgent tasks. Those languages are threatened by the long-standing pressure of Spanish, the growing influence of the media, and the diminishing number of fluent speakers, among other factors. In many cases, it is not just the language but the communities themselves that are seriously threatened (Golluscio 2004).

Members of the indigenous communities have expressed great interest in these preservation tasks, and it can be observed that the language documenting and archiving activities have very positively contributed to their self-esteem. They highly value the fact that they can directly access linguistic material of their own as well as that of other peoples. This is recognized in the agreement, which explicitly states that the indigenous communities have unrestricted access to the resources. The local digital archive provides excellent resources for developing educational projects related both to producing pedagogical materials and training in linguistics and language documentation tools. A fruitful relationship is developing between the local archive and the program on indigenous linguists beginning at the Universidad de Buenos Aires in 2008 supported by the National Ministry of Education. Establishing this archive is also an important prerequisite for new documentation projects focusing on other minority languages in Argentina.

The CONICET Division of Digital Documentation has organized three international academic events under the auspices of the Volkswagen Foundation, CONICET, and the Universidad de Buenos Aires. These activities are helping to promote the archives among potential donors and users, as well as to foster discussion and the search for consensus. All of this contributes to clarifying and legitimizing the function of linguistic documentation.
and archives in South America. The first Symposium on “Language and Culture Documentation in Latin America,” with a focus on “Ethics and Collaborative Research,” as well as the Workshop on “New Technologies and Scientific Collaboration between Europe and Argentina: A View from the Humanities,” were held in 2007 with the active participation of members of the communities, foreign and local researchers, and students. For 2008, the archive team is organizing a second symposium on “Language Contact and Documentation” and training courses in the application of DOBES technology as well as in audio and video recording. The CONICET Division of Digital Documentation has recently been incorporated as a new institutional member of the Digital Endangered Languages and Musics Archive Network (DELAMAN, see http://www.delaman.org) in Latin America. Membership in international networks of digital archives ensures that we keep up to date and provides a productive space for discussion.

In sum, this Argentine endeavor has secured highest-ranking national and international institutional support for a local language archive in this country. Besides its importance for international cooperation and scientific research, the local archive in Buenos Aires seeks to play a role in language and culture maintenance efforts of the indigenous peoples of Argentina.

4. LANGUAGE ARCHIVES AND OTHER DOCUMENTATION INITIATIVES IN BRAZIL.

Brazil is the largest South American country, with the greatest number of indigenous languages (Moore 2007), and language documentation has had a long and diversified history in this country. Three projects on Brazilian indigenous languages participated in the first years of the DOBES program from 2001 to 2005. One of these projects, on Kuikuro or Upper Xinguan Karib, conducted by Bruna Franchetto and others, was based entirely in Brazil, at the Museu Nacional, which belongs to the Federal University of Rio de Janeiro and is one of the oldest academic institutions of this country. The other two, on Awetí, led by Sebastian Drude and others, and on Trumai, led by Raquel Guirardello and others, were hosted partly in European institutions and partly in Brazil, at the Museu Goeldi in Belém do Pará, a federal research institute of the Ministry of Science and Technology, which has a long tradition of research on indigenous languages of Amazonia. The fourth Brazilian DOBES project was run by Sérgio Meira on one Tupian and two Karib languages. Several projects on Brazilian languages funded by the Endangered Languages Documentation Program (ELDP) at the School of Oriental and African Studies (SOAS) in London are also hosted at the Museu Goeldi and Museu Nacional, as well as at the University of Campinas.

The three oldest DOBES projects—one Awetí, Kuikuro, and Trumai—have established a particularly close cooperation, which continues to give rich results, in particular in understanding the regional Amazonian intertribal and multilingual system “Upper Xingú,” to which all three communities studied belong. The end of the DOBES funding did not mean the end of this fruitful cooperation—quite the contrary. Since 2006, the three researchers have been involved in a larger Brazilian project, funded by Edital Universal and the National Council of Technological and Scientific Development, in cooperation with colleagues who work with each of the other Upper Xinguan peoples. This cooperation again involves anthropologists and archaeologists and is also in this sense a natural offspring and
continuation of the comparative work stimulated by the DOBES projects (Fausto et al. In press).

The first digital national language archive in Brazil has been set up in Rio de Janeiro at the Museu do Índio, an institution that belongs to the federal indigenous affairs governmental organ, Fundação Nacional do Índio (FUNAI). For a long time, the Museu do Índio has had a deep, lasting, and well-known tradition of offering the following services for the Indians and the broad public alike: first, to disseminate knowledge; second, to document by archives of objects of material culture, of historical written documents about the Indians and their relationship with the Brazilian state, of images, and of sounds, albeit until now the last is little known and explored; and third, to facilitate research on topics related to the indigenous populations, including research by indigenous people as well as by scientists or journalists. As for disseminating knowledge, besides editing books and other materials, the Museu do Índio hosts a number of exhibitions, which are planned and worked out in direct cooperation with local indigenous organizations. It is worth mentioning here the current exhibition called “Tisakisü: Tradition and new Technologies of Memory,” which was proposed by members of the DOBES Kuikuro team. The exhibition illustrates how the Kuikuro themselves, with financial support from governmental and nongovernmental institutions, take over the ongoing documentation, which started as a DOBES project conducted by a team of scientists. In particular, it reflects the intricate relation of the pretechnological traditions with new technologies incorporated and assumed by the Kuikuro from writing in school, via lasting images and audio recordings, to digital technology such as annotated video recordings. The Kuikuro now use these technologies to build a rich register of their cultural heritage from an insider perspective, but informed by academic research (Franchetto, in press). The exhibition includes videos produced by the so-called “Kuikuro video makers”: five young Kuikuro who began their training during the DOBES project and with the help of the Brazilian NGO Video Nas Aldeias (video in the villages). The films they made are now exhibited abroad and have gained general recognition, including a dozen or so prizes won in and outside Brazil.

With the firm foundation of the profile of the Museu do Índio described above, it was not by chance that it was this institution that volunteered to host the first national language archive, a proposal made by its far-seeing and energetic director in October 2006. The Museu do Índio also has the physical and technological infrastructure appropriate for taking up this ambitious enterprise.

In April 2007, the federal indigenous affairs governmental organ FUNAI and the Max Planck Institute for Psycholinguistics signed a first agreement as a formal basis for transferring the language archiving technology to the digital archive at the Museu do Índio. In October 2007, the Museu Nacional, host of the Kuikuro DOBES project and of ELDP projects, formally adhered to this agreement, and the federal Minister of Research and Technology and the Max Planck Institute for Psycholinguistics signed a similar agreement in order to provide an official basis supporting the second digital archive to be installed in Brazil, at the Museu Goeldi.

Also at the Museu Goeldi, Sebastian Drude is working on a project that builds standards for the language archives of the Museum and is creating a website in Portuguese dedicated to language documentation, providing instructions about data formats, programs, and methods used to build and organize modern language archives. This site will provide
support for documentation activities by international programs such as the ELDP or DO- 
BES, as well as by local initiatives. It will also be useful for those who are transferring 
individual personal and legacy archives to the central digital archives at the Museu Goeldi 
and the Museu do Índio. The website will be especially important for a new national pilot 
program for the documentation of twenty indigenous languages called DOCLIN, which 
has been recently announced by the government and the federal indigenous affairs govern-
mental organ FUNAI and will start in 2008.

Interestingly, in all cases the Brazilian institutions that host the local digital language 
archives are Museums, with or without an academic research profile, but in accordance 
with their traditional mission to curate archives documenting the national and local heri-
tage. Indeed, the coordinator of the linguistics division of the Museu Goeldi, Denny Moore, 
participated in the development of the new area of digital language documentation from the 
beginnings in the late ‘90s, and the division has hosted the largest number of documenta-
tion projects in Brazil so far.

It is hoped that the new Brazilian local language archives will help to preserve and to 
give access to the huge and quickly growing amount of data resulting from the different 
documentation activities. They are central for disseminating data and promoting language 
documentation, which is now taken over and funded by national organs in the above-men-
tioned DOCLIN program in cooperation with national institutions and with the major in-
ternational initiatives. Several linguists, anthropologists, and archivists have been trained 
in the methods of the new field through participating in the documentation projects and in 
many courses taught in various Brazilian universities and at the annual meetings of various 
Brazilian scientific associations.

There are several challenges for the new area. One is a general suspicion of initiatives 
from abroad, as they are easily associated with old colonialist practices. This gives rise to 
diffuse fears that the country and the Indians in particular, could be fleeced of their valuable 
knowledge. For many years, in particular during the military dictatorship, official organs 
have fomented such views, since researchers were a potential source of criticism of their 
policies, which in some cases prejudiced the indigenous groups. As for the researchers, 
there is also a certain culture of individualism that is common in the human sciences in 
general and maybe particularly pronounced in South America, where many colleagues are 
reluctant to share their preliminary results and their raw data, especially if collected in gru-
eling fieldwork. It will take quite a while to change such often rather entrenched attitudes, 
as often only new generations of researchers are open-minded enough to recognize the 
benefits of the new research paradigm. The establishment of local and national archives is 
certainly very helpful in this respect.

In Brazil, the national archives and especially the Brazilian documentation program 
DOCLIN will also have a positive impact on the notoriously problematic infrastructure of 
national institutions. This holds in particular for those regions of this continent-sized coun-
try, where the majority of existing indigenous languages are spoken (the Amazon region 
and the central west of Brazil). Unlike the traditionally more prosperous southeast, these 
regions suffer from a profound lack of resources so that fast internet connections and even 
a reliable energy supply are rare. While people are currently attempting to ameliorate this 
situation, for instance through the national program of digital inclusion, it will take a long
time to improve and sustain the infrastructure, as the problems are of a structural nature and the institutions involved are traditionally rather unstable.

In summary, local archives have been established in Brazilian Museums, with indigenous communities directly involved. These archives are part of a number of initiatives promoting digital language documentation in Brazil, which include a large national language documentation program, the first of its kind in South America.

5. AN EMERGING NETWORK OF ARCHIVES IN SOUTH AMERICA. The initiatives in all three countries are part of a very promising movement towards creating and enriching local archives and documenting the indigenous languages of South America. Creating a network of language archives and related activities attend to a growing demand by an increasingly large number of colleagues in South America, who agree that it is not an appropriate practice to keep precious recordings in personal collections. Even more important is a demand by more and more indigenous groups who are concerned with the imminent loss of central features of their traditional culture and knowledge and in particular, their languages. A standing challenge is to train more local specialists to maintain and update the digital archives, as well as to establish and oversee regular update procedures with larger archives, such as the central DOBES archive in Nijmegen.

Cooperation between the initiatives for language documentation and archiving in South America has been fruitful and is central for all activities in this area. We understand the local archives in Brazil, Iquitos, and Buenos Aires to be the seed of a larger South American—or even Latin American—network to be established, which is open to other archives that might be created in the future. The local archives provide faster and more secure local access to regional language data. In many cases, indigenous communities physically access the centers and create copies of the documentation of their language and culture in order to make them available in their communities. If organized in a network, the local archives can also make the data safer via automatic creation of mutual backups. The digital archives provide also the basis for a South American participation in the coming Common Language Resources and Technology Infrastructure (CLARIN) program, a large-scale European initiative for a network of digital archives (see http://www.clarin.eu). The Peruvian-Argentine-Brazilian documentation and archiving network also enhances resources by planning shared training and disseminating activities. For instance, training courses in corpus management can be held once centrally for the three countries, instead of three times.

Regarding the content of the data stored on the three digital archives, it is interesting to point out that each of them includes documentation of multilingual regional intertribal systems: the People of the Centre in Peru and Colombia, the Chaco in Argentina, and the Upper Xingú in Central Brazil. Such emerging regional multilingual projects are a promising source for future comparative research. Other research possibilities offered by the data stored at the centers include comparative genealogical, typological, and areal studies, such as those on common and specific structural features between Chaco and Amazonian languages or between languages of the same family spoken in distinct areas—for instance between the Arawak languages of the People of the Centre and the Arawak languages spoken in the Upper Xingú. The cooperative research carried out in this network can thus also help deepen our knowledge of the history of the populations and their language contact.
across South America. In general, this network of local archives fosters multidisciplinary research, establishing diverse links among institutions, researchers, students, and members of the communities in distinct regions.

Finally, we want to mention that the Archive of the Indigenous Languages of Latin America (AILLA) at the University of Texas at Austin is another digital archive with a substantial and constantly growing collection of resources for South American languages. As such it is part of the global network of digital archives. AILLA has always made clear that cooperation with regional and local archives, when they should be established, is a central part of its mission, although it does not have the resources to establish such archives itself. AILLA is already pursuing plans to transfer copies of its holdings to each of the three archives discussed above and will continue to serve as a backup repository and archiver of legacy materials, as it has been doing since 2000.
REFERENCES


CITRO, SILVIA 2006. La fiesta del 30 de agosto entre los mocovíes de Santa Fe. Serie Documentos. Buenos Aires: Instituto de Lingüística, FFyL, UBA.


FAUSTO, CARLOS, BRUNA FRANCHETTO, and MICHAEL HECKENBERGER. In press. Language, ritual and historical reconstruction: Towards a linguistic, ethnographical and archaeological account of Upper Xingu Society. In K. David Harrison et al., eds., in press.


GOLLUSCIO, LUCÍA, and HEBE GONZÁLEZ. In press. Contact, attrition and shift in two Chaco languages: The cases of Tapiete and Vilela. In K. David Harrison et al., eds., in press.


Frank Seifart
frank.seifart@berlin.de