Mr. Chairperson, Ladies and gentlemen, it is indeed a great honor for me to be given this opportunity to address you on the occasion of this special meeting. My English name is Sean Zhu, I come from Wuhan city, Hubei province, China. I am a teacher in the College of Computer Science, South-Central University for Nationalities. Sixty percent of the students of our University come from Chinese minorities. I am one of the founders of the Information Processing Institute for Minority Language, which started to operate in January, 2007. My lecture today comes from one of our projects: “Using information technology to salvage and protect Nü Shu”, which is Sponsored by the 2008 science and technology project of The State Ethnic Affairs Commission of the PRC. Nü Shu is a unique script popular among women in several townships of Jiangyong and its neighboring counties in south China’s Hunan province. This script has been mainly handed down for local women’s use only. Invented by women, and used by women, theses italic characters were not known or understood by the outside world until it was discovered by several teachers from our University in the early 1980s.
This is the overview of my speech. Now I will proceed with more specific explanations. First I would like to explain why we pursued researching this field. Then I will show you our initial ideas about how to protect the Nü Shu. Character set, importing methods we have developed, the website construction and the Forest of virtual Steles being constructed to complement that site. These are the four parts of our recent work. At the end of my speech, I will draw conclusions about the significance of our work and share some of our feelings regarding this task with you.
First of all, How did this project begin? Our story began on 2007, March 8, Women's Day.
In the evening of that day, Professor Wu Xianli and me attended the “South Lake Academic Salon”, which was held on the top floor of our University’s library. Professor Wu comes from the Institute of Automation, Chinese Academy of Science. Since that time, he has been working as a special professor in our college, with his main research field being OCR. During this time we have been working together as colleagues, developing a program to accomplish Printed Yi Character Recognition.
To return to my account of that original 2007 meeting, Professor Xie Zhiming was the academic adviser of that Salon. Xie is a philologist, specializing in Nü Shu. In fact, he is one of the several teachers who originally discovered and studied Nü Shu. He now has been studying Nü Shu for 26 years, since 1983.
Professor Xie first introduced the value of Nü Shu to us. Nü Shu’s values can be recognized from three viewpoints: Anthropo-sociologic value, Language’s philological value, and also a literary value. These values will not be more specifically explained because they are not my professional field. However, if you feel interest regarding these values, I recommend professor Li Qingfu who is also now in KunMing.
Xie then told us the predicaments of Nü Shu. Four difficulties need to be considered.

1. Nü Shu’s existent environment is poor, thus hampering its survival.
2. Mysteries of Nü Shu are so numerous, that studying up to this point has not been sufficient.
3. Nü Shu study has been hindered by fake and poor versions being offered as genuine examples.
4. Dissemination of Nü Shu study findings has been caught in an information bottleneck.

All these predicaments are addressed in our paper. But, because we are professionals in information technology, we care the most about the information bottleneck.
The information bottlenecks of Nü Shu’s study lie in the difficulties of collection, statistical analysis, usage, and recognition of the cultural communion tied to Nu Shu. Our project team’s major member, famous Chinese scholar of Nü Shu, professor Xie Zhimin, has, during his 26 year Nü Shu study and desire to keep Nü Shu in its "original ecology " to avoid the fakes, been using the picture-paste method to write articles.

This method is really time consuming and strenuous. A large number of Xie’s research results are difficult to issue because of the copy quality of the document by the time Professor Xie Zhimin has put them together. The domestic and international publication organizations deeply feel this inconvenience. This information bottleneck has seriously influenced the spread abroad Professor Xie Zhimin’s study recording the development of Nü Shu civilization.
In the three days following the salon, I wrote a technology report titled “Several Key Thoughts Regarding Nü Shu’s Informationization”, systematically describing our initial ideas about how we can use information technology to salvage and protect Nü Shu.
Our idea can be describe in two words: “Being Digital”.

To accomplish all these tasks, we first needed a character set of Nü Shu, which could be shown as Nü Shu accessible to microcomputers. This is the foundation for hope of informationizing Nu Shu on a wide scale. Then we needed to design an importing method supporting the character set, which would then make it possible to input, display, modify, copy, cut, paste and print Nü Shu with computers. After that, a website of Nü Shu founded to collect and sort the relevant materials, whether they be letters, pictures or videos. This then allows us to publicize the character set and importing method we have designed, in order to teach Nü Shu and to use Nü Shu in literary creation. This, in turn, then allows us to promote communion between scholars and fans of Nu Shu. At the same time, virtual reality technology will be used to protect the important literatures of Nü Shu in the form of virtual steles. The website and the virtual steles can further publicize Nü Shu within a simulated historical context, including its graphemes, pronunciation, meaning, and the singings which are used as a way to convey the language from generation to generation in the past.
As I have mentioned above, the first step is to Create Nü Shu character set.
The character source of our Nü Shu character is the manuscript of the "Chinese Nü Shu Dictionary", which condenses half a lifetime of energy exerted by Professor Xie Zhimin. Xie gathered 2435 single characters of Nü Shu with every character keeping its original structural appearance and physique strictly. Its phonetic notation and meaning come from the late generally acknowledged successor Gao Yinxian and Yi Lianhua. Each character’s source has been indicated one by one, in order to ascertain validity. This dictionary with this information has now been published by National Publishing Company, sponsored by the General Administration of Press and Publication of the People’s Republic of China.
Secondly, we may need a set of more normal handwritten characters of Nü Shu to serve as the copy for recording or for reproduction. Different from abundant Art “Nü Shu” or Nü Shu calligraphy that is now on the market and reflecting the artistic re-interpretations meant to appeal to contemporary audiences, we hope to keep the characters reflecting what is believed to be their original appearances. The reason is: Since Nü Shu still has numerous mysteries unsolved, any unexpected change or artistic digression from earlier versions may destroy Nü Shu’s primitive physique structure. This may increase the probability of Nu Shu’s mysteries, the cultural information buried within the primitive physiques of the characters, will become indiscoverable.
For this reason, we have specially invited the granddaughter of Gao Yinxian, Ms. Hu MeiYue, to come to Our University as a research consultant. She used the technique of Nü Shu’s handwriting, which her grandmother Gao Yinxian had taught her that year, to write all 2435 Nü Shu characters for us. All characters are read over and examined strictly by Professor Xie Zhimin(VIDO:01:10), ensuring that the every character was written scientifically.

This video was taken on January 20,2008. Five days later, Professor Xie got very serious illness. He lost the ability to speak. But now he has got well, continuing his Nü Shu study. Professor Xie is a 76 years old man with a cancer in his body, but everyday he work more than ten hours. It is the greatest honor in my life to have this chance work with a scholar like him.
An importing method is necessary to help display the Nü Shu characters. In the information age, the Nü Shu importing method not only undertakes the inputting and spreading function, but also can be used to teach and hand down Nü Shu.
As is stated in preceding paragraphs, every character in the character set has been phoneticized with the International Phonetic Symbols according to the pronunciation of Gao Yinxian and Yi Lianhua, but the International Phonetic Symbols are not popularized in China. In order to first spread Nü Shu smoothly in its homeland China, Professor Xie Zhimin concluded his Chinese Nü Shu dictionary with "Nü Shu phonetic symbols as well as the Chinese pinyin alphabet comparison table " , so we “translate” the International Phonetic Symbols of Nü Shu into pinyin according to this table. Then with the help of this Nü Shu pinyin, we have designed the Nü Shu pinyin import method. The goal for using the pinyin importing method is to “know pronunciation, and know how to type".
It is difficult for users to be familiar with the pronunciation of more than 2,400 Nü Shu characters when they use the pinyin import method. In order to help ordinary typists input the Nü Shu manuscript after short training, we are designing the components’ importing method with reference to the Five-Stroke import method used with Han characters(五笔字型). The main idea is to divide the Nü Shu into some parts in accordance with certain rules, and distribute these parts to different letters of the keyboard. If we input these letters in a certain order, Nü Shu characters will then be displayed. The requirement of the components import method is that “simple rules allow the user to know how to type when he sees a character”. Two import methods have been used to input 183 Nü Shu literatures, with more than 20,000 characters.
How to introduce and spread Nü Shu culture to domestic and international personages has perplexed Professor Xie Zhimin for a long time. In his more than 20 years' of Nü Shu study, Professor Xie published papers, wrote monographs, handled lectures, and taught lessons about Nü Shu, but the benefits from all this work have been limited to a narrow area of influence. With the development of information technology over the last 20 years, Internet technology has become the best means by which mass media can be used to spread new knowledge because of its abundant resources, numerous users, and strong communication abilities. Because of this, we are using the Nü Shu character set and importing method developed by ourselves to build a comprehensive Nü Shu website with an advanced, dynamic webpage design technique.
Nü Shu website: [http://womenwritings.scuec.edu.cn](http://womenwritings.scuec.edu.cn)

The website is under construction, and we welcome you to visit it to give your comments.
Professor Xie Zhimin has a grand plan, intending to select 300 most representative Nü Shu’s articles from the Nü Shu literature to be engraved on the stone tablet using Nü Shu, Chinese, and English so the seal cuttings may endure far into the future and be easily translatable. But in reality, this will cost a large number of financial and material resources to realize. So, until such monies can be accumulated for this project, we have thought of virtual reality technology as a means of immediate construction of a stele site, a very active technological research field in recent years.
The process can be divided into four different stages:

We have designed Nü Shu virtual Steles based on FLASH.

Immersion-type virtual reality systems use cave-type audio-visual equipment or helmet display devices to close off the peripheral vision, constrain the sense of hearing and other tactile feelings, producing a kind of illusion that the user is in a fictitious environment and does not have direct relations with the real world of outside. The forest of Steles of Nü Shu in Virtual Reality at this stage will be realized in the virtual reality laboratory of the College of Computer Science, South-Central University For Nationalities. The laboratory has a series of virtual reality equipment worth more than 3 million yuan. Just before I came here, I saw some result of this laboratory. My colleagues have established the framework of this system. So I believe that after this summer, we can see some initial results in this laboratory. We welcome all of you to visit our University, Our college, our institute and laboratory at that time.
Rescuing and protecting Nü Shu with information technology is an interdisciplinary comprehensive task which combines art and science, including the fields of linguistics, philology, ethnology, code studying, statistics, computer applications, and psychological engineering. It requires researchers of different disciplines to cooperate closely, perceiving research as an exploring spirit that is “wading across a stream by feeling our way”摸着石头过河！
During the time we carried out the Nü Shu project, our institute established two goals. The first one is “By cooperating with linguists and anthropologists, we devote ourselves into using information technology to protect civilization’s achievements of all minorities in China, so as to continue richness of Chinese culture’s genes!”
Our Institute’s second goal is this: Taking the spreading of knowledge of minority civilizations as our application background, we are engaged in scientific research on Pattern Recognition, Computer Vision, Machine Learning, Artificial Intelligence, and Virtual Reality.
We have three papers that have been accepted for presentation in three different sessions.

“Using information technology to salvage and protect Nü Shu”, is the title of the paper accepted for today’s Session.

“The Research of Printed Yi Character Recognition”, is the title of the paper accepted for the session on Software Programing for Language Fieldwork and Its Realization. And I will address this topic tomorrow.

“Research on Radical Input Method of Nü Shu”, is the title that has been accepted for the Session on “Diversity of National Culture of Central and Southwest China and Protection of Non-material Culture” Tian Wei will address this topic there the day after tomorrow.
This page show the relationship among these three papers. Generally speaking, our input material is initially minority literature. If this minority language has a character set, like the Yi language, we could use OCR technology to make them been digital. Importing tools help to correct some wrong OCR results. If this minority language does not have character set, like Nü Shu, we would design a character set for it, and then design an importing method to input its characters so we can obtain digital literatures. Then we would build a website and virtual steles to spread these literatures and this minority culture outward to the rest of the world. To state it briefly, we are trying to find an integrated solution for minority literatures.
In the coming years, we will try to design a device named “Minority literatures automated importing machine”. Computer vision and machine learning methods will be used in this device.

Our idea is shown here:

We will design a pipeline for “being digital”, which means “Input literatures, and output computer files”.

There are two key features in this system.

1. Each character can be segmented correctly from initial material, depending on the computer vision.

2. Systems should judge whether this character is in the character set or not, and if not, it should be added to the character set automatically. This involves machine learning.
Our Poster Session has more detail about what I have mentioned, and offers some ideas of what our future work will be. It’s No. is P059-P061, We also prepared some publicity materials and a tryout CD. If you feel interested in what we have done, please contact us to get these materials. You can see the contents from my colleague Liu Sai. We may need you to sign your name to make sure that the CD will only be used for academic research.
Thank you very much for your kind attention. I am looking forward to hearing from you during the discussion period and the coming days in Kun Ming.