

统计之都网站管理手册

谢益辉¹

2010 年 2 月 25 日

¹Email: yihui.xie@cos.name; 个人主页: <http://yihui.name>; COS文集: <http://cos.name/author/yihui/>

目录

第一章 发展纲要	3
1.1 愿景目标	3
1.2 组织结构	3
1.2.1 主站	3
1.2.2 论坛	3
1.2.3 维基	3
1.3 制度原则	3
1.3.1 责任权利	3
1.3.2 组织文化	3
第二章 软件说明	4
2.1 GIT	4
2.1.1 GIT入门	5
2.1.2 GitHub	5
2.1.3 COS与GIT和 \LaTeX	5
2.2 \LaTeX	6
2.2.1 安装 \LaTeX 与 \LaTeX	6
2.2.2 \LaTeX 入门	6
2.2.3 配置 \LaTeX	7
2.2.3.1 中文问题	7
2.2.3.2 与R合作	7

“统计之都”（Capital of Statistics，简称COS）网站成立于2006年5月，其主旨为传播统计学知识并将其应用于实际领域。纵观现今国内统计学理论和应用的发展，一方面我们不难发现统计学在应用领域的巨大潜力—现代管理、咨询、商业、经济、金融、医药、生物等等，无不需要数据的力量，而另一方面我们也不得不承认，国内统计学的应用很大程度上受理论的制约—无论是应用界的人们对统计学基础理论知识的欠缺，还是学术界所研究的理论对应用领域问题缺乏足够的重视。

“统计之都”网站便是基于这样的认识而创建。我们希望，统计理论研究者能充分关注应用问题，而统计应用者也能正确把握统计学基本知识，将统计学这门应用学科真正的潜力开发出来。

第一章 发展纲要

1.1 愿景目标

中国统计学门户网站，免费统计学服务平台

1.2 组织结构

1.2.1 主站

1.2.2 论坛

1.2.3 维基

1.3 制度原则

1.3.1 责任权利

1.3.2 组织文化

第二章 软件说明

2.1 GIT

本文描述了如何使用GIT在统计之都（COS）上合作编写文档，期待众多会员能将自己所学以及在COS的收获分享出来，形成有质量、有水平的小论文或书籍、手册。

GIT是一个合作开发工具，它与版本控制工具SVN有相似之处，但它具有速度快、可离线提交修改等优势。简单来说，我们的日常工作无外乎包括：创建新文档、修改更新、和他人交互；而这个过程在大多数人手里（此处为胡说，无数据无真相）的路径就是：打开Word，写了半页纸，睡觉，第二天继续写了半页，觉得差不多了，邮件加附件发给第二个人，刚发完，想起来有件小事忘了写，于是又打开Word，补上，重新发邮件，于是第二个人就活生生收了两个附件，下载之，读完修改，改完又用邮件发回来，而在这个过程中第一个人发现有地方写错了，并修改过来，此时第二人的新文档已经到了，于是只要肉眼比较两个版本的文档，手工增删改。

邮件发来发去……直到黄花菜都凉了……

版本控制工具本来是为程序开发准备的，比如A写了一段代码，B修改了第3行，C修改了第7-8行，大家的工作在时间上重叠，但内容上可分，此时版本控制工具可以自动合并大家各自的修改，形成新的文档，然后大家各自再从服务器上取回新文档，做自己的修改。把那些肉眼校对的事情交给软件去做而让开发者集中精力写代码—这就是版本控制工具。

总而言之，大家可以各写各的，不必担心冲突。当然，有时候冲突不可避免，此时你也很容易找出冲突所在的位置，手工解决就是了。比如A和B都修改了第3行，如果A坚持认为自己的是对的，那么就把B的修改删掉。

2.1.1 GIT入门

这份文档（英文）描述了在Windows下从安装到使用GIT的过程：

- <http://nathanj.github.com/gitguide/tour.html>

如果想看中文文档，下面这篇PDF文档详细描述了GIT的工作流程：

- http://linuxgem.org/user_files/linuxgem/Image/git-tutor.pdf

2.1.2 GitHub

GitHub是一个提供Git服务的网站（<http://github.com>），注册会员可免费获得300M公开空间，若需要私人空间或者更多服务，则需要额外付费。

- 注册
- 创建密钥
- 创建本地仓库，向服务器提交更改
- 获取他人的工程（通过Fork），如何合并二人的工作？

本文便是基于GitHub协作的一个例子：<http://github.com/yihui/COS>

2.1.3 COS与GIT和 \LaTeX

COS服务器上提供了GIT和 \LaTeX 环境，工作流程为：

- 定期从GitHub取回工程（最开始`git clone`以后都`git pull`就可以了）
- 编译tex文件为PDF（也许会结合R/Sweave）¹
- 自动生成一些更新文件，比如用`git log`生成更新日志

具体的设置待定。

¹目前服务器上的 \TeX Live还没办法编译中文 \LaTeX 文档，所以作者们在自己机器上编好了传PDF吧。

2.2 L_YX

有L_YX的存在，世上居然还有Word横行，真是让人难以理解。L_YX基于L^AT_EX排版，质量一流，而且易用性超强（前提是熟悉L^AT_EX）。

2.2.1 安装L_YX与L^AT_EX

Windows用户最常用的L^AT_EX套装就是MikT_EX，它可以从<http://www.miktex.org>获得；但由于使用中文的原因，我们推荐用C_T_EX，它是一个优秀的本土化的L^AT_EX，主要解决了繁琐的中文配置问题。当然，高手可以尝试T_EXLive等其它L^AT_EX套装自行配置中文环境。简版的C_T_EX就够了，后面需要什么L^AT_EX宏包再另外安装就可以。完全版的C_T_EX打包了所有L^AT_EX宏包，所以体积非常大。

L_YX可从<http://www.lyx.org>下载，注意必须先安装好一个L^AT_EX套装，使得L_YX能找到`latex.exe`的位置才行（安装过程中L_YX会自动寻找）。

2.2.2 L^AT_EX入门

C_T_EX对L^AT_EX的本土化所做的贡献还包括文档的翻译，L^AT_EX初学者可以从“L^AT_EX short”这份文档读起，它本是一份英文文档，叫“一份不太简短的L^AT_EX介绍”。中文翻译还不错，装完C_T_EX之后菜单中能找到PDF文档。

特别警告：一定要对L^AT_EX非常熟悉之后再用L_YX，否则L_YX很容易被用成Word的鬼样子。尤其是要摒弃那些坏习惯，比如手工设置标题的字体字号、手工输入章节号、手工写入图几表几，等等。在L^AT_EX的世界，不要考虑排版问题，大多数问题都是自动的（而且美观的）。如果你还在浪费时间看“Word秘籍”之类的雕虫小技，我等劝君珍爱生命，回头是岸呐。

L^AT_EX本身的麻烦就是不停敲代码呀敲代码，为了写一张幻灯片，得敲上半天各种宏名，有时候实在烦人，觉得做事太低效。L_YX在不损失质量的前提下实现了L^AT_EX的可视化编辑，相当强悍。

2.2.3 配置L_YX

2.2.3.1 中文问题

L_YX本身是基于layout文件设定文档类型的（比如论文、报告、书籍），自带的layout文件在安装目录的Resources/layouts目录下。为了方便使用ctex宏包，我们可以创建几个跟ctex样式对应的layout文件，如定义一个使用ctexart.cls样式（中文论文）的ctex-article.layout：

```
% Do not delete the line below; configure depends on this
# \DeclareLaTeXClass[ctexart]{article (CTeX)}

# Read the definitions from article.layout
Input article.layout
```

然后L_YX菜单：Tools▷Reconfigure，重启L_YX，就可以在Document▷Settings中选择article (CT_EX)了。此时注意文档的语言设置为中文，推荐编码为unicode (XeT_EX) (utf8)，否则L_YX会给你的正文加上cjk环境，这个环境对于ctex宏包来说是多余的；而且用XeT_EX UTF8编码的好处是你可以导出L_YX文档为L^AT_EX文档，然后用XeT_EX编译。选择UTF8编码的时候要在文档类型的选项中填上UTF8，告诉ctex这份文档用UTF8编码处理。

当然，不创建layout文件也可以用中文，L_YX的文档可以设置L^AT_EX导言区（Preamble），在里面写上\usepackage{ctex}也可以。

所有的中文layout文件可以在<http://cos.name/docs/admin/LyX>找到，包括：

- ctex-article.layout
- ctex-book.layout
- ctex-report.layout

2.2.3.2 与R合作

在L_YX中用Sweave也是一件让人激动的事情，想一想，你的正文中只有R代码，当你点一下那个PDF按钮之后，一篇像模像样的统计报告就动态生成出来了……

不过这配置的过程有点让人望而生畏：

- 装R: <http://www.r-project.org> (它是后台做计算和画图的大老板) 建议装R的时候把安装路径中的版本号去掉, 例如直接装到C:\Program Files\R下, 而不要带上R-2.10.1, 原因见后; 安装好R之后把R的bin路径添加到系统PATH环境变量中 (“我的电脑” 右键▷属性▷高级▷环境变量▷PATH, 把C:\Program Files\R\bin之类的路径加到后面, 分号分隔)
- 装Rtools: <http://www.murdoch-sutherland.com/Rtools/> 装好这两个工具之后, 打开一个命令行窗口 (啥? 你不知道什么是命令行窗口? 来人呐, 把这个火星人给我拖出去!), 敲入命令R, 看是否能启动R; 然后敲入R CMD Sweave --help, 看Sweave是否能正常运行。
- 把R安装目录的share\texmf下的文件拷到L^AT_EX宏包的目录下, 比如C:\CTEX\MikT_EX\tex\latex (最好建个目录, 免得乱糟糟), 然后在MikT_EX的设置中刷新文件名数据库 (Refresh FNDB), 使得L^AT_EX能找到Sweave.sty文件。
- 参照CRAN上关于L_YX与Sweave的配置 (<http://cran.r-project.org/contrib/extra/lyx/>) 设置好R和L_YX的合作关系。
- 类似上一节中设置中文的layout文件, 我们也可以设置中文Sweave文档的layout文件。如以下的ctex-article-sweave.layout:

```
#% Do not delete the line below; configure depends on this
# \DeclareLaTeXClass[ctexart, Sweave.sty]{article (Sweave CTeX)}

# Read the definitions from literate-article.layout
Input literate-article.layout
```

- 最后, 为了使中文Sweave文档能顺利运行, 还需设置R的编码环境; 尤其是前面使用UTF-8编码的话, 需要指定options(encoding = "UTF-8") (可以写在R安装目录etc\Rprofile.site文件中)

这样R大致可以和中文Sweave文档合作了。不过还有高级忍术, 需要对L_YX以及R有更深入的理解, 暂且先不介绍了。感兴趣的参见: <http://>

yihui.name/cn/2010/02/misc-issues-in-latex-lyx-r-sweave-pgfsweave/;
这种事情自己能钻研出来，就是说不清楚。