Инструкции по установке Spark-а и всего сопутствующего

Предполагается наличие виртуалки, см. инструкцию из курса по веб-разработке.  
Сначала будет установлен только Spark, с ним буду работать через IPython, для этого буду юзать ещё Anaconda, пока что БЕЗ ВИРТУАЛЬНОГО ОКРУЖЕНИЯ.

Также здесь описан процесс установки Kira.

1. Установка Anaconda
   1. wget UR, взять с сайта последнюю версию
   2. Добвить в .bashrc  
      export PATH="/home/ser/anaconda3/bin:$PATH"
2. Установка Jupiter
   1. Можно создать отдельное окружение (виртуальное), но мы не делаем этого, ибо нет необходимости
   2. Установка Jupyter: conda install jupyter
   3. Запуск: jupyter notebook
3. Возможно, в будущем пригодится: <https://github.com/apache/incubator-toree> - для взаимодействия со Spark-ом
4. Установка Spark
   1. Качаем пакет с Hadoop, я распаковал в Dev/Spark/
   2. Добвить в .bashrc путь, заодно задать SPARK\_HOME  
      export SPARK\_HOME="/home/ser/Dev/Spark/spark-1.6.0-bin-hadoop2.6"

export PATH="$SPARK\_HOME/bin:$PATH"

1. Для связывания Спарка и Юпитера можно юзать пакет findspark, скачаем его с помощью pip
   1. Установка pip   
      sudo apt-get install python-pip
   2. Установка findspark  
      pip install findspark
2. Теперь нужно проверить, что всё работает
   1. Проверка работоспособности Spark  
      pyspark  
      quit() – примечание: ВАЖНО! Не exit()!!!

Установка Kira.

1. <https://github.com/BIDS/Kira> - репозиторий Kira, может быть полезно (обновления если будут)
2. wget <https://github.com/BIDS/Kira/archive/master.zip>
3. Распаковать архив, я распаковал в Dev/Kira
4. Установка Maven, с его помощью Kira собирается:  
   sudo apt-get install maven
5. Если собираем Kira первый раз, то сначала надо установить один jar пакет:  
   mvn install:install-file -Dfile=lib/jfits-0.94.jar -DgroupId="org.esa" -DartifactId="fits" -Dpackaging="jar" -Dversion="0.94"
6. Далее из папки Kira-master:  
   mvn clean package
7. В папке target появится программа
8. ВАЖНО: в папке scratch находится spark версия Kira, собираем её!  
   (при этом выше сборка не факт, что будет нужна, но пусть пока будет)
9. ВАЖНО: либо в SPARK\_HOME указываем путь без bin на конце, либо везде в Kira убираем префикс bin для SPARK\_HOME, пока что выбрал первый путь + в PATH указал $SPARK\_HOME/bin для использования Spark-а отдельно
10. В scala коде есть использование путей к данным в компах авторов, исправить на мои пути к данным (FITS)