Data Integrity

Q: Is it OK to use Excel for data entry?

Data Integrity

- **Q: Is it OK to use Excel for data entry?**
  - Still possible to sort a single column at a time (disaster!)
  - No way of tracking changes made to data
  - Data validation is possible, but limited
  - No way to link related rows (as in a relational database)
  - Often require significant manipulation before or after transfer to stats package

Data Integrity

- **Is it OK to use Excel for data programming?**

Data Integrity

- **Is it OK to use Excel for data programming?**
  - Data programming – any post-entry manipulation of data prior to statistical analysis (e.g., calculating total scores)
  - Unless you’re using macros or VBA, there will be no documentation, or trail, showing what has been done.
  - There is that sorting problem again, as well.
  - Tools like MS Access allow the use of queries that don’t actually ‘change’ the source data.

Data Transfer

- **Q: I took out SSN, address, names, and birthdates, so my file is now de-identified, right?**

Data Transfer

- **Q: I took out SSN, address, names, and birthdates, so my file is now de-identified, right?**
- A: Maybe. All dates, including visit dates, dates of services, etc. must be year only. Ages 90+ can not be there, but must be aggregated to read 90+.

http://www.hhs.gov/ocr/privacy/hipaa/understanding/coveredentities/De-identification/guidance.html#standard

Data Transfer

- **Q: If it’s not a visit registered in hospital software, is the visit date really confidential?**

Data Transfer

- **Q: If it's not a visit registered in hospital software, is the visit date really confidential?**
- A: Anything that could be used to link data back to an individual is potentially problematic, and should be avoided.

Data Security

- **Q: Can data that has been completely de-identified be e-mailed or otherwise shared with colleagues?**

Data Security

- **Q: Can data that has been completely de-identified be e-mailed or otherwise shared with colleagues?**
- A: Sharing may depend on what the original consent form specified.
  - Safe route is to add relevant collaborators to the IRB for the study in question. Then, the data can be shared (SECURELY) with those people.
  - In generally, it’s best to use the ‘need to know’ policy (send them ONLY what they need for their task).
  - Email is a poor method of sharing.

Data Security

- **Q: Is it OK to e-mail a colleague a data file with PHI if I password protect the datafile, and e-mail the password in a separate e-mail.**

Data Security

- **Q: Is it OK to e-mail a colleague a data file with PHI if I password protect the datafile, and e-mail the password in a separate e-mail.**

- A: MS Office passwords can be breached in a matter of minutes via free web tools. NOT SECURE!!!
  - If you have strongly encrypted the file (WinZip or 7-Zip 256-bit encryption), you technically could send via email (however, you would NOT send the password via email!)
  - BUT, it’s more complicated then that. If you are sending PHI data, you need to know how THEY are going to protect it once they decrypt the file.
  - **Email isn’t desirable**.

Data Security

- **Q: I left my laptop in my car and someone broke in and stole it. I had lots of data on there, including PMI. What do I do now?**

Data Security

- **Q: I left my laptop in my car and someone broke in and stole it. I had lots of data on there, including PMI. What do I do now?**
- A: Contact Corporate Security
  - – Ideally, your laptop was encrypted

Data Security

- **Q: When I need to share a datafile with someone, I just give them access to my dropbox. That’s OK, right?**

Data Security

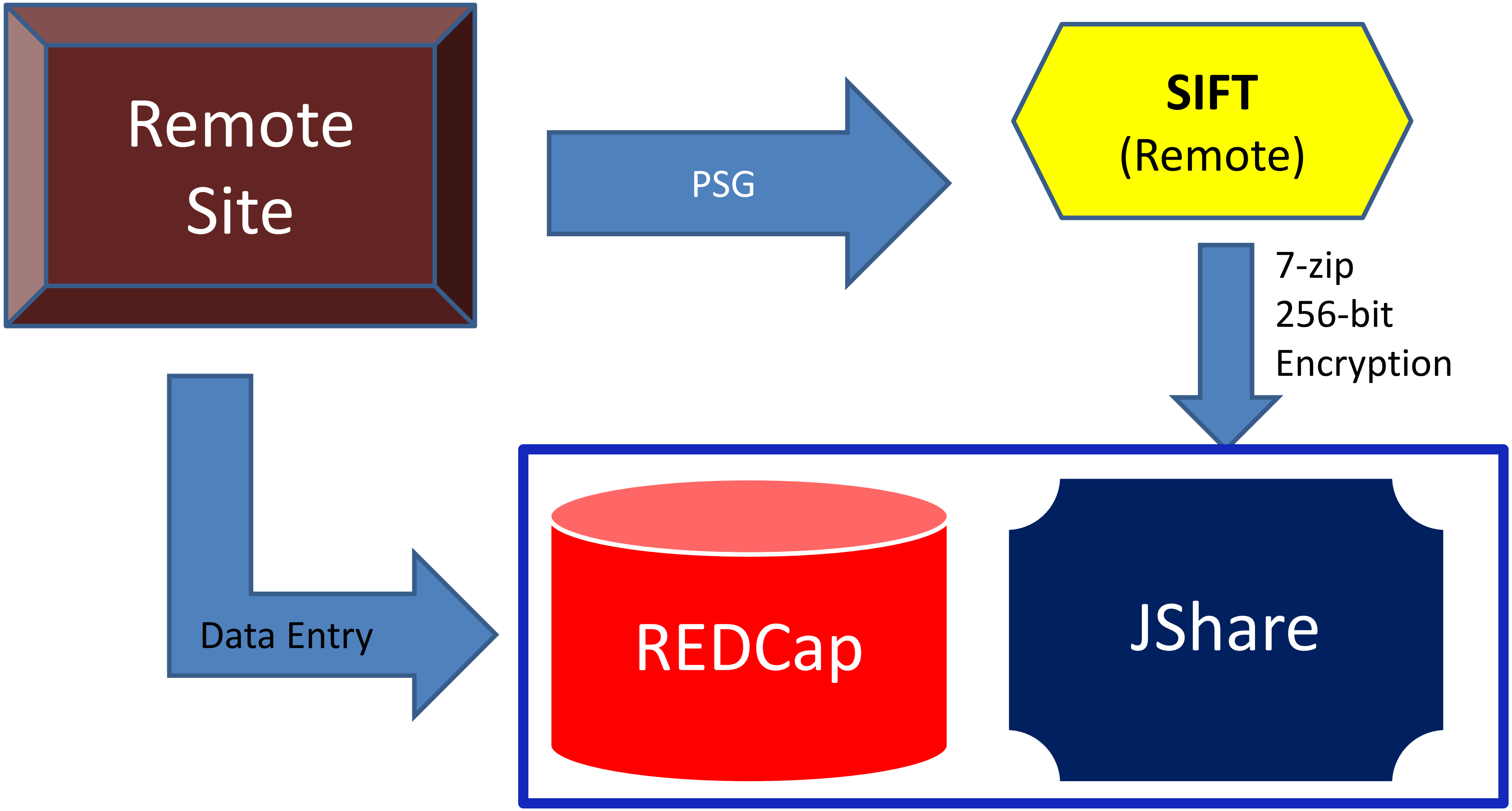
- **Q: When I need to share a datafile with someone, I just give them access to my dropbox. That’s OK, right?**
- A: NO!!! DropBox (and Google drive) is NOT secure. Why?
  - Problem 1: DropBox has access to it
  - Problem 2: In order for you to share it, you must send a link. No authentication is required!
  - There are BETTER options.

Data Security

- **Q: I have a multi-center sleep study that needs to have remote sites submit very large sleep study files to JHU for ‘scoring’. Requires 24-hour turnaround. What is the best way to move the data?**

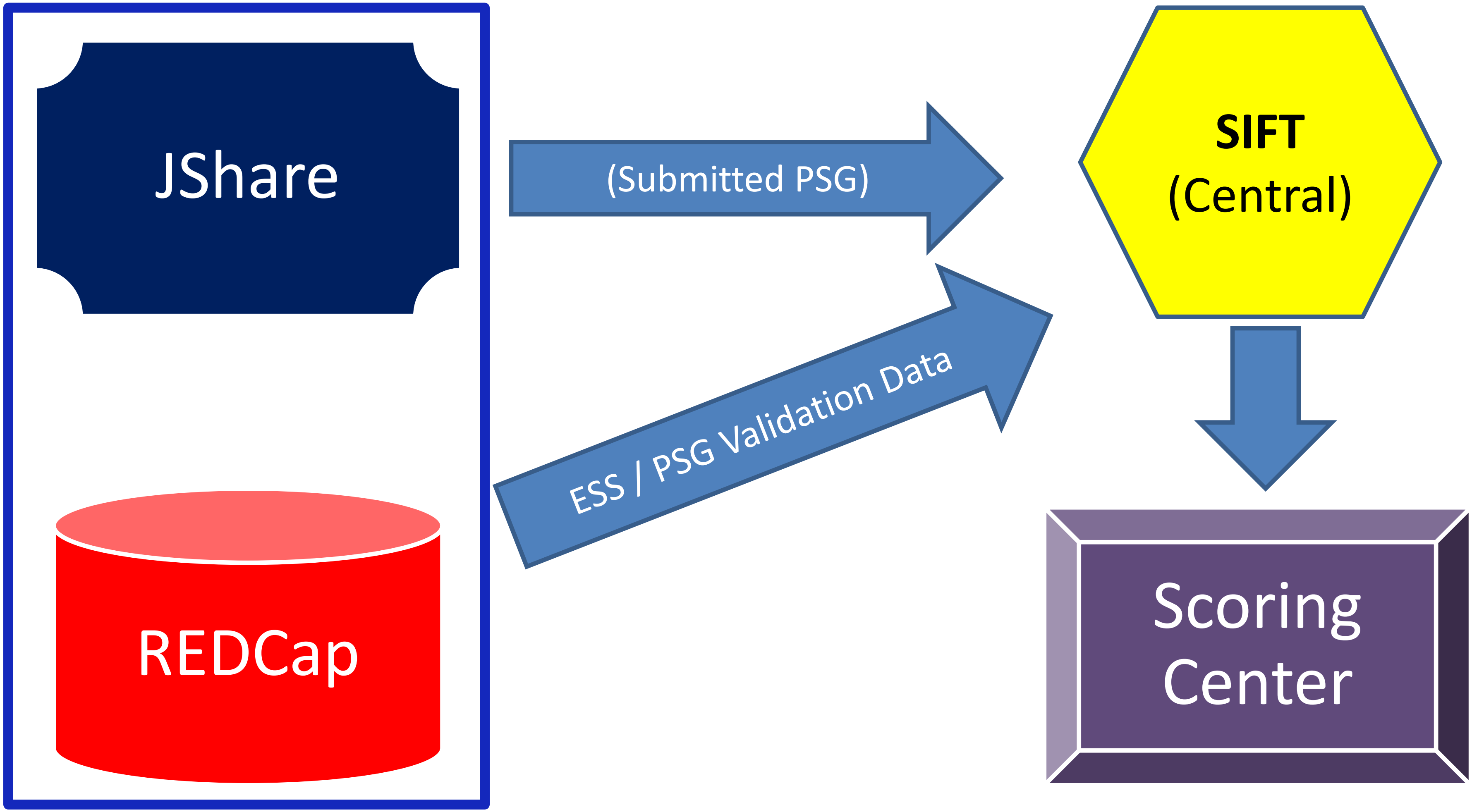
inSleep

Remote Site – Data Collection / Submission



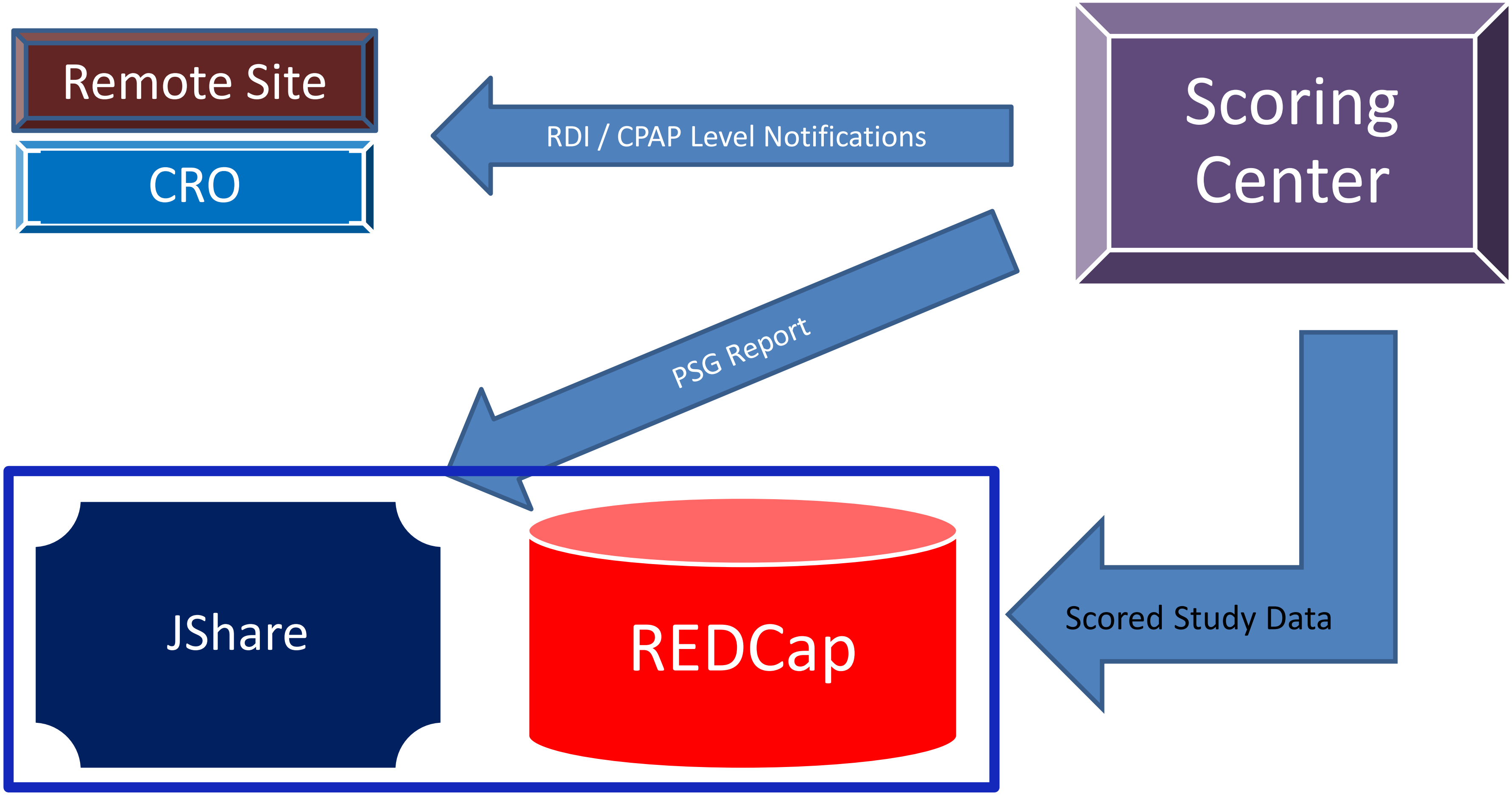
inSleep

Central Site – Queue for Scoring



inSleep

Scoring Center – Scoring / Reporting



Encryption (Software)

- **Q: How do I encrypt a file?**

Encryption (Software)

- **Q: How do I encrypt a file?**
- Windows Zip Compression:
  - Is NOT acceptable encryption
- WinZip:
  - About $30
  - Familiar to most folks
  - 256-bit Encryption
- 7-Zip
  - FREE
  - Easy to use
  - 256-bit Encryption

Data Security

- **Q: If I have to share data with PMI, what is the best way to do it?**

Data Security

- **Q: If I have to share data with PMI, what is the best way to do it?**
  - **GOOD**: Encrypt the file with WinZip or 7-Zip 256-bit encryption. Post to JShare or a suitable 'private' sharing environment (e.g. SpiderOak).
  - **BETTER:** Copy to a HARDWARE encrypted thumb/flash drive and give/send it to them.
  - **BEST:** Create a shared folder on the JHU/WIN Domain where only team members have access. Can access remotely with JHConnect.

Data Security

- Two Areas of Consideration
  - Data at Rest
  - Data In Motion

Data At Rest

- JHU Servers / Workstations
  - Generally secure if on the WIN Domain
- Laptops / Tablets
  - Should use Device Encryption
  - STRONG Passwords
  - Consider device tracking software
  - Consider TrueCrypt
    - FREE
    - Creates a heavily encrypted partition on your hard drive (or thumb drive)

Data At Rest (cont’d)

Portable Storage

- If you are going to use portable devices they should live in an ENCRYPTED state
- CD/DVD’s are convenient for sharing, but tend to get lost and left around. ENCRYPT prior to putting any data on these media.
- Portable/USB Hard Drive – Make sure it’s ENCRYPTED (consider TrueCrypt)
- Flash Drive –
  - Best to use a HARDWARE encrypted version. The entire device is encrypted. Once you authenticate, it works as a normal drive.
  - Consider TrueCrypt for standard USB drives.

Data At Rest (cont’d)

Remote Storage

- JShare is a secure environment. Performance is a bit lacking, but for most applications, it’s a suitable secure alternative for sharing data amongst team members.
- Some remote hosting sites are better than others. SpiderOak does NOT have any access to your encrypted data. If you forget the password… Sorry Charlie.
- Some remote hosting sites are weak (DropBox, G-Drive…). They retain access rights to your data.
- In most instances, you share a link for a file, but there is no password required.

Data At Rest

Sharing Sites

- JShare – the JHU sharing site.
  - Everyone with JHED ID has a JShare folder
  - Can create ‘tickets’ with or without passwords
  - Performance is somewhat lacking
- DropBox
  - Just Say NO!
  - SpiderOak (at least they don’t have access!)

Secure Portability

- Encrypt your laptop!
- Encrypt your tablet!
- Use HARDWARE encrypted thumb drives
  - Imation Defender (bio)
  - Kangoru Defender
  - Kingston DataTraveler
  - Apricorn Aegis (PIN)

Hardware Encrypted Thumb Drives (consider a 'managed' drive)

Closing Thoughts

- Remember… PHI is SELDOM required for analysis. Leave it OUT!
- Just because it doesn’t contain PHI does NOT mean it’s not confidential!

Closing Thoughts

- Try to use a shared folder on the WIN Domain when possible!
- Use JShare until JHU has provided a better alternative.
  - Everyone with a JHED has JShare access
- Use hardware encrypted thumb/flash drives (or an encrypted partition on a standard flash drive [TrueCrypt]).
- Email and CD/DVD are NOT good options!
  - Email is "convenient"... but NOT for sending PHI
  - ALWAYS encrypt data when distributing via email, CD/DVD, telepathy.
  - Just because it doesn't contain PHI does NOT mean it's not confidential!

Helpful Links

- **HIPAA** Privacy Rule: De-Identification Methods
  - http://www.hhs.gov/ocr/privacy/hipaa/understanding/coveredentities/De-identification/guidance.html#standard
- **WinZip**
  - http://www.winzip.com/win/en/index.htm
- **7-Zip**
  - http://www.7-zip.org/
- **TrueCrypt**
  - http://www.truecrypt.org/
- **Eraser** (an EXCELLENT program for securely erasing data on computers, laptops, and portable media)
  - http://eraser.heidi.ie/
- **MS Access**
  - Search YouTube for “MS Access Tutorial” or “MS Access Training”. There are some excellent training tools that are presented in a very visual and clear way. It’s sometimes easier than trying to fit in a ‘course’.