***Data-Science-Production*** *server (priority)*

**Purpose:** to act as an organized data storage and model-fitting environment for ACE data scientists to:

1. **run** final models only(i.e., for incorporation into products such as Flexible Analytics)
2. **archive** final code, metadata, and relevant train/test/output datasets (HIPAA)
3. **interface** with production/engineering teams (e.g., through use of Docker)

**Requirements:**

* Server location must be future proof (i.e., within **Watson Health Cloud** ecosystem) <https://sbybz221073.cloud.dst.ibm.com:9443/ProcessPortal/login.jsp>
* Large storage capacity (to accommodate multiple products with all archived versions)
  + Short-term estimate: **8TB** (for 2018)
  + Longer-term estimate: **32-64TB** (for 2020 and beyond)
* Processing capacity
  + Short-term: e.g., **16-32GB** RAM to facilitate standard **Python/R** environment
  + Medium-term (mid-2018): e.g., above environment with additional linked capacity (e.g., **512GB**) to facilitate **Spark**

**Related info:** While the Data-Science-Production server will contain full code/data archives, ACE data scientists will also store working and final codebases within the Watson Health organization on IBM Whitewater GitHub: <https://github.ibm.com/Watson-Health>

***Data-Science-Scratch*** *server (early-mid 2018)*

**Purpose:** toact as a team sandbox environment for model development (limited data storage).

**Requirements:**

Moderate storage capacity (to accommodate 20-30 users on ongoing basis)

* + Estimate: **8-16TB** (e.g., average per-user limit of 300-500GB)

Processing capacity

* + Estimate: **16-32GB** RAM to facilitate **Python/R** environments