

1. Project title
  - UCPATH Reporting Instance Data Extraction Framework
2. Submitter's name, title, and contact information
  - Walter Stokes, Director, Data and Platform Services, UC Berkeley, [walter@berkeley.edu](mailto:walter@berkeley.edu). 510-664-4084
3. Names of project leader(s) and team members
  - Quin Bligh (UCB)
  - Carolin Mathew (UCOP), Gregory Boyer (UCOP), Vivek Vennam (UCOP), Joe Lavu (UCOP), Rajani Prakash (UCOP), Jay McGehee (UCOP)
  - Victor Kalchev (UCSD)
4. A simple short statement (even just one sentence) summarizing what the project does
  - The RI Extract Framework (also known as Data Pump Framework, or DPFW) pulls data from the AWS based UCPATH Reporting Instance, and moves it to different UC campuses, where the data can then be imported into a local database, for the specific campus to use for localized reporting.
5. A project narrative that:
  - Describes the problem being solved and the project goals
    - The UCPATH system pushes data into a Reporting Instance (RI) for reporting purposes. Individual campuses wanted to pull data from the RI for local reporting purposes. Recognizing that one solution would benefit all the UC Campuses, Quin developed an extraction process for any campus that has a local Oracle database as a target, then extended it to produce output that was relevant for other databases as well. (i.e. SQL Server)
      1. Goal 1: Create a single extraction mechanism that would benefit all the UC Campuses, while minimizing the amount of duplicative work required of each campus.
      2. Goal 2: Recognizing that not all campuses are alike, develop a mechanism that will feed different target databases, i.e. both Oracle and SQL Server.
  - Emphasizes the solution and innovation, rather than technical detail
    - The UCPATH system moved from an Oracle Data Center to Amazon's AWS environment, and replicates data into a Reporting Instance (RI) for reporting purposes. In the previous iteration of UCPATH, all of the UC Campuses were receiving their specific campus' data via a data loader mechanism. With the change to AWS, instead of the previous method of pushing data to the campuses, UCOP opted to change to a pull methodology, which meant that all the campuses would have to create individual solutions to pull data from the RI.
    - Realizing that building **one good solution** that could be used at **all the campuses** would be preferable than building **10 different solutions**, Quin created this extraction framework, and collaborated with different personnel at UC Berkeley, UC Merced and UC Santa Cruz to test the solution.
  - Tells how the solution has impacted customers/users

- This project has allowed all of the UC campuses to use this extraction process created once, and NOT have to create one on their own.
  - It has greatly facilitated the local reporting process at each campus, allowing the functional business analysts to receive that campus' UCPATH data, without having to wait for a customized extract solution to be built.
  -
- Explains how project success was measured
  1. Adoption of DPFW by campuses (August 2020 - April 2021). All but one campus has gone live in production with DPFW.
- Highlights collaborations with other locations, departments, or teams
  1. UC San Diego is contributing a solution for AWS/Oracle-RDS that utilizes the DPFW as the core engine.
  2. Campuses provided feedback on functionality and reported defects.
  3. Early adopters (UCR, for example) assisted in on-boarding new campuses, greatly facilitating the implementation.
- Provides the timeframe of deployment
  - Quin created and deployed this solution over the course of 8-10 months in 2020, in the midst of the COVID pandemic to boot.
- Briefly describes the technology utilized
  - UCPATH is a Peoplesoft HCM ERP, which was migrated from Oracle Managed Cloud Services to Amazon Web Services (AWS). The Reporting Instance (RI) is an Oracle database, replicated from the primary instance via Oracle Golden Gate. The extract framework pulls changed data for each campus into a staging schema, and then exports data from the Reporting Instance using Oracle's Data Pump. This output goes into both an Oracle format (for subsequent import into an Oracle database) and a csv / delimited format (for subsequent import into a non-Oracle database. (i.e. SQL Server)
  - This flexibility allows different UC campuses to be served by this framework; for example the target database at UC Berkeley is Oracle. The target database at UCLA is SQL Server, but both campuses can use the same extract framework, cutting down on repetitive work.
  - In addition, the framework has been changed so that it supports target databases that reside in AWS as well. Quin collaborated with Victor Kalchev at UCSD in order to extend the framework, and create a wrapper around the DPFW that will allow it to interact with target databases in AWS' RDS service
- Optional: Provides relevant screenshots or urls (included in the 5 page limit)

