State Enrollment Reporting and Data Quality



Data Services

March 2019
Update September 2020
Paula McDaniel, Data Analytics Learning and Education Instructor
Washington State Board of Community and Technical Colleges, Data Services



Table of Contents

Course Description	1
Course Goal	1
Course Objectives	1
Data Quality	3
What is Data Quality	3
Data Quality Dimensions	3
Accurate Data	3
Consistent Data	3
Complete Data	4
Unique Data	4
Timely Data	4
Valid Data	4
Why is Data Quality Important	4
Data Flow	5
Data Re-Engineering Process	7
Accuracy Double Check	7
The Data Warehouse Key	7
Derived Fields	7
What Enrollment Data is Being Reported to SBCTC?	7
Tickler Applications	7
Using the Legacy Tickler Application	8
Legacy Tickler Process	8
PeopleSoft Tickler	8
PeopleSoft Tickler Process	9
The Coding Manual	9
Using the Coding Manual	9
Revisions	10
Legacy and PeopleSoft Coding	10
Sections	11
Coding Elements	12
Course and Class Attributes	13
Student Groups	14
Student Attributes	15
FTE tables in Campus Solutions	16

Checking the Data For Errors	16
Quality Assurance Reporting System (QARS)	16
QARS and SBCTC Master Data Views	16
Using the Quality Assurance Reporting System	17
Logging into QARS	17
My Reports	17
dataLink and GoldenGate	20
Near real-time production data	20
Adding Tables Which Will Be Replicated by DataLink	20
dataLink, GoldenGate and metaLink	20
Legacy Data Access	20
MetaLink Data Dictionary	20
Logging into the metaLink Data Dictionary	21
Using the metaLink Data Dictionary	22
Record Information	23
Field Information	24
MetaLink Query, Report and Pivot Grid Search	25
Search Criteria Boxes	27
All Pillars	27
Modules	27
Folders	27
Field Names	28
Free Text Search	28
Combine Search criteria	29
Search Results	29
Queries Tab Results	30
Reports Tab Results	32
Pivot Grids Tab Results	34
View Script Search	35
Resources	36
Adding Information to MetaLink	37

COURSE DESCRIPTION

This course is designed to impart knowledge regarding coding and compliance with a focus on data quality with the intent to teach users how to use the available applications and documentation to ensure accurate coding and hence, accurate State and Federal reporting. This course will cover:

- Data Quality What it is and Why it's Important
- Data Flow
- The Coding Manual
- FTE Calculations
- Quality Assurance Reporting (QARS)
- DataLink
- State Board Master Data Views and the Data Warehouse
- metaLink

COURSE GOAL

The goal of this course is for participants:

- Have a solid understanding of why Data Quality is so important to us, our institutions and our students and why it is vital that everyone who touches the data plays a part in maintaining its quality.
- To gain the knowledge and skills necessary to work within the boundaries of the coding manual
 to deliver complete and accurate data for State and Federal reporting whether it be delivered
 from Legacy or gathered from PeopleSoft.
- To define how the coding of classes and students will differ from Legacy to PeopleSoft, to fully learn how to use the applications and documents provided by the State Board in order to successfully manage and maintain accurate data.
- To provide a high level view of the processes and applications affecting the data used in State and Federal reporting.

COURSE OBJECTIVES

At the end of this course participants will be able to:

- Explain the definition, components and importance of Data Quality.
- Accurately define the flow of how Legacy data sent by the colleges is received by the State Board and how PeopleSoft data is extracted by the State Board.
- Describe how the data is re-engineered in order to set the derived data elements necessary for calculating FTE's.
- Summarize how the Legacy and PeopleSoft Tickler applications are used to inform the correct people of upcoming jobs that need to be run and when data is due to be released.
- Define the purpose of the Coding Manual.
- Describe how the Coding Manual is set up.
- Successfully recognize PeopleSoft coding elements including Class Attributes, Student Groups and Student Attributes
- Describe how the values within the coding elements affect State and Federal reporting.
- Summarize the difference between local and global coding standards in PeopleSoft coding.
- Explain the data governance process in the creation of local PeopleSoft codes.

- Describe how MIS Edit Checks works to help Legacy colleges ensure quality data.
- Define how QARS can help benefit each PeopleSoft institution in maintaining quality data.
- Define how State Board Master Data Views function and allow for QARS reporting.
- Explain how the State Board Master Data Views are the only source of PeopleSoft data for State and Federal reporting.
- Log in and use the QARS reporting system.
- Explain the advantage and value of dataLink .
- Define dataLink.
- Relate the process of adding tables to be replicated by dataLink.
- Explain the relationship between GoldenGate, dataLink and metaLink.
- List the 4 areas of metaLink users can access to find information.
- Detail the two main sections of the metaLink data dictionary .
- Explain the concept of metadata.

DATA QUALITY

With any system the quality of the data pulled out will only ever be as good as the quality of the data entered in. We've all heard the saying "garbage in = garbage out". The reverse is, of course, also true... "excellence in = excellence out". Washington's College system is striving for excellence in all things, including data quality. Due to the importance of the data being reported, excellence in data quality is paramount in achieving timely and accurate reporting.

What is Data Quality

What is "Data Quality?" According to Thomas C. Redman in his book: Data Driven: Profiting from Your Most Important Business Asset. "Data quality refers to the condition of a set of values of qualitative or quantitative variables. There are many definitions of data quality but data is generally considered high quality if it is "fit for [its] intended uses in operations, decision making and planning." (Redman, 30 December 2013). What is the intended use of the data generated by the college system? While it is also used for other purposes; one of the primary uses of PeopleSoft and Legacy data is for State and Federal reporting which is how funding for colleges is established. It can't get more important than that! Our students rely on us to report accurate and consistent data in order to receive the highest amount of consideration possible to use for their benefit.

Data Quality Dimensions

The State Board is in the process of implementing data quality reviews and will be using Data Quality Dimensions to achieve this goal. Data Quality Dimensions is the term that is used to describe an attribute/feature/characteristic of data that can be measured or assessed to determine data quality. Some are mandatory and others are optional depending on the type of data being measured.

The picture to the right shows some of the Data Quality Dimensions that are used to gauge data quality. This list is not exhaustive but it is comprised of generally accepted Data Quality Dimensions. For this course we will discuss the following Data Quality Dimensions:

- Accuracy
- Consistency
- Completeness
- Uniqueness
- Timeliness
- Validity

Accurate Data

The definition of accurate data is "the degree to which data correctly describes the "real world" object or event being



described". ¹ For example, if a European writes his birth date in the European date format, this would be a different date entirely in the United States and not a correct description of the student's date of birth and thus, inaccurate. Accuracy is determined by taking data from a source that has been deemed trustworthy and comparing it to the extracted data. Accuracy is mandatory because if data is inaccurate if may not be "fit for use".

Consistent Data

Consistency is determined by measuring the data against itself or a counterpart in another record or database. It is defined as "the absence of difference, when comparing two or more representations of a

¹ "United Kingdom, Dama UK Working Group, The Six Primary Dimension for Data Quality Assessment (October 2012), 12"

thing against a definition".² To continue with our birthday example we would compare the birthday from one record to that from another record to determine if it is indeed the same in both records. If they are the same then the data is consistent. Consistency is mandatory, however, keep in mind that it is entirely possible to have consistent data that is neither valid nor accurate.

Complete Data

Complete data is defined as "the proportion of stored data against the potential of 100% complete" It is determined by measuring either the absence of blank or NULL fields or by measuring the presence of blank of NULL fields. If a field is required then, of course 100% completeness will be achieved. In that case accuracy and validity would still need to be checked. If the field is not mandatory then completeness is optional.

Unique Data

The definition of Unique Data is that "no thing will be recorded more than once based upon how that thing is identified". If there is a student "John Smith" and a "Johnny Smith" with 2 different EMPLID's and the same address despite there being only one "John Smith" enrolled at an institution then the data being used is not unique and would need to be addressed. Uniqueness can be optional or mandatory depending on the situation.

Timely Data

Timely Data is defined as "the degree to which data represents reality from the required point in time".

⁵ It measures the difference in time between required entry into the system and the event happening.

For example, if a student submits an updated emergency contact number on June 4th, 2017 and the new number isn't recorded until June 8th, 2017 and there is a service level agreement for 2 days then the breach of the agreement is 2 days. Note: it can only be untimely if there is a rule as to what constitutes timeliness. Thus, Timely Data can be optional or mandatory depending upon the situation.

Valid Data

"Data is valid if it conforms to the syntax (format, type, range) of its definition" ⁶ It compares the rules for a field against the entered data. If a field requires 2 letters and 2 numbers with no spaces and ZZ11 is entered then it meets the requirements for the field as a valid entry. It may not be accurate, consistent or unique, but it is valid. Validity is mandatory.

Why is Data Quality Important

Though, per the SBCTC policy manual, the Registrar is ultimately responsible for the quality of the data, it is very important for all users to work together to ensure accuracy. If a user is entering data into a system, some things to ask are:

- "How is what I'm entering affecting the funding for students and programs at my institution?"
- "How does what I'm entering affect other users and reporting in the system?"
- "What is the ultimate purpose of the data I'm entering?"

² "United Kingdom, Dama UK Working Group, The Six Primary Dimension for Data Quality Assessment (October 2012), 13"

³ "United Kingdom, Dama UK Working Group, The Six Primary Dimension for Data Quality Assessment (October 2012), 8"

⁴ "United Kingdom, Dama UK Working Group, The Six Primary Dimension for Data Quality Assessment (October 2012), 9"

⁵ "United Kingdom, Dama UK Working Group, The Six Primary Dimension for Data Quality Assessment (October 2012). 10"

⁶ "United Kingdom, Dama UK Working Group, The Six Primary Dimension for Data Quality Assessment (October 2012), 11"

- "What steps can I take to ensure that the data I enter is accurate and complete?" If the user is not responsible for entering data but perhaps the user's job description entails using or reviewing information, the user can still participate actively in ensuring data quality.
 - Does the data look like it is correct?
 - Are the numbers logical?
 - Compare numbers to what was seen last year.

There are multiple methods that can be used to help ensure accurate, complete, consistent, unique and timely data. Why does it matter? Darby Kaikkonen from Internal Research at the Washington State Board of Community and Technical Colleges shares a story below of a real life example of why correct coding is so important. The bottom line is it comes down to funding for the colleges.

In our performance funding system we award points and subsequently dollars on the basis of a couple of different things. One way is a specific dollar amount for the total number of points earned, and the other way is a dollar amount for the number of points earned per student. The latter metric is designed to reward efficiency in getting as many students through as many milestones/points as possible. The point per student metric requires careful attention to coding because of the denominator issue. We don't want to include students in that denominator who have no intention of earning those milestone points. Examples of these students are those who are taking one or two non-credit/non degree-related courses, those in parent education classes, and those who take state-supported "senior" classes. To effectively manage that denominator the state board came up with coding rules for which students could be removed from that calculation. It is in a college's best interest to have proper coding on students for those scenarios so they don't inadvertently inflate that denominator, thereby causing them to have a low ratio.

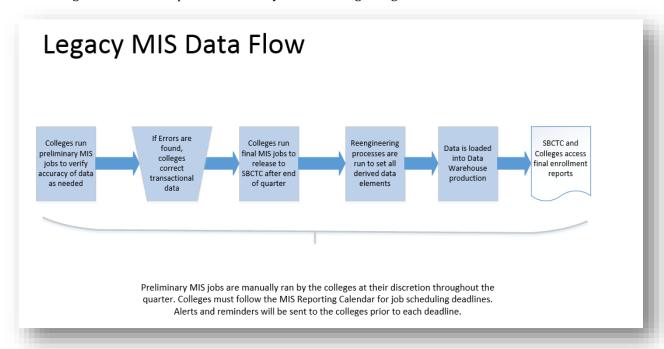
Where this is especially important is in the funding calculation. Everything is awarded based on a share. When it comes time to calculate the awards, we look and see what college's share of the total is, and that is the percent of the funding they earn. So for example if a college had 5% of the total points per student metric, they would get 5% of the money. This is the reason why if there is a coding error it is not possible to go back and fix it. If a college believes there was a mistake and their share should have been higher, that means the difference in what they got and what they think they should have got has to come out of all of the other college's shares. Once the allocation goes out this is not reasonable or right to do. Because of the importance of coding, when we made changes to SAI (performance funding system) a couple years ago we held multiple workshops for the colleges to understand the implications. One college had students that fit the criteria for being excluded from point per student, but they did not have the correct intent code on them. Because of that we did not pull them out of the calculation, thereby causing their ratio to be very low. This resulted in a lower amount of money had they had their coding correct. In follow up analysis we calculated that had the students been correctly coded it would have resulted in about \$7,000 more in their allocation. Not a tremendous amount of money, but the bottom line was that we did not recalculate and give them an extra allocation because the money had already been allocated to other institutions. They are fortunate that it wasn't more, and I think that is a good moral of the story. Coding is so important because of how interlinked everything is and it is not possible to go back and fix things after they are finalized, even if it was a million\$!

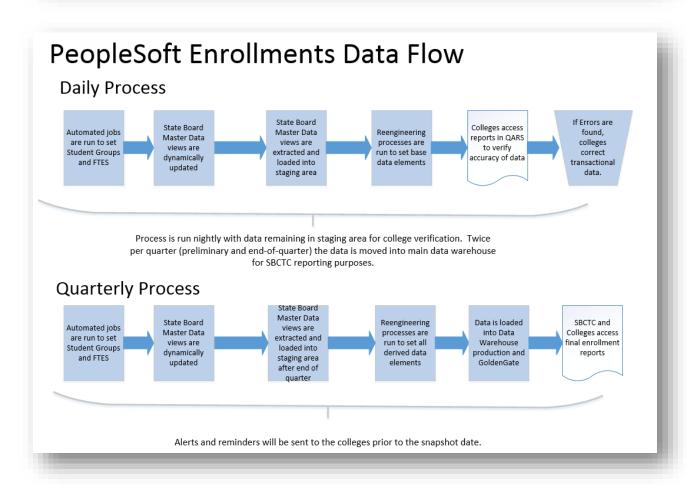
DATA FLOW

How does data get from a college to the State Board? What happens to the data once it arrives at the State Board? The data flow is slightly different depending on whether it is coming from Legacy or PeopleSoft however the basic principle is the same. The data is provided by the college and re-engineered at the State Board then loaded into the data warehouse for use in State and Federal Reporting. In the legacy system when the college manually runs the MIS finals process, a dataset is processed and transferred to the State Board. In PeopleSoft there is no job that processes a dataset. Instead, Data Services

transferred to the State Board. In PeopleSoft there is no job that processes a dataset. Instead, Data Services takes a snapshot of the full student, class and enrollment data. In both systems the data is extracted after the end of the term being reported. The re-engineering process that happens at the State Board adds data elements to simplify reporting and research but does not in any way, shape or form change the data being reported by the college whether it be in Legacy or PeopleSoft. Instead what the re-engineering process does is double check the data for accuracy, DW Key, derived fields.

The following charts detail the process in each system from beginning to end.





Data Re-Engineering Process

What happens to the data once it is received by the State Board? Whether received from Legacy or PeopleSoft the data must eventually make it to the Data Warehouse where it will be used for State and Federal as well as historical reporting. To do this the data is put through a re-engineering process. The re-engineering process that happens does not in any way, shape or form change the data being reported by the college. Instead, the purpose of the re-engineering process is to double check the data for accuracy, assign the Data Warehouse, or DW Key which is explained later in this chapter and add data elements to simplify reporting and research.

Accuracy Double Check

First the data is double checked for accuracy. If a record comes back as state funded and part of corrections course then there is an error and, if time permits, the data will be sent back to the college for fixing or Data Services staff will work with the college to manually correct it.

The Data Warehouse Key

The Data Warehouse Key or DW Key is the unique identifier used in the data warehouse for students. It is designed to be unique to each person based on the student's SID/EMPLID or social security number to facilitate longitudinal analysis which is basically tracking students across institutions and over time.

Derived Fields

As the data is being collected into one database from multiple systems and is being consumed by multiple parties (Data Services, IR, etc) there are instances where some elements must be derived in order to maintain consistency and cohesiveness. For example, in Legacy race is tracked as a code of 1 or 2 with a hispanic indicator while in PeoplesSoft each race is listed with a checkbox and multiple selections are allowed. To combine these in the Data Warehouse, data elements are added to allow for derived fields to be used to create homogeneous data. Again, the original data is not changed in any way.

What Enrollment Data is Being Reported to SBCTC?

Enrollment data extracted includes all classes offered during the quarter with one or more students enrolled as of the census date. The census date can vary depending on the type of class. Even though the data is reported after the end of the quarter, it includes any students who are enrolled as of the census date for the course, whatever that census date may be. It is NOT only reporting on who is enrolled at the end of the quarter, unless that happens to be the census date for the course. Students who drop the class before the census date are NOT included, but students who withdraw from the class after the census date ARE included.

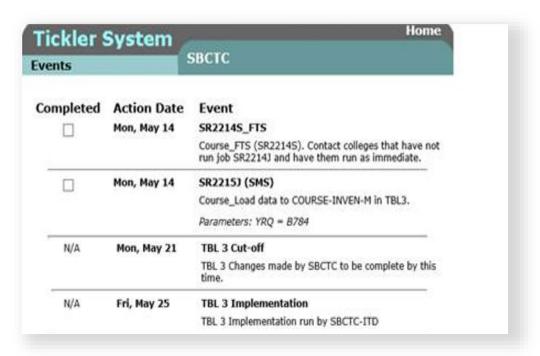
Tickler Applications

A Tickler is an application that is set up to act as a reminder to perform a certain task. At the State Board we have set up two Tickler apps to assist colleges in managing the data flow and reporting to SBCTC; one for PeopleSoft and one for Legacy. Both apps include reminders for program, course, class, student, and personnel related jobs and data. All reminders are tied to events. Some events are mandatory which require the user to manually mark them as complete and some are informational and do not require any response from the user.

The colleges have traditionally provided two contacts to be "tickled" with these reminders. Generally there is one person on the Student Management System (SMS) side of things and one on the Payroll/Personnel Management System (PPMS) side. However, with PeopleSoft it is possible to add more than two contacts to receive the tickler email messages.

Using the Legacy Tickler Application

The Legacy Tickler app is used exclusively by Data Services for reminders. In Legacy a mandatory event is an action that must be performed, such a running a job to create an extract to send to SBCTC Data Services. Mandatory events have a 'Completed' checkbox to the left of the Action Date in the Legacy Tickler app. Information events are marked N/A and do not have a 'Completed' checkbox.



Legacy Tickler Process

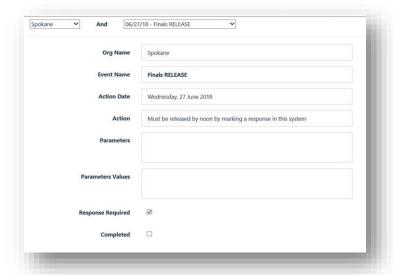
- The Tickler sends a reminder email to both designated college contacts for the event.
- If the event is required, the colleges complete the task and mark off in the app that the event is complete
 - The Legacy Tickler is set up to send:
 - An initial reminder
 - A follow up reminder
 - A final reminder
- If events are not marked complete Data Services will touch base with the designated contacts to ensure they have completed the required tasks.

PeopleSoft Tickler

While the PeopleSoft Tickler also has at least two contacts designated by the colleges, there is a hierarchy. The first person tickled is normally someone in Admissions. The second contact is tickled if the first does not mark the event complete. The second contact is normally the Registrar. Additionally, it is possible to add more than two contacts per college, if desired.

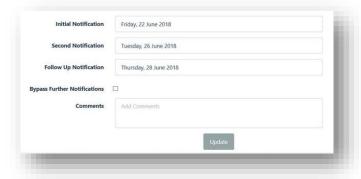
The PeopleSoft Tickler was designed so that it could be used by both the Data Services team and/or the Functional ERP team. While the ERP team may use the Tickler to remind colleges about jobs that need to be ran, Data Services uses is exclusively to remind colleges to prepare their data for upcoming snapshots. So, in the scope of this course, a mandatory event does not require the college to run a job in the system as it does in Legacy, instead the college will ensure that the data is clean and ready to be snapshotted by SBCTC. If the event is mandatory, there will be a check in the checkbox 'Response Required'. If it is informational the checkbox will be blank. Once the data is ready for extraction the college will check the 'Completed' checkbox in the application.

The PeopleSoft Tickler application is closely related to the Quality Assurance Reporting System (QARS) which allows colleges to pull reports from PeopleSoft that help find incorrect or incomplete data. See the <u>QARS</u> section of this manual for more information.



PeopleSoft Tickler Process

- If the event is required, the colleges prepare their data and mark off in the app that the event is complete and the data ready to be extracted.
- The PeopleSoft Tickler is set up to send:
 - An initial reminder to the first contact.
 - A second reminder to the first contact.
 - o A follow up reminder to both contacts.
- If events are not marked complete Data Services will touch base with the designated contacts to ensure they have completed the required tasks.



THE CODING MANUAL

The Coding Manual was developed to assist colleges in coding programs, courses, and students for State and Federal reporting purposes. These codes are used for FTE calculations as well as allocation of funds to each institution. It is extremely important to get the codes right to ensure accurate State and Federal reporting and the manual has been designed to help with that.

Using the Coding Manual

The coding manual is an online resource found on the SBCTC website. The format is a PDF and it can be found here: Washington State Board for Community and Technical Colleges Student and Course Coding Manual

(https://www.sbctc.edu/resources/documents/colleges-staff/data-services/data-warehouse/sbctc-student-and-course-coding-manual.pdf)

There is a Table of Contents which allows for quick and easy navigation to each section.

Revisions

A history of each revision is found at the very beginning of the manual to simplify the process of keeping track of any coding changes.

		Revised coding for identifying PLA Credit by Exam in PeopleSoft
9/13/16	2.1	Revised Continuing Education section to include "W" Intent Added a Student Intent code section
1/23/17	2.2	Updated the Visa Codes for the International Student (SINT) Student Group criteria Updated the CIP Codes for BEdA and WABERS+ Added the coding for co-enrolled WRT and WF students Added additional instructions for entering DOC numbers Added Vet Benefit Code changes
7/19/17	2.3	Removed "Drop-Out" from the High School Reengagement title Updated OER description and definition Added additional PeopleSoft coding information to Worker Retraining and WorkFirst sections

Legacy and PeopleSoft Coding

An important detail of the Coding Manual is that it includes both Legacy and PeopleSoft coding. This allows users of any system access to coding information that will help them to ensure data that is:

- Accurate
- Consistent
- Complete
- Unique
- Timely
- Valid

As the data in the data warehouse spans many years, the legacy coding will eventually be used as a historical perspective once all institutions are live on PeopleSoft.

It is also designed to assist PeopleSoft users in the transition to the new system and new processes.

Below is an example of both PeopleSoft and Legacy coding found in the Coding Manual.

Туре	Legacy Fee Pay Status	PS Waiver or Residency Exception	
Senior Citizens – Audit	10	WI0	
Unemployed Waiver	21	W2I	
Underemployed Waiver	22	W22	
Native American non-resident waiver	25	SF02	
State Employees	51	W5I	
Senior Citizens - Credit	56	W56	

Sections

The Coding Manual includes general coding for classes, students and enrollments such as required and intent coding. It also includes specific program coding such as workforce and basic skills as well as student specific coding like veterans or foster-care coding.

Each section provides a holistic coding description of individual program areas or specific types of students. It goes well above and beyond the basic information of "use this code here". But instead, describes how the code used impacts how the data is reported and the reasons and requirements why a particular code should be used. Information that may be included in each section is:

- Calculations
- Descriptions
- PeopleSoft Coding
- Legacy Coding
- All applicable Program Coding to include: Student, Class, Program, etc.
- Any other pertinent information

Below, is an example taken from the Coding Manual on how FTE's are calculated.

FTE Calculation

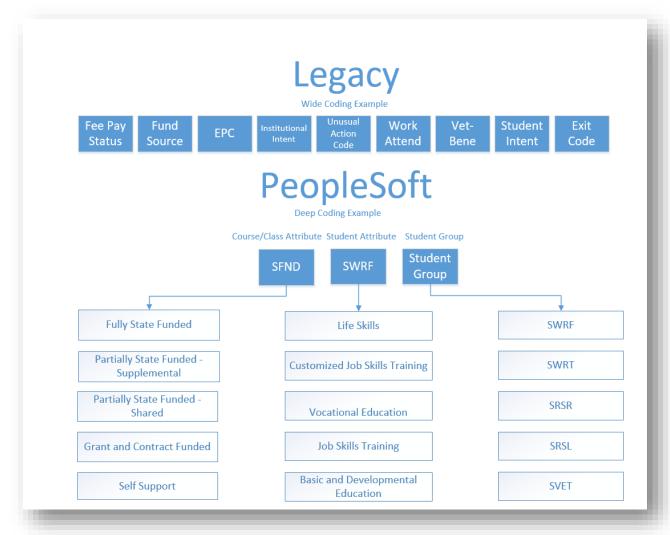
The formula to calculate FTES (full-time equivalent student) for reporting is:

FTES = Census-day Enrollment x Credit Equivalent ÷ 15 The divisor of 15 is constant.

FTES are calculated for each fund source enrollment category, as ALL enrollments (enrolled courses) during the quarter should generate an FTE value and be associated to a fund source enrollment category.

Coding Elements

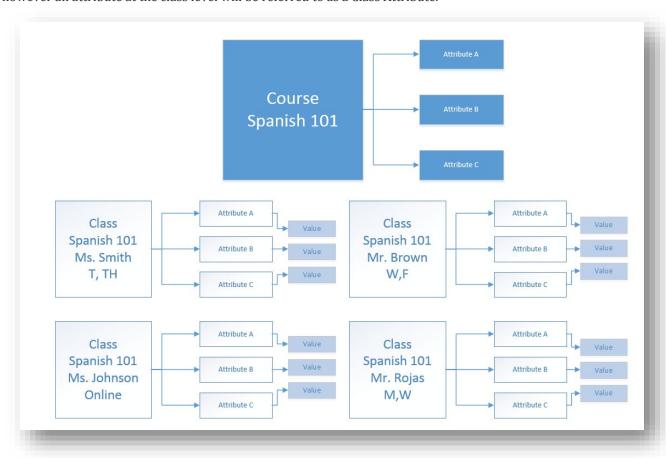
There are three main coding elements for PeopleSoft enrollment reporting, Class Attributes and Student Groups/Attributes. These replace the multiple coding elements available in Legacy. Think of the difference between the two systems as though Legacy has very wide coding capabilities with many, many data elements while PeopleSoft is has very deep coding capabilities with only three main coding elements which have unlimited values.



Course and Class Attributes

Class attributes are general characteristics that define the course and class offerings. They are primarily used for state reporting and institutional research purposes. Though the attributes that are pulled into the data warehouse are "called" course attributes, in reality it is the class attributes which are used. A course will have many classes under it and each class will pull the attributes of its "parent" course while still allowing separate attributes to be added to the individual class. Remember, outside of the State Board, such as in PeopleSoft documentation or online, they will always be referred to as course attributes. However, the ONLY attributes which are used by the Data Warehouse will be at the Class level, or the Class Attributes.

The attribute at the course level will be referred to by the Data Warehouse team as a Course Attribute, however an attribute at the class level will be referred to as a Class Attribute.



When attributes are configured, attribute values can also be specified. Therefore, parent attributes can be defined that have one or more attribute values. An example of a class attribute is fund source (SFND) which has various attribute values.

Class Attribute:	Description:
SFND	Funding Source

Class Attribute Values:	Description:
1	Fully State Funded
2	Partially State Funded - Supplemental
3	Partially State Funded - Shared
4	Grant and Contract Funded
5	Self-Support

It is actually the <u>value</u> of the class attribute which defines the characteristic of the class. Though there are some course attributes which do not have values defined such as OER, the majority have values. This allows a single attribute to have a much wider scope.

Course Attributes and State Reporting

The State Board has set up special course attributes to be used for State reporting and fund allocation. All of these State Board course attributes will begin with "S" so each institution will know that these course attributes need to be added to the course in order for it be picked up by the system correctly.

Global vs. Local Course Attributes

All course attributes are global meaning that every institution would see and have to sort through all other institutions "local" attributes. Yet, because the value is what defines the attribute there is a simple solution to this issue. Each institution will have their own course attribute and then be able to request the creation of their own values. It is highly recommended that each institution have a data governance committee with an appointed representative to make the request for the new value after it has been approved by the committee. This will allow for the naming conventions used in the creation of new values to stay on the same path and not veer off to where there is just a huge list of values not in any particular order. For example, if Tacoma has the course attribute of 220 they could then create codes to represent particular student paths such as PATH_A, PATH_B, PATH_C, etc. This would allow for grouping of like values and ease of use for end users.

Student Groups

Student groups are used for several functions in PeopleSoft. Some student groups are informational, so that users can identify specific student characteristics. Other student groups are attached to functions like financial waivers. The purpose of student groups is to designate a group of students by demographic, program, and/or enrollment information for reporting purposes.

Student Groups and State Reporting

Student groups are not associated with a specific Year and Quarter, so students will need to be deactivated once they are no longer in the program. Students must be deactivated in the quarter in which they should no longer count. If this process is not followed it could create errors in State Reporting.

Student groups starting with "S", for System Defined, have been created by SBCTC and are used for state-level reporting. Some student groups are "mass assigned", or assigned automatically, based upon student characteristics or information gathered at admissions. Others are "manually assigned" by college staff. More information regarding student groups can be found in the <u>Coding Manual</u> section of this course. Because these codes are for the purpose of State-level reporting when using a code beginning with "S" it is required that the usage meet the required purpose and definition of the code.

Global vs. Local Student Groups

Student groups are local in nature, meaning that they are tied to an institution code so each institution will only see student groups tied to their own college. However, every college will also have State Board defined student groups which will begin with 'S' that are used with State and Federal Reporting. Colleges are able to request the creation of these local student groups in addition to the special State reporting student groups. These local student groups can be used for purposes individual to each institution. As there is no maximum of student groups that can be assigned to a student this provides for truly robust reporting functionality. In order to set up a new student group a ticket must be submitted to ERP Support. As with course attributes, the recommended process is to have an institution specific data governance committee approve the request before it is sent to the State Board.

Student Attributes

Student attributes can work alone or in conjunction with student groups and are used to provide additional, more flexible structure to track students based on Program and Career. Students must have a term-activated active plan stack during the quarter being reported as student attributes are associated to plan stacks. The effective date for the program and the student attribute must be the same. For details on adding a student attribute, please review the Adding student attributes Quick Reference Guidelines (QRG) at http://ctclinkreferencecenter.ctclink.us/m/56654/1/496697-adding-and-updating-student-attributes

Student Attributes and State Reporting

Student attributes that begin with the letter "S" are, like course attributes, designated by SBCTC and are used for State reporting purposes and fund allocation.

As with course attributes, student attributes also allow for specifying attribute values. Therefore, a parent student attribute can be defined which has one or more attribute values. An example of a student attribute code is SWRF (Work First) which has the various student attribute values:

Student Attribute:	Description:
SWRF	Work First

Student Attribute Value	Description
SF06	Life Skills
SF01	Customized Job Skills Training
SF02	Vocational Education
SF03	Job Skills Training
SF05	Basic and Developmental Education

Student Attributes and Student Groups

There are three student groups which work in conjunction with student attributes; they are: SWRT (Worker Retraining), SWRF (Work First), and SAPR (Apprenticeship).

The SWRT (Worker Retraining) student group will be automatically applied to any student which has the following student attributes values.

Student Attribute Value	Description
SW01	Dislocated Worker
SW02	Stop Gap Employment
SW03	Receiving UI bens Not Dislocated
SW04	Displaced Homemaker
SW05	Was Self-Employed
SW06	Boeing Dislocated Workers
SW07	Vulnerable Worker
SW08	Disaster Impacted Workers
SW09	Displaced Veterans
SW10	Active Duty Military

The SWRF (Work First) student group will be automatically applied to any student which has the following student attributes values:

Student Attribute Value	Description
SF06	Life Skills
SF01	Customized Job Skills Training
SF02	Vocational Education
SF03	Job Skills Training
SF05	Basic and Developmental Education

The SAPR (Apprenticeship) student group will be automatically applied to any student which has the following student attributes values:

g stadent attributes vardes.	
Student Attribute Value	Description
SAPRFA	Registered Federal Apprenticeship
SAPRSA	Registered WA State Apprenticeship
SAPRAAJAC	Aerospace Joint Apprenticeship (AJAC)
SAPRHS	High School Apprenticeship
SAPRPA	Pre-Apprenticeship (not used for mass
	assign into Student Group)

Global vs. Local Student Attributes

Student attributes are local in nature, so they are also tied to an institution and each institution will only see the student attributes tied to their own college. Yet, student attributes that begin with 'S' created by SBCTC will also be created for each institution so they can be used globally. The local student attributes can be used for purposes individual to each institution while the global 'S' attributes should be used exclusively for the required purpose and definition of the code.

As there is no maximum of student attributes that can be assigned to a student this provides for truly robust reporting functionality. In order to set up a new student attribute a ticket must be submitted to ERP Support. As with other attributes, the recommended process is to have an institution specific data governance committee approve the request before it is sent to the State Board.

FTE TABLES IN CAMPUS SOLUTIONS

One of the most widely anticipated features of PeopleSoft by many Legacy users is the ability to have a more in-depth view of how FTES are calculated. PeopleSoft has a custom built table with FTE information at the student/class enrollment level where users are able to view the criteria used in FTE calculations. In Legacy FTES are set at the class level, not the enrollment level, during the MIS finals processes. The FTE table in Campus Solutions is called CTC_STDNT_FTE.

CHECKING THE DATA FOR ERRORS

Quality Assurance Reporting System (QARS)

QARS (pronounced cars) replaces Legacy MIS Edit Checks. It allows colleges a way of preventing coding mistakes that affect State and Federal level reporting.

QARS and SBCTC Master Data Views

Master data is a single source of common business data used across multiple systems, applications, and/or processes. This is the data that the State Board uses in the compilation of State and Federal

required reports. With the Legacy system there is a series of flat files, or static information that is provided to the State board and then processed to a usable format for reporting. It is a manual and time-consuming process. PeopleSoft has changed that by providing dynamic data delivered by the SBCTC Master Data Views. A view is simply a process of taking information from multiple tables and compiling it all into a single table.

These master data views allow for the QARS or Quality Assurance Reporting System to function. QARS gives colleges the opportunity to find and fix coding errors before it is too late to do so, ensuring each college has the chance to accurately report all the information which is pertinent to their receiving all of the allotted funding due. Once the deadline for data changes has passed the data is extracted from PeopleSoft via the Master Data views and passed to the Data Warehouse. The State and Federal reports are compiled with the data stored in the Data Warehouse.

The source of the data is the SBCTC Master Data Views built using PeopleSoft application data. A View is a way of combining data from many different tables into a single table. So the SBCTC Master Data Views store all of the pertinent information needed for State and Federal reporting

Using the Quality Assurance Reporting System

Current quarter data is live or refreshed nightly in this system (the refresh frequency is listed in QARS for many reports). Some coding changes made today will not be reflected until tomorrow, while others are immediate. The source of any historical data is the SBCTC Data Warehouse.

There are two categories of reports: Summary and Detail. The Summary data contains aggregate headcount or FTE values. The Detail data will contain lists of classes and or students who are found to have coding issues.

There are two levels of quality assurance: Errors and Warnings. It is critical that any errors are resolved as any classes or students shown with errors will **not** be included in the SBCTC Data Warehouse. Any classes or students shown with Warnings will be included in the SBCTC Data Warehouse, though the quality of that record may affect reporting.

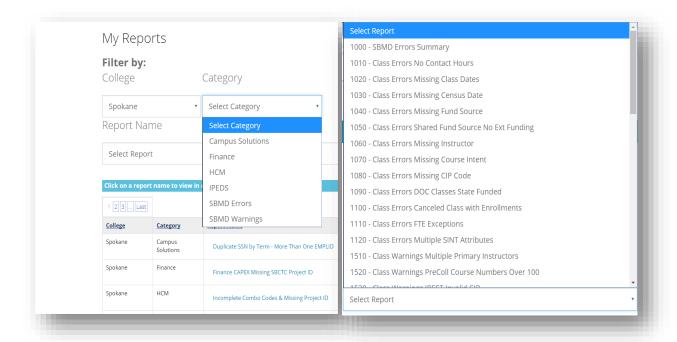
Logging into QARS

The Quality Assurance Reporting System is an online resource found at https://qars.sbctc.edu.

A login is needed to access the reports found in the QARS System. To receive a login, send an email with the request to: Lsager@sbctc.edu or dataservices@sbctc.edu

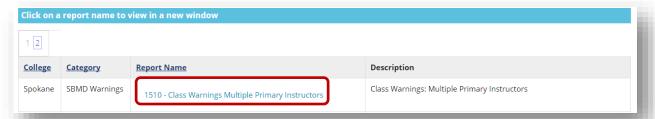
My Reports

Once users have access to QARS they are able to pull reports based on College and Category or they can view all available QARS Reports. The reports are numbered to let users know which report should be looked at first. For example, users need to correct "Missing Census Date" or "Missing Fund Source" before checking the "FTE Exceptions" report. Since both census date and fund source are required to generate FTE, correcting the errors on these reports will correct some of the errors on the "FTE Exceptions" report. There is also a "summary" report that allows users to see the number of records on each report, so that users do not have to open reports that do not have any records listed.

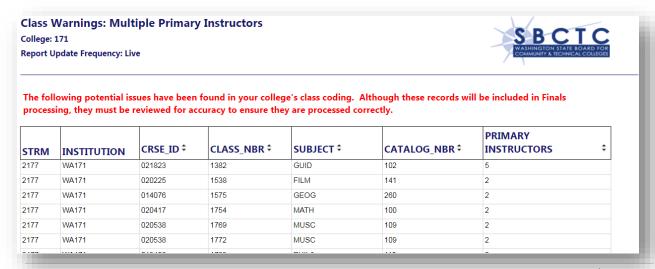


Select the desired report and click on the Report Name hyperlink.

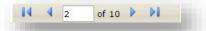
If the report does not provide additional selection criteria then it will pull up in a new window. However, some reports can be run by different criteria such as term. If that is the case then select the term from the drop down window and then select the View Report Icon. Only three terms are available in the system.



The Report Header will display the name of the Report, the college code for which is it being run as well as the Report Update Frequency. Some QARS reports are live data and some are updated nightly.



The data is broken down into pages which are displayed separately. Navigate by clicking the blue arrow icons



Search for text displayed on the page by typing in the search box and clicking enter or Find.



To download the Report click on the download icon and select the desired format.



DATALINK AND GOLDENGATE

Oracle GoldenGate is software that provides real time data integration and replication. This allows ctcLink users access to live data and the ability to make fast, real time decisions based on current information without overutilization of system resources or interrupting users. This process is referred to as datalink. CtcLink has made the decision to disallow reporting in the production environment as this can severely impact system resources and negatively affect users of the system. This presented a challenge for users who need access to live, real-time production data. This challenge has been addressed and solved by dataLink. For more information on dataLink see the Data Services website: https://www.sbctc.edu/colleges-staff/data-services/datalink.aspx

Near real-time production data

dataLink replicates live data with a 2 to 3 second delay so analysts and researchers are able to quickly and easily report on near real-time information.

dataLink works by only looking for information that has been changed within PeopleSoft and then replicating that information to a pre-defined target reporting instance. For ctcLink there will be one reporting database per institution for each institution's data and one for the State Board with all data. Each institution will then be able to access all of their mission critical information in near real-time, experiencing only the 2 to 3 second delay in getting the latest data. Simply stated GoldenGate is the software which makes this real-time data replication possible and dataLink is database where their replicated data is stored.

Adding Tables Which Will Be Replicated by DataLink

While PeopleSoft has thousands of tables only a small percentage of them are replicated by dataLink to increase efficiency. However, new tables are very easy to add to the list of replicated tables so never fear that needed data will not be available in dataLink. Just submit a request to Data Services to replicate the necessary table if it is not already being replicated.

dataLink, GoldenGate and metaLink

GoldenGate is also the software that allows for the ctcLink data dictionary metaLink to be updated on a nightly basis. As new records are replicated or queries/reports/views added or updated to dataLink, GoldenGate ensures that those changes are then also reflected in metaLink. This keeps the data dictionary accurate and relative increasing its value to all users. MetaLink is discussed in detail in the next module.

In fact this dynamic duo is how registrars, researchers, etc access and consume all production data outside of the PeopleSoft application. metaLink, QARS, the Data Warehouse and local college Operational Data Stores are all powered by GoldenGate and dataLink.

Legacy Data Access

Legacy users are currently able to access production data outside of the system by accessing a FTP site where they are able to download the files for consumption. However, there is a plan in place to use GoldenGate to replicate Legacy production data as well so Legacy users will be able to enjoy some of the same benefits that GoldenGate has to offer.

METALINK DATA DICTIONARY

With any PeopleSoft implementation one of the most difficult thing for users responsible for Query and Report development is to determine what data lives where and/or to understand the meaning and usage of each record and field. The sheer number of available tables and fields can be overwhelming for anyone attempting to learn the data structure. In addition end users reading reports may have

questions about the exact nature of the data they are looking at. What each particular date field means, for example.

The obvious answer to this challenge is, of course, a Data Dictionary. A Data Dictionary is a repository of detailed information about the data itself such as relationships to other data, definitions, usage, format, etc. Oracle defines it as a collection of tables with metadata. What is metadata? It is data about data. An example would be the field EMPLID. What type of field is it? Is it a character field, a numeric field? How many characters can it contain? What are valid entries into the field? The Data Dictionary gives users a place to search all of that information and more regarding the data that they use daily. In addition, the Data Dictionary also provides a place for information regarding data relationships.

To meet the need for a Data Dictionary for our Query Developers and end users to have access to the information they need regarding PeopleSoft data structure and metadata, Greg Gamble, of the Application Development team at the State Board developed a web application called metaLink. It is a .Net site on Framework 4.0, using C sharp on code pages. SQL Server is the Database. Having the web app developed in house allows us amazing flexibility.

The MetaLink app will support the college system by providing key information about the data stored within the tables and columns in Golden Gate, which are replicated from the ctcLink production database. The key information is critical to the accurate use of PeopleSoft data for reporting and/or analysis.

The key features of MetaLink allows designated users to:

- Document structures and properties for a given record or field.
- View completed data dictionary reports.
- View the ctcLink reporting catalog.
- Search for available PeopleSoft data queries.

LOGGING INTO THE METALINK DATA DICTIONARY

The Data Dictionary does require a login as it details PeopleSoft data structure which is proprietary. The Data Dictionary provides the ability to search records by pillar or to search our State Board Master Data which is used for our data warehouse.

In order to gain access to MetaLink you must first request a password. Please submit all requests for MetaLink passwords to: Carmen McKenzie or Paula McDaniel. cmckenzie@sbctc.edu or pmcdaniel@sbctc.edu.

Navigate to http://dataservicesmetalink.sbctc.edu

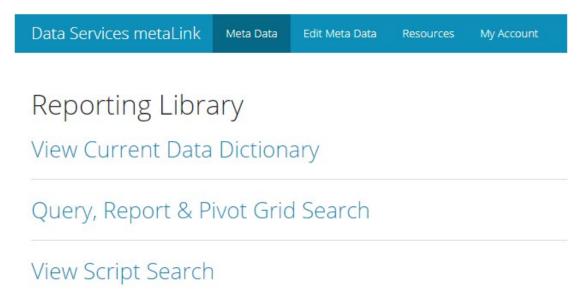
Click on Login in the upper right hand corner of the screen and enter in your login credentials.

You will now be able to view metadata, edit metadata or edit your account information. Click on Meta Data to go to the Reporting Library.

Data Services metaLink Meta Data Edit Meta Data My Account

USING THE METALINK DATA DICTIONARY

To view record information, click on View Current Data Dictionary. Note: at this point in the training we are focusing on finding record and field information. The other areas of metaLink will be explained later in this class.



Select a pillar then choose to either display only common tables or only prompt tables and then select the table and click View Report. Note that record names in MetaLink are slightly different than the record names in PeopleSoft query as all record names in MetaLink begin with PS_.



There are two main sections: Table Information and column Information. Table Information holds details regarding the record while column Information holds details regarding the fields of the table.

Record Information

Record details are found under the header TABLE INFORMATION. The available fields are:

- Table Name Name of the record.
- Alternative Table Name The laymen name for the record. For example, the record STDNT_ENRL is commonly referred to as the Student Enrollment table.
- Table Description Description of the record and what it is used for. Any pertinent information regarding the record.
- Business Use Description Description of the business use of the record.
- Parent Child Business Rules A parent-child hierarchy is a hierarchy in a standard dimension that
 contains a parent attribute. A parent attribute describes a self-referencing relationship, or self-join,
 within a dimension main table. Parent-child hierarchies are constructed from a single parent
 attribute.
- Pillar the Pillar the record belongs to.
- Module the Module the record pertains to.
- Prompt Table indicates if the record is used as a prompt table.
- Effective Dated Table indicates if the record contains the EFFDT field.
- Commonly Used Table indicates if the record has been flagged as commonly used.

TABLE INFORMATION	
Table Name	PS_BEN_HEALTH_SRCH
Alternative Table Name	
Table Description	HEALTH_BENEFIT is a record that identifies all health-related benefit elections for an employee. These benefits are related to the plan types in the '1x' series such as Medical, Dental, and Vision. Warning! For technical reasons, this record's schema shows that it is keyed by EMPL_RCD. This is not logically or functionally accurate. Throughout all of Base Benefits, when in the context of benefit enrollment, EMPL_RCD must be interpreted as an alias for BENEFIT_RCD_NBR. All benefit enrollment-related records are logically keyed by BENEFIT_RCD_NBR, so here EMPL_RCD actually contains the value of the appropriate BENEFIT_RCD_NBR. It is CRITICAL that all joins, comparisons, and updates to this record be against BENEFIT_RCD_NBR. Thus a typical driving join for benefit participation is PS_PER_ORG_ASGN_BENEFIT_RCD_NBR = PS_BEN_PROG_PARTIC.EMPL_RCD, and a typical join for benefit enrollment is PS_PER_ORG_ASGN_BENEFIT_RCD_NBR = PS_HEALTH_BENEFIT.EMPL_RCD. The security and search views used in Base Benefits perform this "mapping" internally.
Business Use Description	
Parent Child Business Rules	
Pillar	SYSADM_HCM
Module	BA - Benefits Administration
Prompt Table	No
Effective Dated Table	No
Commonly Used Table	No

Many of the tables only contain information on the Table Name and Pillar which is why it is so important to add information if you have it. If we look at the table BEN_HEALTH_SRCH in the HCM pillar we can see some of the great information that has been added to metaLink about the record.

The information in the Table Description field of this record provides users with vital knowledge about which fields to use to create a join with another record. Without this information a user could inadvertently join on the wrong field and thus create a query with incorrect or misleading results.

Field Information

Field details are found under the header COLUMN INFORMATION. The available fields are:

- Alternative column Name The laymen name for the column. For example, the column name STRM is commonly referred to as Session/Term.
- Description The description of the field and what it is used for. Any pertinent information regarding the field should be included.
- Primary Key Indicates if this field is part of the primary key for the record (table). The combination of primary key fields represent the uniqueness of the row in the record.
- Data Type A classification identifying one of various types of data, such as integer, text or decimal that determines the possible values for that type.
- Size The length or scale of the field value. Size is directly related to the Data Type.
- Nullable Indicates of the field can contain null or blank values.
- Data Classification The WA State Office of the Chief Information Officer (OCIO) data classification. Category 1 thru 4 are used with Category 1 being public information and category 4 being confidential information requiring special handling.
 - For a detailed description of each category, see the Data Classification Brief
 here: http://www.sbctc.edu/resources/documents/colleges-staff/data-services/data-warehouse/DataClassificationBrief.pd
- History: Includes any historical changes to the field such as inclusions or exclusions of values.
- Accountable Commission: Designated by the Data Governance Committee and indicates which system commission is ultimately accountable for this specific data element.
- Examples of Use: Provides examples of how/when to use the field. Information about auto-joins and how they might affect the query, etc.
- Converted: The table and field name of the legacy source record if converted during implementation.
- Common column: Used only by SBCTC.
- Notes Any other notes, pertinent information regarding the field.
- Valid Values field values and their meanings.



You will see how the information presented could be useful. For example, when studying the tables, we see that the STDNT GRPS table does not have an Effective Status yet the STDNT GRPS HIST does have

Effective Status. This means that whenever you need a query to display students in an inactive status you must use the STDN'T_GRPS_HIST table. Also, note the STDNT_GROUP_TBL defines student group codes making it an excellent choice for a prompt table.

METALINK QUERY, REPORT AND PIVOT GRID SEARCH

One of the great things about MetaLink is that it allows users to search for queries, reports and pivot grids across a variety of fields from a single input source. Users are able to search across all pillars, by a single Folder or by all Folders, across Description, Definition and field name. The plethora of available options and the ease of searching is one of the best features of MetaLink. In PeopleSoft query for example, the only way to search by query definition is by creating a query to do so while MetaLink allows query definition searching with no additional effort required.

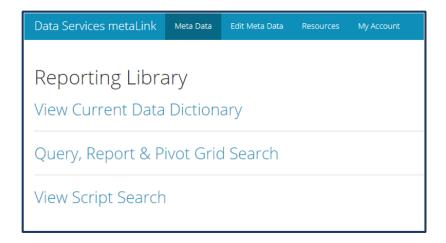
To search MetaLink for existing queries, reports and pivot grids in Production click on the MetaData tab if you are logged into the system.



If you are not logged in as a logon is not required for this section simply click on Search Tools.

Data Services metaLink	Search Tools	Resources	

Next click on View Query, Report and Pivot Grid search hyperlink.



From here you will enter your search criteria.

It is important to note that users enter the search criteria in one page to search three different objects; queries, reports and pivot grids. However, each search criteria may not actually apply to all three objects. For example, Modules will search reports only while Folders will only search queries. However, Pillars will search queries, reports and pivot grids. If the search criteria does not apply to a certain tab

then an exhaustive list of complete inventory will display instead of a filtered list of search results. How will a user know which search criteria applies to which tabs? There is a Definitions box to the right of the search criteria which explains which search criteria affects which tab.

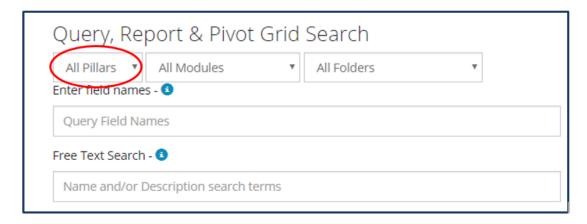
Definitions

- · Pillar Will search Queries, Reports and Pivot Grids.
- · Module Will search Reports.
 - Pivot Grid and Queries will display ALL current inventory.
- · Folders Will search Queries.
 - Pivot Grid and Reports will display ALL current inventory.
- Field Names Will search Queries and Pivot Grids.
 - · Reports will display ALL current inventory.
 - Use a comma (,) as an OR separator to search multiple fields.
 - AND separator searches are not available.
- Free Text Will search Queries, Reports and Pivot Grids.
 - Free Text Search allows using up to four search terms.
 - Use a comma as an OR separator to search using multiple search terms.
 - Use a single space as an AND separator to search using multiple search terms.
- * All search boxes are separated by AND conditional operator logic. *

Search Criteria Boxes

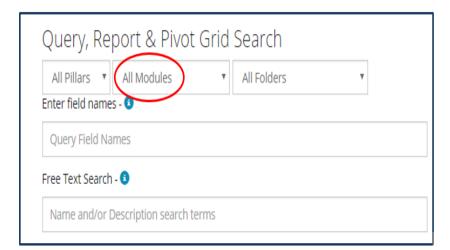
All Pillars

If you want to search a specific pillar you are able to select it in the "All Pillars" box. If you want to search across all three pillars simply leave "All Pillars" selected. Queries, reports and pivot grid results will be returned based on your selection.



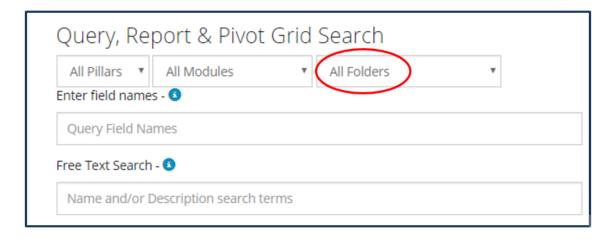
Modules

Selecting a module will narrow the search to only return Reports based on the selected module. The search will produce results in the Queries and Pivot Grids tabs, however the results will simply be a list of all current queries and pivot grids, not a list of results based on Module. Instead, the only tab which will reflect a filtered list of results will be the Reports tab.



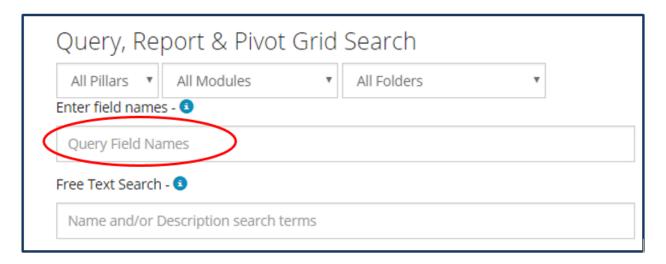
Folders

In the next box you are able to search for queries in either a specific folder or across all folders if 'All Folders" is selected. The search will produce results in the Reports and Pivot Grids tabs, however the results will simply be a list of all current reports and pivot grids, not a list of results based on Folder. Instead, the only tab which will reflect a filtered list of results will be the Queries tab.



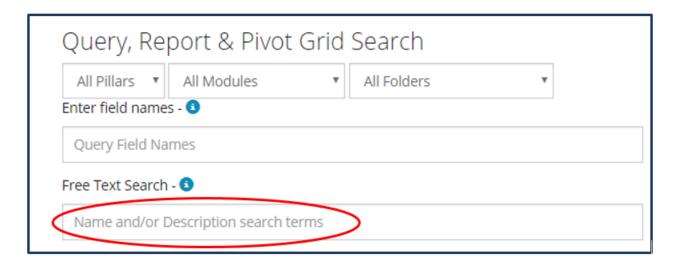
Field Names

In the Fields box users are able to enter field names to find results. Commas act as an "OR" separator allowing for multiple fields to be searched, however it is not possible to search multiple fields based on an "AND" logic at this time. The search will produce results in the Queries and Pivot Grids tabs, results in the Reports tab will simply be a list of all current reports, not a list of results based on the Field/s entered.



Free Text Search

The "Free Text Search" box will search across all columns in all three tabs returning any results which contains the word or words entered. Up to five terms can be separated by a comma for an "OR" logic separation or a space which acts as an "AND" separation logic.

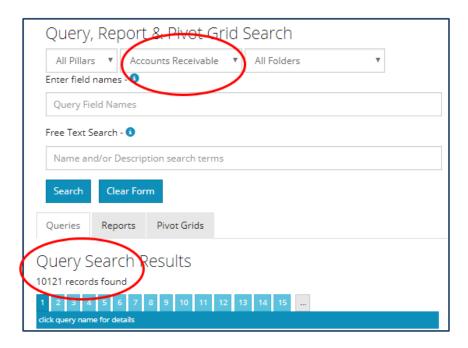


Combine Search criteria

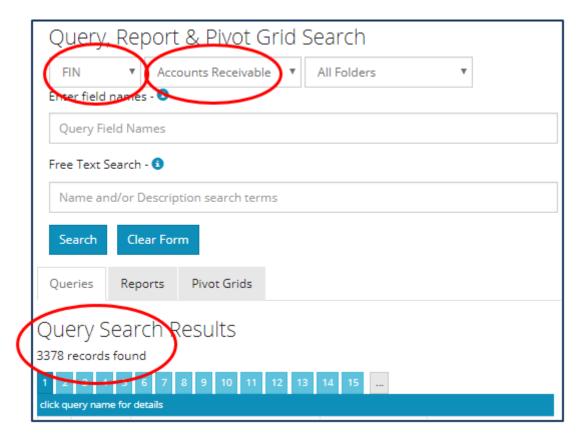
In the same way as the Advanced Search in PeopleSoft, all of the fields used for searching in MetaLink can be combined with other fields to narrow down your search results to find specifically the queries, reports or pivot grids you are most interested in. Just make sure to keep in mind that some of the search fields are designed to only search either queries, reports or pivot grids.

Search Results

Each tab will either display results or a message saying no results found if the search applies to that tab. If the search does not apply to the tab then a complete list of current inventory will display. For example, if a search is performed on the Modules box then a complete, exhaustive list of all queries will display as the Modules search does not apply to queries as shown below.



It gets really interesting however, if another search criteria has been selected. Following the same example, let's say that we selected the Pillar of FIN and then selected a Module. Then, the list of queries (and pivot grids) will be exhaustive and complete WITHIN the Finance module.



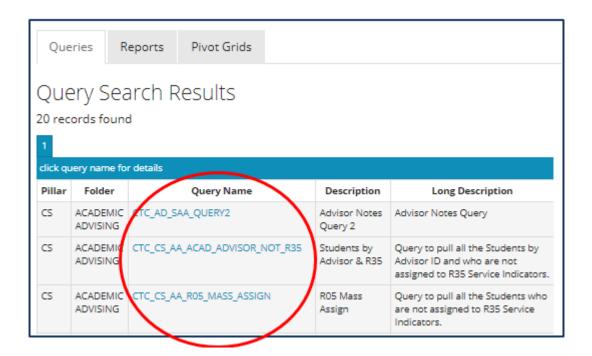
It can be a bit confusing at first but with practice it will soon make sense!

Queries Tab Results

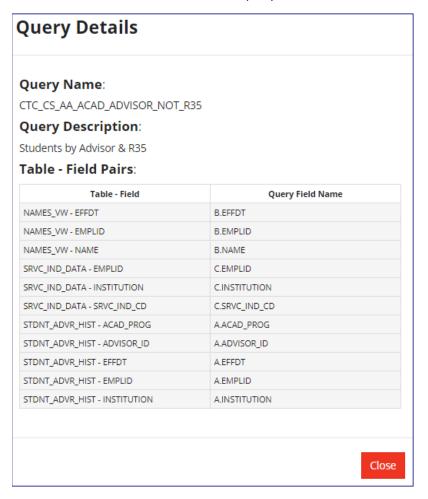
The Queries Tab will show queries that match the search criteria (or all queries as explained above). The list will include the:

- Pillar
- Folder
- Query Name
- Description
- Long Description (Definition)

The query name will be In the form of a hyperlink which can be clicked to bring up additional details about the query.



The additional details available via the hyperlink are the table/field pairs which shows which records and fields were used in the creation of the query.

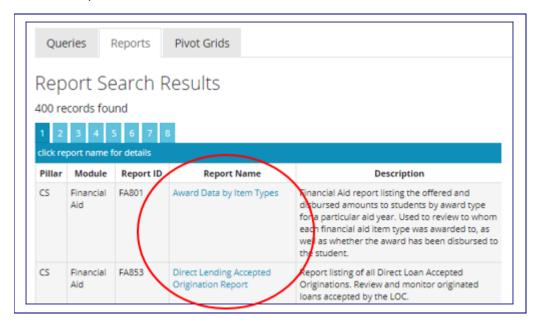


Reports Tab Results

The Reports Tab will show reports that match the search criteria (or all reports as previously explained). The list will include the:

- Pillar
- Module
- Report ID
- Report Name
- Description

The report name will be In the form of a hyperlink which can be clicked to bring up additional details about the report.



The additional details available via the hyperlink are:

- Report ID
- Business Process Number
- CEMLI Number
- Report Name
- Report Description
- Report Type
- PeopleSoft Navigation



Pivot Grids Tab Results

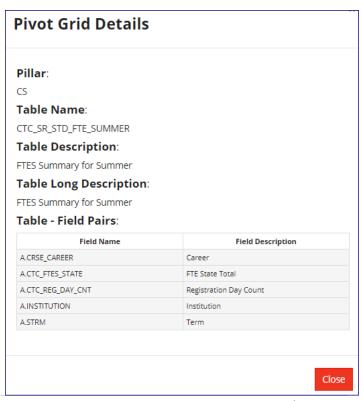
The Pivot Grids Tab will show pivot grids that match the search criteria (or all pivot grids as previously explained). The list will include the:

- Pillar
- Object Name (Pivot Grid Name)
- Description
- Long Description

The pivot grid name will be In the form of a hyperlink which can be clicked to bring up additional details about the pivot grid.



The additional details available via the hyperlink are table/field pairs which show which records and fields were used in the creation of the pivot grid.

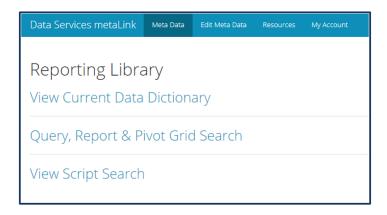


VIEW SCRIPT SEARCH

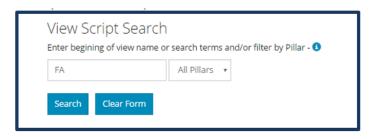
The MetaLink Script Search allows for searching for the underlying SQL script used in the creation of Views. Users can search by View Name and/or by Pillar. Click on Meta Data to go to the Reporting Library.



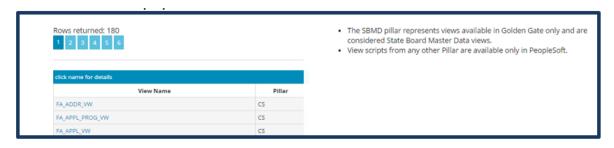
Click on View Script Search.



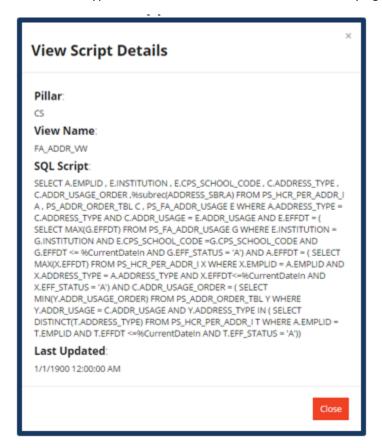
Here you will enter your search terms.



The results will display below.



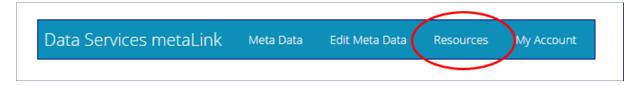
Click on the hyperlink of the View Name to see the underlying SQL code.



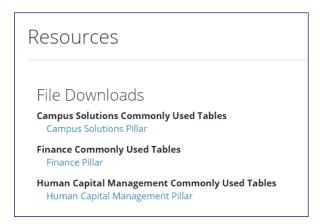
RESOURCES

While it is now possible to search commonly used tables in metaLink there is also a Resources page where full documents of commonly used tables are available by pillar.

To access these documents simply click on Resources when accessing the home page.



This brings the user to the Resources page where documents are available for download.



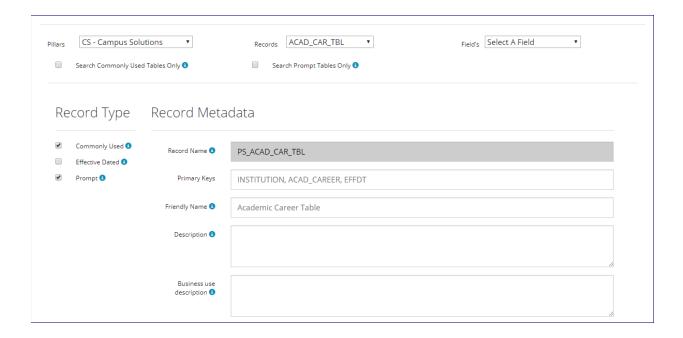
ADDING INFORMATION TO METALINK

How is it populated? Well, while we in Data Services are definitely the responsible party for making sure it is populated, we are also harnessing the knowledge and expertise of our users... all of our users, not just Query Developers. The app is basically a wiki utilizing crowdsourcing as its source of information. We leverage our communities of practice, functional analysts, project staff... anyone who knows (or finds out something interesting) about the data is able to share their knowledge. As we gain information about data structure and business uses of different fields and records we have an obligation to all our fellow users to share it. It is quick and easy to add in information.

Click on Edit Meta Data



To update the record – select only the Pillar and record. To update a field within a record – select the Pillar, record and field you would like to modify. Type in the new information.



Scroll to the bottom of the screen when finished and click Save Data.