ASSIST Articulation Data Extract Specifications

August 25, 2004 – Revision

PLEASE NOTE: This document is an update of the earlier ASSIST Articulation Data Extract Specifications (dated 03/11/2002) and provides information on an additional data extract, "CSU and UC courses referenced in articulation agreements". Also, there is a related document available on the ASSIST Information Center titled "Policies for ASSIST Data Extracts" that provides further information on use restrictions for ASSIST data extracts.

Background

All California Community College, California State University and University of California campuses maintain a variety of course transfer and articulation information in ASSIST. The ASSIST database is comprehensive, historical and fully relational – that is to say that the information stored in ASSIST is not simply textual documents that are displayed on a web site, but dynamic data that is assembled for display each time a report is requested. The ASSIST database was designed to facilitate data entry and updating course articulation as well as to display various types of course articulation information. The ASSIST database was also designed to support data extracts so that campuses could use raw data from ASSIST in other campus-based information systems. This document provides the specifications for these ASSIST Articulation Data Extracts.

These specifications are divided into three sections. Section 1 describes the general format of the extract files. Section 2 describes the contents of each extract file along with any other important information. Section 3 describes each of the data elements in detail including valid values and other important characteristics.

It is important to note that while ASSIST does include a broad range of detailed information, some information systems that could benefit from integrating with ASSIST data require a different level of detail, or different information all together, than can be provided by ASSIST. It is critical that campuses who intend to use ASSIST data in their local systems perform detailed analysis to fully understand what data is needed, whether ASSIST data is adequate to meet local needs, how ASSIST and ASSIST-related data will be updated in the local system, and any other key policies and practices related to using data from ASSIST.

Section 1 - General Format of ASSIST Articulation Extract Files

Any college or university that has developed or is developing a system that could use data from ASSIST generally will be customized to meet local needs. This means that each campus would potentially use different structures to store course and articulation related information. Since there is no universally agreed upon structure for storing course and articulation information, an ASSIST extract format that is tailored to the needs of each campus is not feasible. Instead, a

generic format is utilized for the extract data. This requires that individual campuses convert the extract files into a format suitable for their specific systems.

The ASSIST database includes a variety of transfer/articulation related information that have potential use in degree audit systems. There are six types of data that can be extracted, each into a separate file, including:

- UC Transferable Community College Courses
- CSU Transferable Community College Courses (CSU Baccalaureate Level)
- IGETC Applicable Community College Courses
- CSU-GE/Breadth Certification Applicable CSU and Community College Courses
- CSU and UC Courses Referenced in Articulation Agreements
- Instances of Course-to-Course Articulation (articulation between pairs of institutions)

While not all institutions require all of these types of data, they are discussed completely in this document. The general format of each of these files is similar. The specific data elements to be included are different from file to file. All data elements are represented as variable length data, although some data element descriptions identify a fixed length. Each file is provided as an ASCII text file with data elements delimited by tildes (~) and records delimited with a carriage return and line feed. Textual data elements are not surrounded by quotation marks. Generally, records within each file are sorted by:

1) Institution-Code
2) Course-Prefix
3) Course-Number
4) Beginning-Term

Some of the source data in ASSIST also includes textual information such as footnotes, addenda, comments, and articulation entered in textual form rather than fielded form. Since the various local campus information systems that will use these files generally cannot make use of textual information, this information is not included in the extract files. Specific implications of this limitation are discussed separately for each of the files.

The Articulation Agreement data in the ASSIST database is organized by academic year. In the extract data, Beginning-Term and Ending-Term values are deduced from the academic year of the source data. Academic years are assumed to begin with the Fall term and end with the Summer term. Term names are represented by a code, as follows: Fall = 10, Winter = 01, Spring = 04, Summer = 07. If the academic year 97-98 appeared in the source data, for example, the extract process would generate the Beginning-Term value 199710 and the Ending-Term value 199807. It is important to note that the Beginning-Term and Ending-Term values do not imply courses being created or removed from an institution's curriculum. Rather, the Beginning-Term and Ending-Term information refers to the effectiveness of the agreement from which the data was extracted.

In order to request ASSIST articulation data extracts, campuses will submit an email request to the Coordination Site. Coordination Site staff will then process the request and provide the requested data either via email attachments or FTP.

Section 2 - Extract File Descriptions

Following is a more detailed description of each of the extract files including a general description of the extracted data, a brief description of the data source, any special comments about the data, and an ordered list of the data elements included in each record of the extract. The specific descriptions of each data element are available in Section 3 of this document.

Each of the first five data extract file types represent information about individual courses (as opposed to the sixth file type that represents information on the relationship between courses from two institutions). Over the life of a given course the prefix, number, title and units may change. Each change to any of these data elements results in a different version of the complete course information for a specific time period. In order to most accurately represent these changes over time, the first five data extract files include additional rows of output for each version of the complete course information with corresponding qualifications of begin and end terms. Additional versions of courses are only included in an extract if they meet the criteria for the data extract type. For example, if a course that is currently UC transferable had a different title and unit value during a time when the course was not UC transferable, the earlier version of the course would not be included in the first data extract file (UC Transferable Community College Courses).

File # 1: UC Transferable Community College Courses

Description: This extract file includes data for all community college courses that are UC transferable.

Data Source: These are the official data as approved and maintained by the University of California Office of the President.

Comments: This file includes a separate row of output for each version of a course name, title and/or units. This file includes courses which are also in other files. If the UC-Credit-Limit-Flag data element is a 'Y', this course has some kind of a textual comment or footnote which describes a limitation that cannot be codified.

Data Elements:

1. Institution-Code *
2. Course-Prefix *
3. Course-Number *
4. Beginning-Term *
5. Ending-Term
6. Course-Title
7. Course-Units (Minimum)

8. Course-Units (Maximum)
9. Unit-Type
10. UC-Credit-Limit-Flag
11. UC-Eligibility-Code

Note: Data Elements marked with an asterisk (*), taken together, can be used to uniquely identify each record.

File # 2: CSU Transferable Community College Courses (Baccalaureate Level)

Description: This extract file includes data for all community college courses that are CSU Baccalaureate Level (CSU Transferable).

Data Source: Community college ASSIST curriculum data maintained by the individual community colleges via the ASSIST Curriculum Update System.

Comments: This file includes a separate row of output for each version of a course name, title and/or units. This file includes courses which are also in other files.

Data Elements:
1. Institution-Code *
2. Course-Prefix *
3. Course-Number *
4. Beginning-Term *
5. Ending-Term
6. Course-Title
7. Course-Units (Minimum)
8. Course-Units (Maximum)
9. Unit-Type

Note: Data Elements marked with an asterisk (*), taken together, can be used to uniquely identify each record.

File # 3: IGETC Applicable Community College Courses

Description: This extract file includes data on all community college courses that have been approved for use in the Intersegmental General Education Transfer Curriculum (IGETC).

Data Source: This is the official data as approved by the California State University Chancellor's Office and the University of California Office of the President and maintained by the ASSIST Coordination Site.

Comments: This file includes a separate row of output for each version of a course name, title and/or units. The beginning and ending terms of each record indicate the period in which the respective course version can be used for the given IGETC code. Courses with multiple IGETC codes will have a separate record for each code in which the course is applicable as the beginning and ending terms may be different for each IGETC code. This file includes courses which are also in other files.

An additional sort value for this data is the IGETC code.

Data Elements:

1. Institution-Code *
2. Course-Prefix *
3. Course-Number *
4. IGETC-Code *
5. Beginning-Term *
6. Ending-Term
7. Course-Title
8. Course-Units (Minimum)
9. Course-Units (Maximum)
10. Unit-Type

Note: Data Elements marked with an asterisk (*), taken together, can be used to uniquely identify each record.

File # 4: CSU-GE/Breadth Certification Applicable Courses

Description: This extract file includes data on all community college courses that have been approved for use in CSU General Education/Breadth Certification (per CSU Executive Order #595).

Data Source: These are the official data as approved by the California State University Chancellor's Office and maintained by the ASSIST Coordination Site.

Comments: This file includes a separate row of output for each version of a course name, title and/or units. The beginning and ending terms of each record indicate the period in which the respective course can be used for the given CSU GE/Breadth code. Courses with multiple CSU GE/Breadth codes will have a separate record for each code in which the course is applicable as the beginning and ending terms maybe different for each CSU GE/Breadth code. This file includes courses which are also in other files.

An additional sort value for this data is the CSU-GE/ Breadth code.

Data Elements:

1. Institution-Code *
2. Course-Prefix *
3. Course-Number *
4. CSU-GE-Certification-Code *
5. Beginning-Term *
6. Ending-Term
7. Course-Title
8. Course-Units (Minimum)
9. Course-Units (Maximum)
10. Unit-Type

Note: Data Elements marked with an asterisk (*), taken together, can be used to uniquely identify each record.

File # 5: CSU and UC Courses Referenced in Articulation Agreements

Description: This extract file includes data for all CSU and UC campus courses that are referenced in the ASSIST database.

Data Source: CSU and UC campus ASSIST curriculum data is maintained by the individual universities via the ASSIST Curriculum Update System.

Comments: This file includes a separate row of output for each version of a course name, title and/or units. This file includes courses which are also in other files.

Data Elements:

1. Institution-Code *
2. Course-Prefix *
3. Course-Number *
4. Beginning-Term *
5. Ending-Term
6. Course-Title
7. Course-Units (Minimum)
8. Course-Units (Maximum)
9. Unit-Type

Note: Data Elements marked with an asterisk (*), taken together, can be used to uniquely identify each record.

File # 6: Course-to-Course-Articulation

Description: This extract file includes data on instances of course-to-course articulation (articulation between pairs of institutions). This extract file includes course-to-course articulation of single courses and course sequences. It does not include articulation records that are

exclusively textual in nature. When an instance includes both course(s) and textual information the courses are included in the extract and the corresponding Text-Flag data element is set to "Y". The actual text is not included in the data extract.

Data Source: ASSIST Articulation Agreement database. This data is maintained in ASSIST by the receiving institution. Only finalized articulation is included in this file.

Comments: Any instance of articulation can have one or more courses on the sending or receiving side of the relationship. When this occurs there will be multiple rows of output for the instance where each row includes the Nth sending and receiving course in the instance. For example if an instance has 2 sending courses and 3 receiving courses, there will be 3 rows of output for the instance and each row is further uniquely identified by the Item-Position-Number data element. All Item-Position-Numbers for a given ID-Number with a specific pair of institutions in a specific year can be grouped together to represent the complete instance of articulation. Since the order of multiple course items within an instance is important (in order to correctly correlate conjunctions) this data element is computer generated and always begins with 1 for a given instance.

When there are multiple sending or receiving courses, it must be noted that it CANNOT be assumed that the Nth sending course relates directly to the Nth receiving course. In order to correctly interpret an instance, the complete set of sending courses must be considered to be related to the complete set of receiving courses. This is a very important concept to understand for local information systems that will be doing automated processing of data using the Course-to-Course Articulation data extract.

Instances of articulation that apply to multiple majors will have separate rows of output for each major to which the instance of articulation is applicable.

Depending on how the receiving institution maintains articulation agreements in ASSIST, the degree to which this extract file matches the actual articulation may vary greatly. For example, if a receiving institution uses only textual information in agreements, no course-based data can be extracted. If a receiving institution uses a combination of course-based and textual information, the extracted data may only represent part of the information in the full agreement. If a receiving institution uses course-based articulation exclusively, the extracted data should match the full agreement

closely. This will need to be assessed on a campus-by-campus basis by reviewing how data is entered into ASSIST.

Data Elements:

1. Institution-Code * (Receiving Institution)
2. Institution-Code * (Sending Institution)
3. Beginning-Term *
4. Receiving-Institution-Major *
5. Receiving-Institution-Major-Position *
6. Receiving-Institution-Major-Conjunction
7. Include-By-Department-Flag
8. ID-Number *
9. Item-Position-Number *
10. Ending-Term
11. Unit-Type (Sending Institution)
12. Text-Flag (Sending Institution)
13. Course-Prefix (Sending Course Prefix)
14. Course-Number (Sending Course Number)
15. Course-Title (Sending Course Title)
16. Course-Units (Minimum Sending Course Units)
17. Course-Units (Maximum Sending Course Units)
18. Conjunction (Sending Course Conjunction)
19. Unit-Type (Receiving Institution)
20. Text-Flag (Receiving Institution)
21. Course-Prefix (Receiving Course Prefix)
22. Course-Number (Receiving Course Number)
23. Course-Title (Receiving Course Title)
24. Course-Units (Minimum Receiving Course Units)
25. Course-Units (Maximum Receiving Course Units)
26. Conjunction (Receiving Course Conjunction)
27. Comment-Footnote-Flag

Notes: Data Elements marked with an asterisk (*), taken together, can be used to uniquely identify each record.

Records within this file are sorted by:

Institution-Code (Receiving Institution)
Institution-Code (Sending Institution)
Beginning-Term
Receiving-Institution-Major
Receiving-Institution-Major-Position
ID-Number
Item-Position-Number

Section 3 - Definition of Data Elements

The following are definitions of each data element. Many of these data elements appear in multiple files.

Name:	Institution-Code
Description:	Educational Testing Service institution code
Comments:	This is a six-digit code
Valid Data Values:	Any valid ETS code that has been registered with the ASSIST Coordination Site

Name:	Beginning-Term
Description:	The beginning term of the information expressed as a 4-digit year followed by a 2-digit term (10 = Fall, 01 = Winter, 04 = Spring, 07 = Summer). For example: 199704 represents Spring '97. Maximum length: 6 characters.
Comments:	Values for this element in some extract files may not be 100% accurate as it may be deduced (by the extract programs) from the academic year of a file in which the course is stored. Also note that the beginning term for some non transferable, CSU-transferable and CSU-GE courses indicates the first term the course was represented in ASSIST. The institution may have actually begun teaching this course prior to the indicated Beginning-Term.
Valid Data Values:	yyyy01 yyyy04 yyyy07 yyyy10 000000 (No beginning term is defined – Fall 1950 is implied in the data)

Name:	Ending-Term
Description:	The ending term of the information expressed as a 4-digit year followed by a 2-digit term (10 = Fall, 01 = Winter, 04 = Spring, 07 = Summer). For example: 199704 represents Spring '97. Maximum length: 6 characters.
Comments:	Values for this element in some extract files may not be 100% accurate as it may be deduced (by the extract program) from the academic year of a file in which the course is stored.
Valid Data Values	yyyy01 yyyy04 yyyy07 yyyy10 999999 (No ending term is defined – Fall 2070 is implied in the data)

Name:	Course-Prefix
Description:	Discipline abbreviation of the course. For example: MATH. Together, the Course-Prefix and Course-Number comprise the course name.
Comments:	Course prefixes match the format of course names as they appear on computer-produced transcripts from the institution (if the campus actually entered the information into ASSIST correctly). Each institution is responsible for defining and updating the course prefixes they use on transcripts.
Valid Data Values:	The length of any course prefix can vary depending on what each institution uses, however, there are currently no local course prefixes in ASSIST that are over 13 characters in length.

Name:	Course-Number
Description:	Number information for the course, including any number prefix and number suffix. For example: 101, 10A, or V12H. Together, the Course-Prefix and Course-Number, concatenated with a space between them, comprise the course name.
Comments:	Course number information in the ASSIST database may not exactly match the format of course numbers as they appear on computer-produced transcripts from the institution. However, course numbers from ASSIST can be matched with course numbers on transcripts using a simple algorithm that the campus using the extracts must write based on the following information. Internally, ASSIST stores course number information in three parts: number prefix, number and number suffix. Number prefixes cannot contain digits. Numbers are stored as positive integer numbers (with no leading zeros). Number suffixes may contain any characters including spaces and digits (as long as the first character of the suffix is not a space or digit). In ASSIST articulation extracts, these three parts are concatenated together with no spaces. To match course numbers between ASSIST articulation extracts and transcripts, either the ASSIST course number information can be expanded (with extra spaces and leading zeros) or the transcript course number information can be reduced (by eliminating redundant spaces and leading zeros). This work must be done by the campus using the extract data.
Valid Data Values:	See above paragraph.

Name:	Course-Title
Description:	Descriptive, textual title of a course. For example: Physical Anthropology
Comments:	Course titles included in these extract specifications are not expected to be used during any automated evaluations. They are

	included to aid in problem solving and human interpretation of data. Course titles in ASSIST are not expected or required to match with computer-produced transcripts.
Valid Data Values:	There is no standard format or length for valid course titles, but the source database has a maximum length of 255 characters.
Name:	Course-Units
Description:	The units of credit the student receives upon completing the course.
Comments:	Refer to the Unit-Type data element to determine whether the Course-Units refers to quarter or semester units. Since a variable number of units can often be received for completing a course, this data element always occurs twice — once to express the minimum number of units that can be earned, and again to express the maximum units. If the minimum and maximum are the same, the course can only be taken for a fixed number of units.
Valid Data Values:	Any non-negative integer Any non-negative real number. Decimal points are supplied explicitly
Name:	Unit-Type
Description:	Specifies whether the institution uses the quarter or semester system, and therefore provides information for interpreting unit values.
Comments:	
Valid Data Values:	Q (Quarter) S (Semester)
Name:	UC-Credit-Limit-Flag
Description:	Indicates if a credit limit applies to a given course.
Comments:	If a credit limit does apply, it is expressed as textual information in the source data. This textual information is not included in the extract since it is not represented in a consistent, codified format that could be used accurately by other evaluation systems.
Valid Data Values:	Y N
Name:	UC-Eligibility-Code
Description:	Courses marked with "B", “E”, “H”, “M”, or “S” will satisfy the five different areas of the 1998 transfer course requirements.
Comments:	The following are the valid codes and their respective area. Information within parentheses is not part of the data, but rather a description of that data value. Multiple codes can apply to the same

course. When this occurs each code is included with all codes concatenated together with no separators.

Valid Data Values:

B	(Behavioral and Social Sciences)
E	(English)
H	(Humanities)
M	(Math)
S	(Biological and Physical Sciences)

Name: IGETC-Code

Description: The area of IGETC applicability for a specified course.

Comments: The following are the valid area codes and their respective descriptions. Information within parentheses is not part of the data, but rather a description of that data value.

Valid Data Values:

1A	(English Composition)
1B	(Critical Thinking - English Composition)
1C	(Oral Communication {CSU requirement only})
2A	(Math)
3A	(Arts)
3B	(Humanities)
4A	(Anthropology and Archaeology)
4B	(Economics)
4C	(Ethnic Studies)
4D	(Gender Studies)
4E	(Geography)
4F	(History)
4G	(Interdisciplinary, Social & Behavioral Sciences)
4H	(Political Science, Government & Legal Institutions)
4I	(Psychology)
4J	(Sociology & Criminology)
5A	(Physical Science)
5B	(Biological Science)
6A	(Languages Other Than English)
8A	(Interim sequence for 1B - Critical Thinking)
8B	(Interim sequence for 1B - English Composition)
8C	(Interim sequence for 1B - English Composition - 2nd Quarter)

Name: CSU-GE-Certification-Code

Description: The area of CSU-GE/Breadth Certification applicability for a specified course.

Comments: The following are the valid area codes and their respective descriptions. Information within parentheses is not part of the data, but rather a description of that data value.

Valid Data Values:

A1	(Oral Communication)
A2	(Written Communication)
A3	(Critical Thinking)

B1	(Physical Science)
B2	(Life Science)
B3	(Laboratory Activity)
B4	(Mathematics/Quantitative Reasoning)
C1	(Arts {Art, Dance, Music, Theater})
C2	(Humanities {Literature, Philosophy, Foreign Language})
D1	(Anthropology and Archeology)
D2	(Economics)
D3	(Ethnic Studies)
D4	(Gender Studies)
D5	(Geography)
D6	(History)
D7	(Interdisciplinary Social or Behavioral Science)
D8	(Political Science, Government and Legal Institutions)
D9	(Psychology)
D0	(Sociology and Criminology)
E	(Lifelong Understanding and Self-Development)

Name: Receiving-Institution-Major
Description: An abbreviated textual name for the major to which the given instance of articulation applies.
Comments: A single instance of articulation may apply to multiple majors.
Valid Data Values: Major Abbreviations are defined in the ASSIST database for each university and no standard abbreviations are used. Universities are welcome to work with the Coordination Site to define major abbreviations in such a way as they will map directly to major abbreviations or codes used in local campus information systems. A blank value for this data element implies that the instance of articulation is not associated with any specific major.

Name: Receiving-Institution-Major-Position
Description: Indicates an ordinal position for the instance within the specified major.
Comments: When an instance of articulation applies to a major, the order of instances can be important in order to evaluate the Receiving-Institution-Major-Conjunction data element. A blank for this data element implies that the instance of articulation is not associated with any specific major.
Valid Data Values: Any non-negative integer.

Name: Receiving-Institution-Major-Conjunction
Description: The conjunction between multiple instances of articulation that apply to a major
Comments: When an instance of articulation applies to a major, the receiving institution can provide further information that describes how the current instance relates to the next instance (when ordered by the

Receiving-Instance-Major-Position data element). The AND and OR conjunctions imply an AND or OR relationship between the current and next instances. The NONE conjunction implies that the current and next instances are to be grouped together, visually, to imply that they are the same instance. The LINE and PAGE conjunctions are for display purposes and do not imply that there is any relationship between the current and next instances. A blank for this data element implies that the instance of articulation is not associated with any specific major

Valid Data Values: AND
OR
NONE
LINE
PAGE

Name: Include-By-Department-Flag
Description: Indicates whether or not the current instance of articulation is to be presented as departmental articulation in ASSIST.
Comments: This data element is independent of the Receiving-Institution-Major data element. When this data element is Y, it can generally be assumed that the instance of articulation applies throughout the Receiving Institution (unless there are also instances for the course(s) that apply to one or more majors).
Valid Data Values: Y
N

Name: ID-Number
Description: Primary key used to uniquely identify the instance of articulation.
Comments: This data element is used in the source database to uniquely identify an instance of articulation. The ID-Number is only unique in the context of a pair of institutions and an academic year. The same ID-Number may be used for a different pair of institutions or for a different academic year.

ID Numbers included in these extract specifications are not expected to be used during any automated evaluations. They are included to aid in problem solving and human interpretation of data and generate a business key for the # 5 data extract.
Valid Data Values: Any non-negative integer.

Name: Item-Position-Number
Description: Indicates an ordinal position for courses within a single instance when there are multiple sending and or receiving courses.
Comments: Any instance of articulation can have one or more courses on the sending or receiving side of the relationship. This data element is used to order each course when there are multiple courses. All

Item-Position-Numbers for a given ID-Number with a specific pair of institutions in a specific year can be grouped together to represent the complete instance of articulation. Since the order of multiple course items within an instance is important (in order to correctly correlate conjunctions) this data element is computer generated and always begins with 1 for a given instance.

When there are multiple sending or receiving courses, it must be noted that it CANNOT be assumed that the Nth sending course relates directly to the Nth receiving course. In order to correctly interpret an instance, the complete set of sending courses must be considered to be related to the complete set of receiving courses. This is a very important concept to understand for local information systems that will be doing automated processing of data using the Course-to-Course Articulation data extract.

Valid Data Values: Any non-negative integer.

Name: Text-Flag
Description: Indicates if textual information appears within the instance of articulation in place of course information.
Comments: When articulation agreements are formatted, textual information can be used in place of either sending or receiving course names. This textual information is not included in the extract since it is not represented in a consistent, codified format that could be used accurately by other evaluation systems. The presence of a Y in this data element indicates that at least one course was replaced by textual information.
Valid Data Values: Y
N

Name: Conjunction
Description: The conjunction between multiple courses from the same institution within a single instance of articulation in an articulation agreement.
Comments: When an instance of articulation included multiple sending or receiving courses, the conjunction provides additional information on the relationship between courses from the same institution. Due to the nature of how courses and conjunctions are entered and stored in ASSIST, the conjunctions can only be interpreted correctly when multiple courses are ordered using the Item-Position-Number data element.

Depending on how data has been entered into ASSIST, it is possible that there could be some confusion in correctly interpreting this data element since ASSIST does not support parenthetical logic.

Valid Data Values:	AND OR NONE
Name:	Comment-Footnote-Flag
Description:	Indicates if a comment or a footnote applies to a given instance of articulation.
Comments:	When articulation agreements are formatted, textual comments and/or footnotes can be used to provide additional information. This textual information is not included in the extract since it is not represented in a consistent, codified format that could be used accurately by other evaluation systems. The presence of a Y in this data element indicates that either a comment or a footnote was attached to this instance of articulation in the source data.
Valid Data Values:	Y N