

# caArray Roadmap

The list below outlines the priority features ordered and grouped by priorities.

## Development Focus - Features planned for 2.5.0

The following features are planned for caArray 2.5.0:

- Large data support
  - Upon upload, store files (as-is) on the file system instead of in the database.
  - Provide an Uploader client that allows the User to select a large number of files from a local directory and upload to caArray, with support for “pause” and “resume” and failure recovery. (E.g., Flickr Uploadr)
  - Break import into smaller transactions to avoid long-running transactions and to avoid MySQL transaction import limit on parsed data values.
- Other
  - Plugin architecture that supports easy addition of new parsers as well as different file storage options
  - Update the technology stack (Jboss 5.1, Java 6)

## Development Focus - Features planned beyond 2.5.0

The following features are next in line after caArray 2.5.0:

- Grid security and Single sign-on to support TRANSCEND
- Audit trail to support TRANSCEND
- File-level security (if necessary for TRANSCEND)
- Maintenance and Bug fixes including the Collaborator group defect which prevents collaborators from seeing files that have been uploaded but not yet imported.

## Major Features in Scheduled and Past Releases

The table below outlines the major features in **scheduled and past release(s)**.

Version	Development Focus	Release Date
2.4.1	<p>This release represents a feature release of the caArray 2 software. Highlights of caArray 2.4.1 are:</p> <ul style="list-style-type: none"> <li>• Visibility into the queue of import and validation jobs, and where your job is in relation to other jobs in the system. Ability to cancel a queued job. For details about this feature, refer to the <a href="#">Submitting Data to an Experiment</a> chapter in the <a href="#">caArray User's Guide</a>.</li> <li>• Search for experiments by publication author or PubMed ID.</li> <li>• Organisms are now restricted to terms from the NCBI Taxonomy.</li> <li>• Bug fix to allow proper import of Agilent array design files where the chromosome name is not in a strictly numeric or X/Y format.</li> <li>• The installer no longer depends on an internet connection.</li> <li>• Fix to bug where experiments with multiple array designs may cause copy number data to get imported as "Imported (Not Parsed)".</li> </ul>	September 2011
2.4.0	<p>This release represents a feature release of the caArray 2 software: Focus is around data parser. The following data parsers will be included in caArray 2.4.0:</p> <ul style="list-style-type: none"> <li>• Agilent:           <ul style="list-style-type: none"> <li>◦ GEML/xml array designs (aCGH, gene expression and miRNA)</li> <li>◦ Raw TXT data files (aCGH, gene expression and miRNA)</li> </ul> </li> <li>• Nimblegen - Community Code Contribution:           <ul style="list-style-type: none"> <li>◦ NDF array designs</li> <li>◦ Pair Report (raw and normalized) data files</li> </ul> </li> <li>• Illumina:           <ul style="list-style-type: none"> <li>◦ BGX/TXT array designs</li> <li>◦ Gene expression: Sample Probe Profile TXT files with unique Probe_Id</li> <li>◦ Genotyping: Processed matrix TXT files with unique IlmnID values</li> </ul> </li> <li>• Affymetrix:           <ul style="list-style-type: none"> <li>◦ AGCC/Command Console formats for CDF, CEL and CHP files.</li> <li>◦ CNCHP files with copy number and LOH data.</li> </ul> </li> <li>• Copy number data in MAGE-TAB Data Matrix format</li> </ul>	August 2010
2.3.1	<p>This release represents a feature release of the caArray 2 software: Focus is around data export to GEO. Highlights of caArray 2.3.1 include:</p> <ul style="list-style-type: none"> <li>• caArray will allow the user to export an experiment into GEO SOFT format. This SOFT package can then be submitted to GEO.</li> <li>• The initial release will support single-channel experiments that have Affymetrix CHP data files, and that are based on array designs (platforms) already in GEO.</li> <li>• Focussing on CHP files allows us to package the native data files in the SOFT package as opposed to having to derive data tables to include in the SOFT submission.</li> </ul>	February 2010
2.3.0	<p>This release represents a feature release of the caArray 2 software: Focus is around new programmatic interface to caArray application. Highlights of caArray 2.3.0 include:</p> <ul style="list-style-type: none"> <li>• 15578: Redesign Grid API to support backwards compatibility via external data model</li> <li>• 18908: Keep an audit trail of permissions changes made to an experiment.</li> </ul>	October 2009

2.2.1	<p>This release represents a minor release of the caArray 2.x software: Focus is around access control and ownership of experiments. Highlights of caArray 2.2.1 include:</p> <ul style="list-style-type: none"> <li>• 19324: Genepix data files must require an accompanying SDRF and sample names must not be mandated.</li> <li>• 15443: Ability to change experiment ownership</li> <li>• 19532: System administrator should be able to change collaborator group ownership.</li> </ul>	April/May 2009
2.2.0	<p><b>80 features/issues were addressed</b></p> <ul style="list-style-type: none"> <li>• Highlights: <ul style="list-style-type: none"> <li>◦ Sample search: <ul style="list-style-type: none"> <li>▪ The user can now search for biomaterials (Samples and Sources) in the system.</li> <li>▪ The keyword search can be applied to one of a list of standard or predefined categories, or to arbitrary categories previously imported into the system via Characteristic[] columns in a MAGE-TAB SDRF document.</li> </ul> </li> <li>◦ Export experiment annotations in MAGE-TAB format: <ul style="list-style-type: none"> <li>▪ The user can generate and download a MAGE-TAB file set describing the annotations of an experiment.</li> <li>▪ The generated MAGE-TAB includes the relationship between biomaterials, hybridizations and data files, and also characteristics of the biomaterials.</li> <li>▪ The generated MAGE-TAB does not contain information about experimental factors, protocols, publications or people.</li> </ul> </li> <li>◦ Bulk-update of experiment annotations using MAGE-TAB import: <ul style="list-style-type: none"> <li>▪ When a MAGE-TAB file set is being imported, the System recognizes references to existing biomaterials and hybridizations by name.</li> <li>▪ Biomaterial linkages can be modified in an additive way via MAGE-TAB import. Existing linkages and biomaterials cannot be deleted via MAGE-TAB import.</li> <li>▪ Biomaterial characteristics can be modified or new characteristics added via MAGE-TAB import.</li> </ul> </li> <li>◦ Support for GEO and ScanArray data: <ul style="list-style-type: none"> <li>▪ Data files can be designated as GEO SOFT or GEO GSM and then imported (without parsing) and associated with samples.</li> <li>▪ ScanArray CSV data files can be imported (without parsing) and associated with samples.</li> </ul> </li> <li>◦ Enhanced usability of permissions management workflow: <ul style="list-style-type: none"> <li>▪ While setting sample-selective access for an experiment, the user can do a keyword search for Samples against a set of categories, select the Samples of interest, and grant access to the selected Samples as a whole.</li> </ul> </li> </ul> </li> </ul>	January 2009
2.1.1	<b>Become part of the Life Science Distribution</b>	October 2008