The Rocks Have Names

Integrating IGSN with an Existing Digital Sample Catalog

University of TX Bureau of Economic Geology's Core Research Centers

Aaron Averett – Research Scientist Associate

TEXAS Geosciences
Bureau of Economic Geology
Jackson School of Geosciences
The University of Texas at Austin

Bureau of Economic Geology

Lots and Lots and Lots of rocks

- The Bureau of Economic Geology is a leader in the curation of rock cores, cuttings and other geological samples
  - Over 550,000 individual samples in approximately 1,000,000 boxes
  - Three core research centers across Texas

Austin

Houston

Midland

Bureau of Economic Geology

Photo credit: Aaron Averett, Beverly DeJarnett, Nathan Ivicic - Bureau of Economic Geology

What is SESAR?

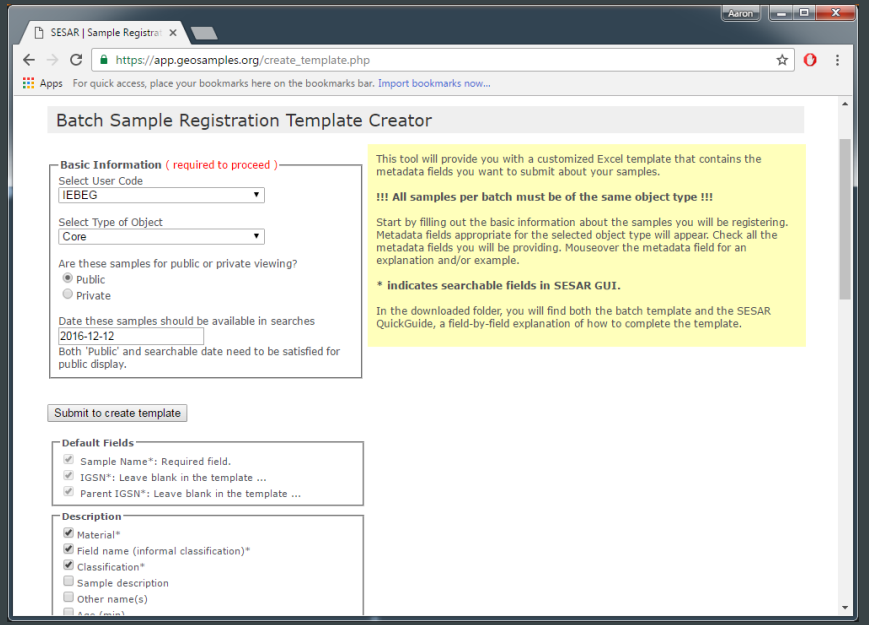
- "System for Earth Sample Registration"
  - Hosted by Lamont-Doherty Earth Observatory at Columbia University
  - Global registry of geologic samples
- What is an IGSN?
  - "International Geo Sample Number"
  - Globally unique identifier for a particular sample provided by SESAR
  - Useful for citing specific samples in literature unambiguously
  - http://www.geosamples.org

Lets do this!

- Option 1: XML-Based Web Service
  - Returns IGSNs immediately
  - Implies a software development project
- Option 2: Upload Excel files to GeoSamples.org
  - Generate template via MySESAR
  - Export relevant records from your catalog
  - Copy data to empty template
  - Submit file to MySESAR
  - Wait for response (can take a few days)
  - Join IGSNs to samples in your catalog

Customizing the Template

- SESAR format captures least common denominator for geologic samples – little application-specific information is supported
- Pick the fields you can fill
- Few fields are *required*…
  …but it looks like they are. Read carefully.



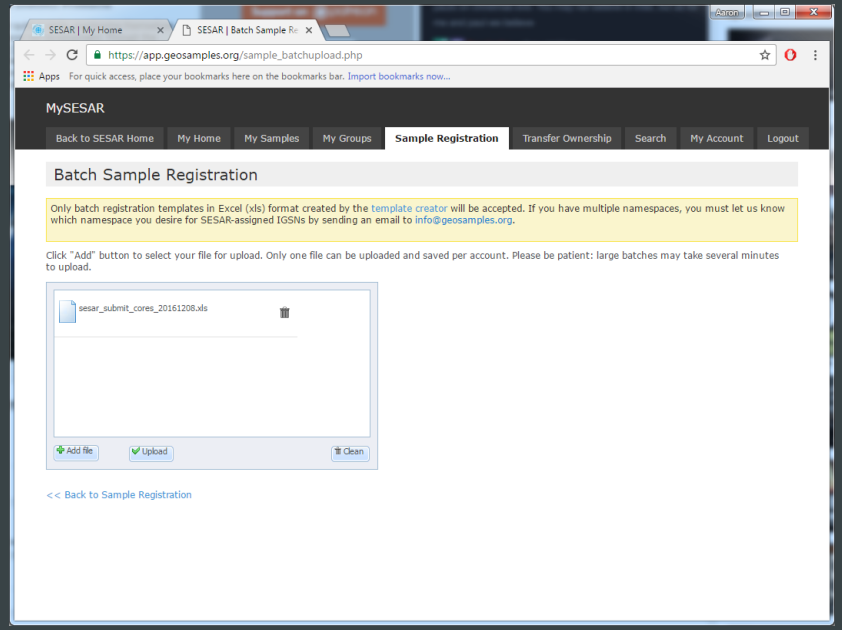
Prepare Your Data

- Few fields are required, but many have constraints
- Ex: Longitude must be between -180 and 180
- Easiest to use SQL Server, and export finished data
- Export to excel, rename columns, copy to empty template file

```
SELECT * FROM wellinfo where longitude < -115 OR longitude > -92

UPDATE wellinfo SET state_abbr=null
UPDATE wellinfo SET state_abbr = LEFT(RIGHT(county, LEN(county) - CHARINDEX(',', county)), 10) WHERE CHARINDEX(',', county) > 0

SELECT * FROM wellinfo where state_abbr IS NOT NULL
update wellinfo SET state_name = t2.[US State] FROM wellinfo AS t1 INNER JOIN [state_abbrev] AS t2 ON LTRIM(RTRIM(t1.state_abbr)) = t2.[Abbreviation]

SELECT * FROM wellinfo where state_abbr IS NOT NULL AND state_name IS NULL

update wellinfo set state_name = 'Texas' Where state_abbr is NULL AND state_name IS NULL


UPDATE wellinfo SET sesar_description = 'This is a sample of type ' + [sample_type] + ' from an oil or gas well.' WHERE [api_number] is NULL

alter table wellinfo ADD sesar_current_archive NVARCHAR(255)
alter table wellinfo ADD sesar_currevent_archive_contact NVARCHAR(500)

UPDATE wellinfo SET sesar_currevent_archive_contact =
'Physical: Bureau of Economic Geology, The University of Texas at Austin, 10100 Burnet Rd., Bldg 131, Austin, TX 78758-4445' + CHAR(13) + CHAR(10) +
'Email: nathan.ivicic@beg.utexas.edu; brandon.williamson@beg.utexas.edu'
 WHERE sesar_currevent_archive_contact IS NULL--cnum LIKE 'O%' OR cnum LIKE 'S%' OR cnum LIKE 'T%' OR cnum LIKE 'U%' OR cnum LIKE 'X%'


 SELECT * into wellinfo_public FROM wellinfo WHERE well_use_restricted = 0

 SELECT * into sesar_core FROM wellinfo WHERE well_use_restricted = 0 AND sample_type NOT like '%cuttings%' AND sample_type LIKE '%core%'

 UPDATE sesar_cuttings SET latitude = NULL, longitude= NULL WHERE longitude IS NULL OR latitude IS NULL

 alter table sesar_cuttings ADD Depth_scale VARCHAR(10)

 UPDATE sesar_cuttings SET Depth_scale = 'Ft.' WHERE top_depth IS NOT NULL AND bot_depth IS NOT NULL
```

Submit Your Data

- Splitting file into parts may be necessary
- Practical limit to file size is ~18,000 records
- Registration requires manual work at SESAR's end
- Results arrive via email



JOIN Response to Your Catalog

- Returned Excel file contains assigned IGSNs
- Load into SQL Server, join to samples table

1	Object Type:	Cuttings	All Public:	yes	Searchable date:	2016-12-09	User Code:
2	Sample Name	IGSN	Parent IGSN	Material	Sample description	Geological unit	Collection method
3	R61998	IEBEG1NIK		Rock	This is a sample of type CUTTINGS from		Coring
4	R61999	IEBEG1NIL		Rock	This is a sample of type CUTTINGS from		Coring
5	R61987	IEBEG1NIM		Rock	This is a sample of type CUTTINGS from		Coring
6	R61475	IEBEG1NIN		Rock	This is a sample of type CUTTINGS from		Coring
7	R61464	IEBEG1NIO		Rock	This is a sample of type CUTTINGS from		Coring
8	R61465	IEBEG1NIP		Rock	This is a sample of type CUTTINGS from		Coring
9	R61467	IEBEG1NIQ		Rock	This is a sample of type CUTTINGS from		Coring
10	R61468	IEBEG1NIR		Rock	This is a sample of type CUTTINGS from		Coring
11	R61469	IEBEG1NIS		Rock	This is a sample of type CUTTINGS from		Coring
12	R61470	IEBEG1NIT		Rock	This is a sample of type CUTTINGS from		Coring
13	R61471	IEBEG1NIU		Rock	This is a sample of type CUTTINGS from		Coring
14	R61472	IEBEG1NIV		Rock	This is a sample of type CUTTINGS from		Coring
15	R61427	IEBEG1NIW		Rock	This is a sample of type CUTTINGS from		Coring
16	R61474	IEBEG1NIX		Rock	This is a sample of type CUTTINGS from		Coring
17	R61460	IEBEG1NIY		Rock	This is a sample of type CUTTINGS from		Coring
18	R61476	IEBEG1NIZ		Rock	This is a sample of type CUTTINGS from		Coring
19	R61477	IEBEG1NJ0		Rock	This is a sample of type CUTTINGS from		Coring
20	R61478	IEBEG1NJ1		Rock	This is a sample of type CUTTINGS from		Coring
21	R61479	IEBEG1NJ2		Rock	This is a sample of type CUTTINGS from		Coring
22	R61482	IEBEG1NJ3		Rock	This is a sample of type CUTTINGS from		Coring
23	R61485	IEBEG1NJ4		Rock	This is a sample of type CUTTINGS from		Coring
24	R61473	IEBEG1NJ5		Rock	This is a sample of type CUTTINGS from		Coring
25	R61441	IEBEG1NJ6		Rock	This is a sample of type CUTTINGS from		Coring
26	R61542	IEBEG1NJ7		Rock	This is a sample of type CUTTINGS from		Coring
27	R61429	IEBEG1NJ8		Rock	This is a sample of type CUTTINGS from		Coring
28	R61432	IEBEG1NJ9		Rock	This is a sample of type CUTTINGS from		Coring
29	R61433	IEBEG1NJA		Rock	This is a sample of type CUTTINGS from		Coring
30	R61434	IEBEG1NJB		Rock	This is a sample of type CUTTINGS from		Coring
31	R61435	IEBEG1NJC		Rock	This is a sample of type CUTTINGS from		Coring
32	R61436	IEBEG1NJD		Rock	This is a sample of type CUTTINGS from		Coring
33	R61437	IEBEG1NJE		Rock	This is a sample of type CUTTINGS from		Coring
34	R61438	IEBEG1NJF		Rock	This is a sample of type CUTTINGS from		Coring
35	R61463	IEBEG1NJG		Rock	This is a sample of type CUTTINGS from		Coring
36	R61440	IEBEG1NJH		Rock	This is a sample of type CUTTINGS from		Coring
37	R61462	IEBEG1NJI		Rock	This is a sample of type CUTTINGS from		Coring
38	R61442	IEBEG1NJJ		Rock	This is a sample of type CUTTINGS from		Coring
39	R61443	IEBEG1NJK		Rock	This is a sample of type CUTTINGS from		Coring
40	R61444	IEBEG1NJL		Rock	This is a sample of type CUTTINGS from		Coring
41	R61446	IEBEG1NJM		Rock	This is a sample of type CUTTINGS from		Coring
42	R61447	IEBEG1NJN		Rock	This is a sample of type CUTTINGS from		Coring
43	R61448	IEBEG1NJO		Rock	This is a sample of ty	WILCOX	Coring
44	R61451	IEBEG1NJP		Rock	This is a sample of type CUTTINGS from		Coring
45	R61453	IEBEG1NJQ		Rock	This is a sample of type CUTTINGS from		Coring

Alternately - Submit to the Web Service

- Immediate response
- Allows integration with your application; automatic registration?
- Complex setup but less work in long term
- No (documented) update or delete functions – you'd better be sure your data is correct!

Using the Web Service

- Compose XML in SESAR's schema
- Craft HTTP request
- Submit Request
- Capture and parse XML response
- Handle response data
  - Save successful result (200)
  - Respond to error (400, 401)

LibSESAR_CSharp

- C# library encapsulates XML/HTTP submission logic
- Open source – available on GitHub
- MIT License
- Unit test code shown

```
[TestMethod]
public async void TestSuccessfulSubmit()
{
    //Create the sample submission request
    LibSESAR_CSharp.SESARSampleSubmissionRequest request = new LibSESAR_CSharp.SESARSampleSubmissionRequest();

    //Defaults to the test instance - be sure to set this to the production instance in your app
    //request.ServiceEndpoint = LibSESAR_CSharp.Constants.SESAREndpointProductionUrlBase;

    //Create the sample "collection" object
    request.Samples = new LibSESAR_CSharp.Models.samples();

    //Set your user credentials
    request.UserName = "aaron@myorg.edu";
    request.Password = "notR3allymp455w0rd";

    //Create and populate a sample
    LibSESAR_CSharp.Models.samplesSample sample = new LibSESAR_CSharp.Models.samplesSample();
    sample.name = "Q00001";
    sample.user_code = "IEBEG";
    sample.sample_type = LibSESAR_CSharp.Models.sample_type.Toothpick;
    sample.material = LibSESAR_CSharp.Models.material.Rock;
    sample.is_private = LibSESAR_CSharp.Models.is_private.Item1;
    //More fields here...

    //Add the sample(s) to the request.
    request.Samples.sample = new LibSESAR_CSharp.Models.samplesSample[1] { sample };

    //Execute the request against the web service
    await request.DoRequest();

    //Get the response object from our request object
    LibSESAR_CSharp.Models.SampleSubmissionResponse response = request.RequestResultModel;

    //We had a successful HTTP request, right?
    if(response == null || response.StatusCode != (int) System.Net.HttpStatusCode.OK)
    {
        Assert.Fail("Request did not succeed.");
    }

    if(response.SampleList.Count == 0)
    {
        Assert.Fail("Response has no sample records.  Did they change the service?");
    }

    //Success!  Access the IGSN as response.SampleList[0].IGSN;

}
```

GitHub Repo

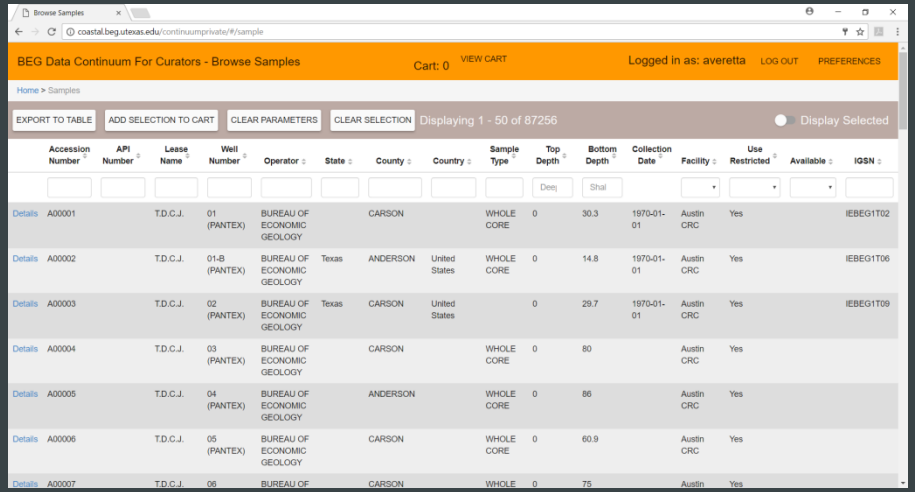
https://github.com/ut-beg/LibSESAR

What makes this hard?

- Documentation is sparse
- Data quality issues
  - Missing values
    - Go get them
    - Leave them out
- Lots of undocumented (but reasonable) requirements and limitations – Ex: can't specify depth units
- Web service requires a small software development effort
- Web service lacks JSON support – really verbose code (potentially)
- Requires a robust catalog application

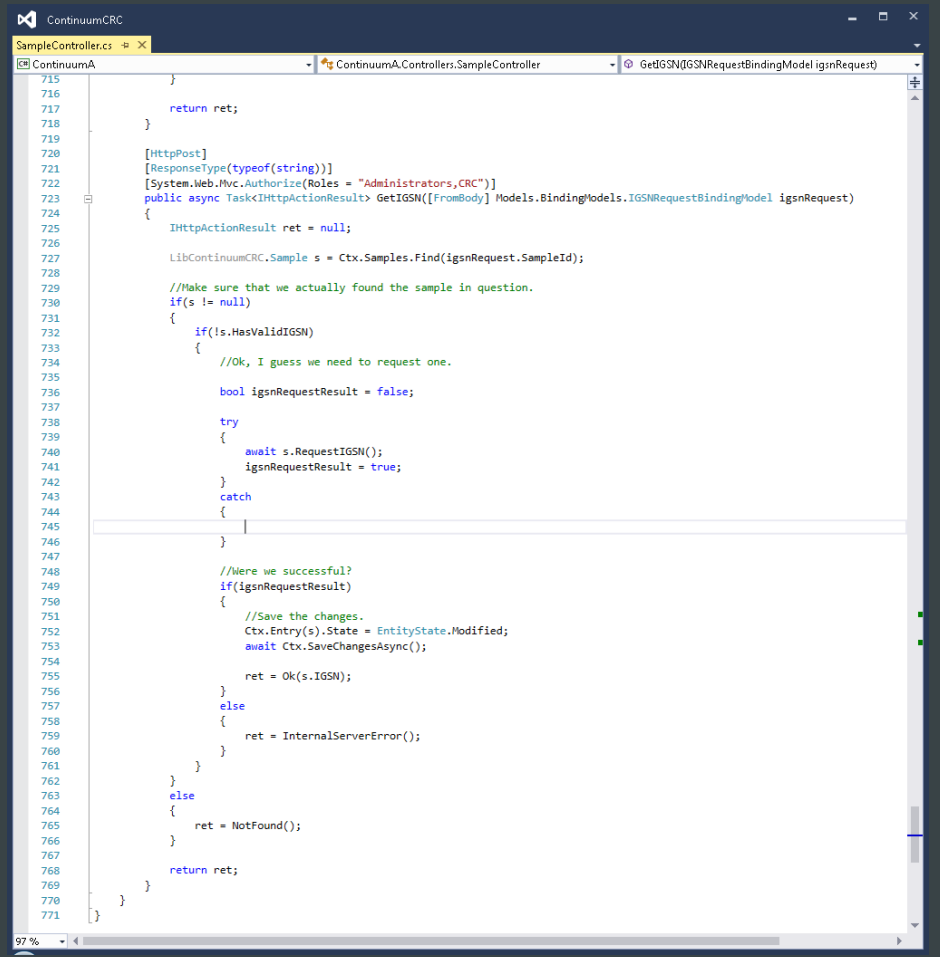
Continuum

- Enterprise sample and log catalog
- Shared across all facilities
- Empowers curators to manage collection
- Automates common management tasks
- IGSN Integration: Two parts
  - Server side
  - Client side



Server Side: Step 1

- API controller function
- Accepts Sample PK
- Check for IGSN registration allowed
- Call core library register function (async)
- Update record with new IGSN
- Notify client of result

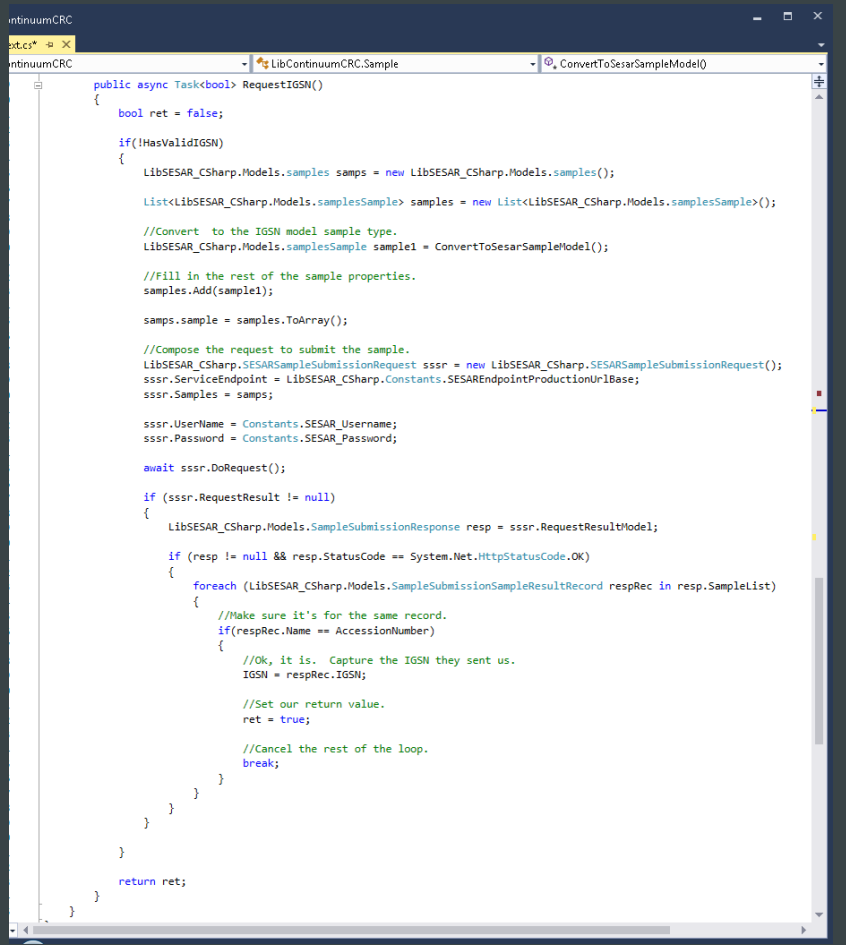



```
            }

            return ret;
        }

        [HttpPost]
        [ResponseType(typeof(string))]
        [System.Web.Mvc.Authorize(Roles = "Administrators,CRC")]
        public async Task<IHttpActionResult> GetIGSN([FromBody] Models.BindingModels.IGSNRequestBindingModel igsnRequest)
        {
            IHttpActionResult ret = null;

            LibContinuumCRC.Sample s = Ctx.Samples.Find(igsnRequest.SampleId);

            //Make sure that we actually found the sample in question.
            if(s != null)
            {
                if(!s.HasValidIGSN)
                {
                    //Ok, I guess we need to request one.

                    bool igsnRequestResult = false;

                    try
                    {
                        await s.RequestIGSN();
                        igsnRequestResult = true;
                    }
                    catch
                    {

                    }

                    //Were we successful?
                    if(igsnRequestResult)
                    {
                        //Save the changes.
                        Ctx.Entry(s).State = EntityState.Modified;
                        await Ctx.SaveChangesAsync();

                        ret = Ok(s.IGSN);
                    }
                    else
                    {
                        ret = InternalServerError();
                    }
                }
            }
            else
            {
                ret = NotFound();
            }

            return ret;
        }
    }
}
```

Bureau of Economic Geology

Server Side: Step 2

- Compose submission request
- Submit HTTP request
- Handle response



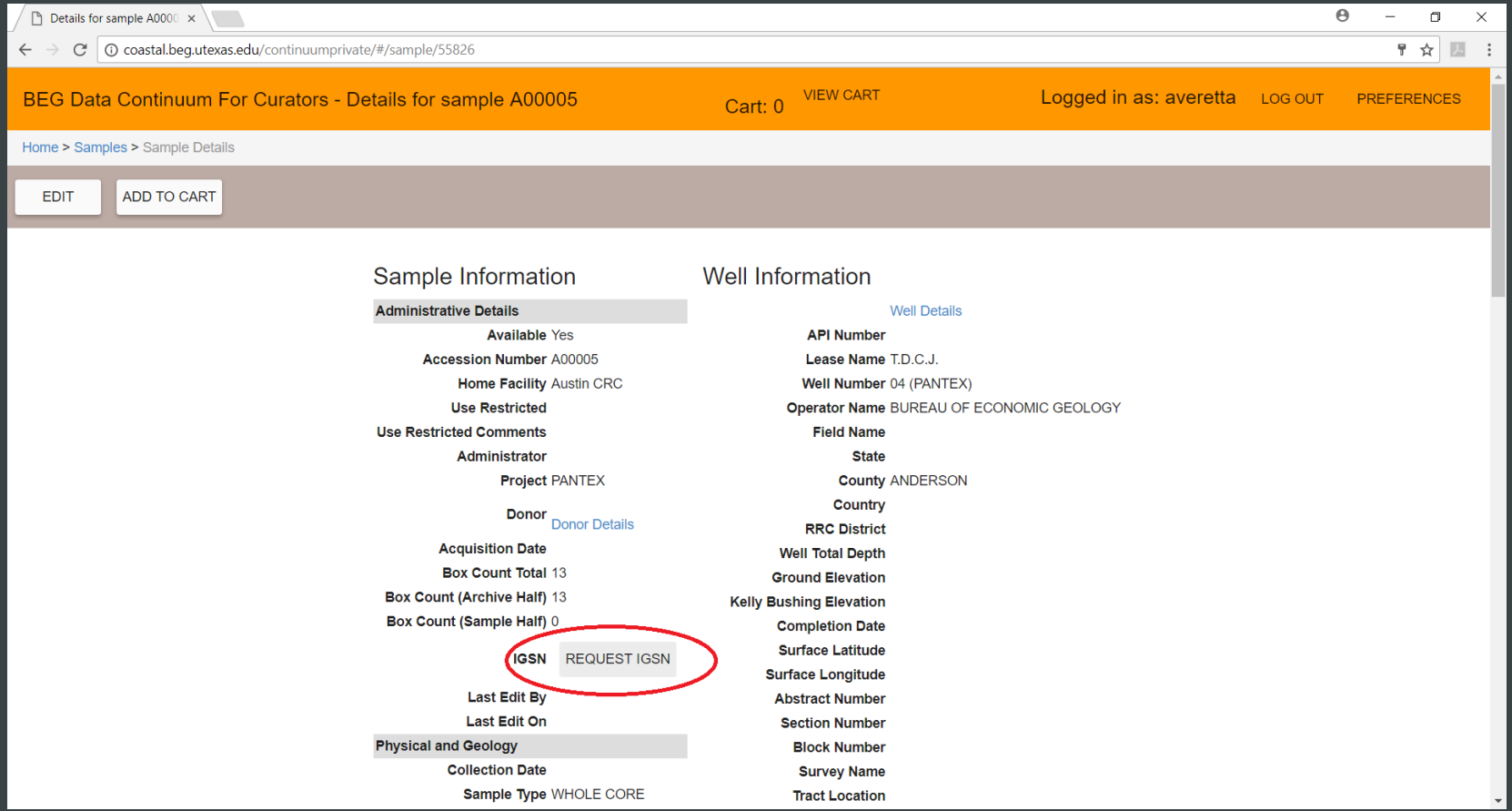
```
ntinuumCRC
ext.cs*
ntinuumCRC                    LibContinuumCRC.Sample                    ConvertToSesarSampleModel()

        public async Task<bool> RequestIGSN()
        {
            bool ret = false;

            if(!HasValidIGSN)
            {
                LibSESAR_CSharp.Models.samples samps = new LibSESAR_CSharp.Models.samples();

                List<LibSESAR_CSharp.Models.samplesSample> samples = new List<LibSESAR_CSharp.Models.samplesSample>();

                //Convert  to the IGSN model sample type.
                LibSESAR_CSharp.Models.samplesSample sample1 = ConvertToSesarSampleModel();

                //Fill in the rest of the sample properties.
                samples.Add(sample1);

                samps.sample = samples.ToArray();

                //Compose the request to submit the sample.
                LibSESAR_CSharp.SESARSampleSubmissionRequest sssr = new LibSESAR_CSharp.SESARSampleSubmissionRequest();
                sssr.ServiceEndpoint = LibSESAR_CSharp.Constants.SESAREndpointProductionUrlBase;
                sssr.Samples = samps;

                sssr.UserName = Constants.SESAR_Username;
                sssr.Password = Constants.SESAR_Password;

                await sssr.DoRequest();

                if (sssr.RequestResult != null)
                {
                    LibSESAR_CSharp.Models.SampleSubmissionResponse resp = sssr.RequestResultModel;

                    if (resp != null && resp.StatusCode == System.Net.HttpStatusCode.OK)
                    {
                        foreach (LibSESAR_CSharp.Models.SampleSubmissionSampleResultRecord respRec in resp.SampleList)
                        {
                            //Make sure it's for the same record.
                            if(respRec.Name == AccessionNumber)
                            {
                                //Ok, it is.  Capture the IGSN they sent us.
                                IGSN = respRec.IGSN;

                                //Set our return value.
                                ret = true;

                                //Cancel the rest of the loop.
                                break;
                            }
                        }
                    }
                }

            }

            return ret;
        }
    }
```

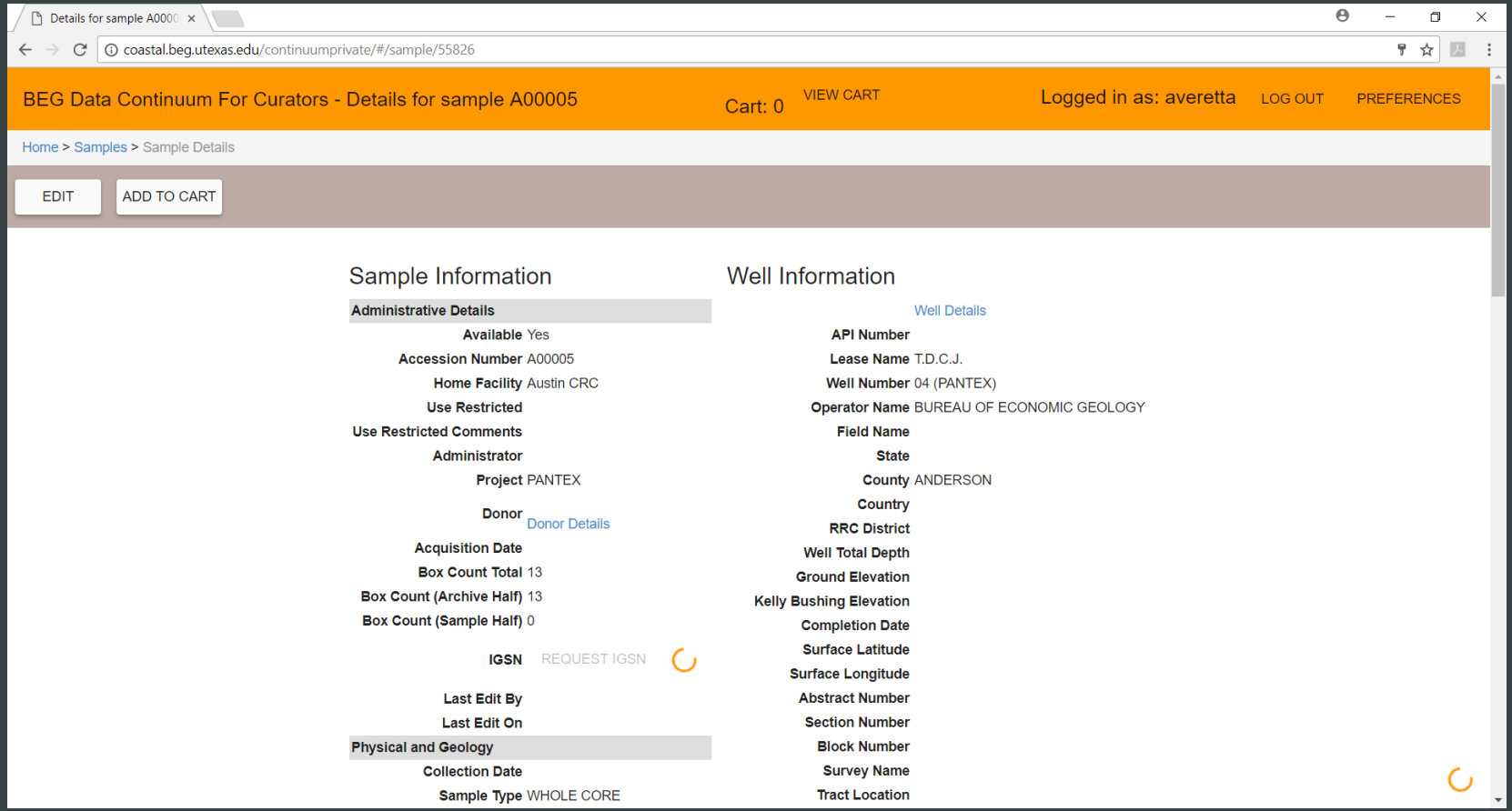
Bureau of Economic Geology

Client Side: UI

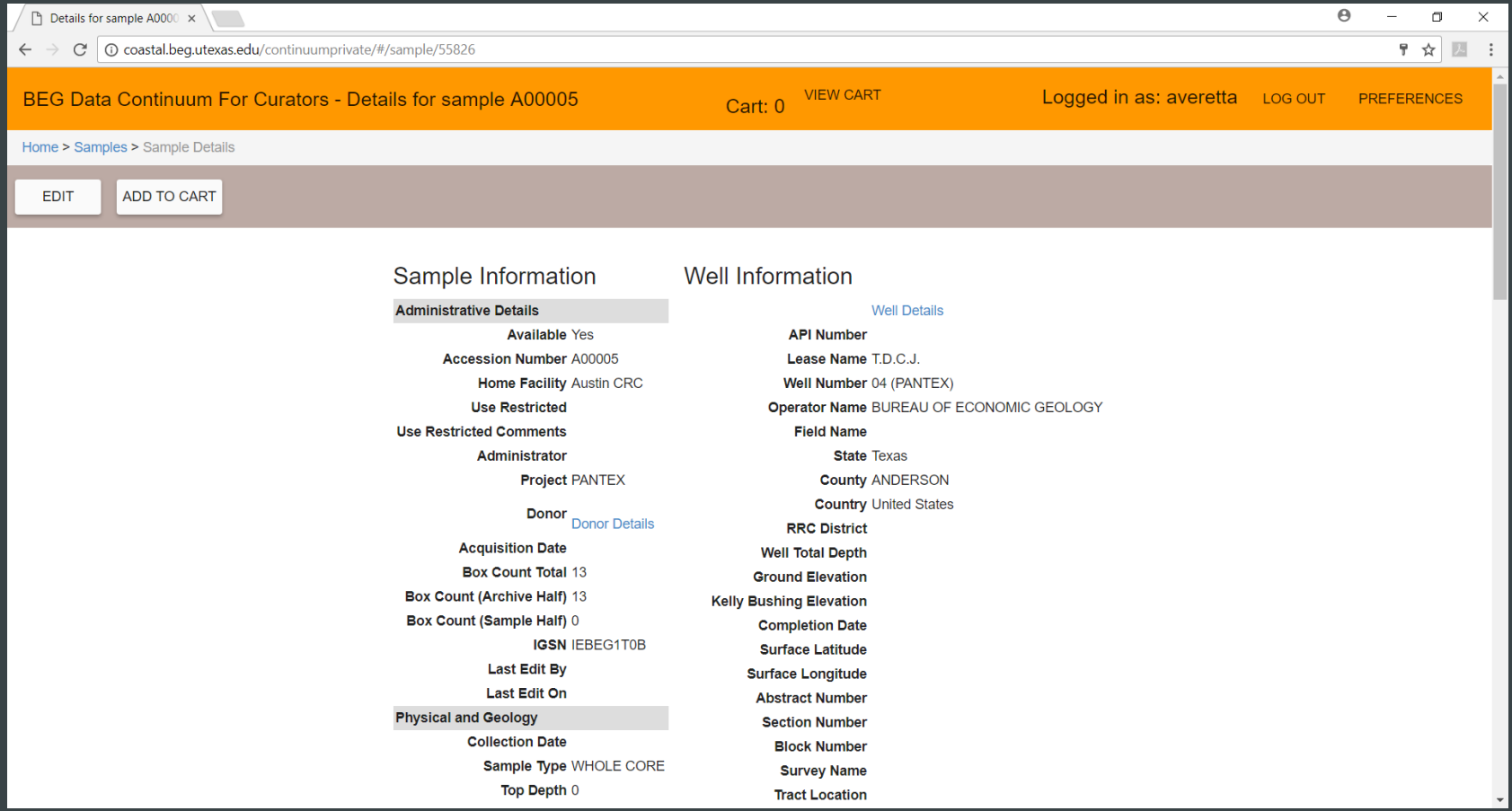
- Angular MVC framework
  - Javascript, HTML
- Controller function calls Web API asynchronously
- Promise API
  - Resolve: Update UI
  - Reject: Notify user
- Refresh displayed sample

```
$scope.requestIGSN = function (ev) {
    $scope.requestIgsnIsWorking = true;

    //Call Web API Controller function, handle promise resolution.
    ContinuumSamplesService.requestIGSN($scope.sample).then(

        //Successful registration - update loaded sample properties
        function (result) {
            ContinuumSamplesService.fetchObjectWithId($scope.sample.SampleId).then(
                function (result) {
                    $scope.requestIgsnIsWorking = false;
                    $scope.sample = result; //Update the UI
                },
                function (result) {
                    $scope.requestIgsnIsWorking = false;
                });
        },
        function () {

            //Registration failed
            $scope.requestIgsnIsWorking = false;

            $mdDialog.show(
                $mdDialog.alert()
                .parent(angular.element(document.body))
                .clickOutsideToClose(true)
                .title("Error")
                .textContent("An error occurred while attempting to register this sample for an IGSN.")
                .ariaLabel("IGSN Registration Error")
                .ok("Ok")
                .targetEvent(ev));
        });
};
```


Details for sample A0000
coastal.beg.utexas.edu/continuumprivate/#/sample/55826
BEG Data Continuum For Curators - Details for sample A00005
Cart: 0
VIEW CART
Logged in as: averetta
LOG OUT
PREFERENCES
Home > Samples > Sample Details
EDIT
ADD TO CART
Sample Information
Administrative Details
Available Yes
Accession Number A00005
Home Facility Austin CRC
Use Restricted
Use Restricted Comments
Administrator
Project PANTEX
Donor
Donor Details
Acquisition Date
Box Count Total 13
Box Count (Archive Half) 13
Box Count (Sample Half) 0
IGSN
REQUEST IGSN
Last Edit By
Last Edit On
Physical and Geology
Collection Date
Sample Type WHOLE CORE
Well Information
Well Details
API Number
Lease Name T.D.C.J.
Well Number 04 (PANTEX)
Operator Name BUREAU OF ECONOMIC GEOLOGY
Field Name
State
County ANDERSON
Country
RRC District
Well Total Depth
Ground Elevation
Kelly Bushing Elevation
Completion Date
Surface Latitude
Surface Longitude
Abstract Number
Section Number
Block Number
Survey Name
Tract Location

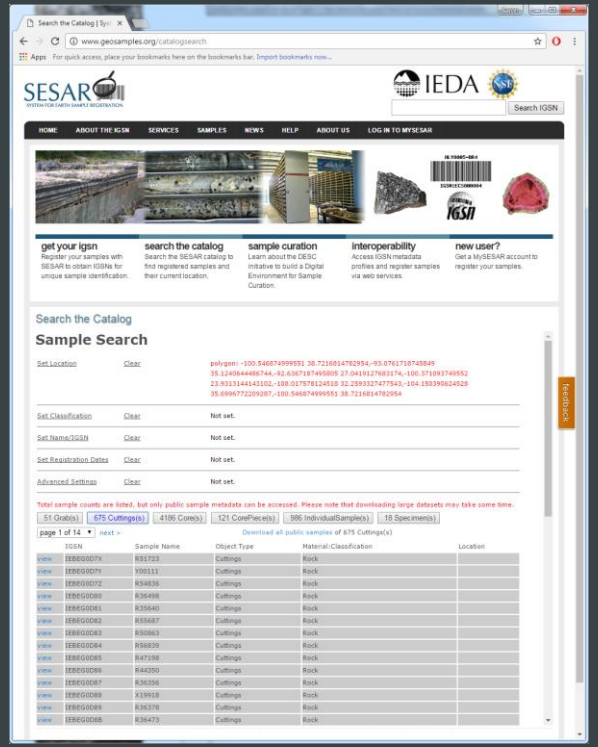


Bureau of Economic Geology


Details for sample A0000
coastal.beg.utexas.edu/continuumprivate/#/sample/55826
BEG Data Continuum For Curators - Details for sample A00005
Cart: 0
VIEW CART
Logged in as: averetta
LOG OUT
PREFERENCES
Home > Samples > Sample Details
EDIT
ADD TO CART
Sample Information
Administrative Details
Available Yes
Accession Number A00005
Home Facility Austin CRC
Use Restricted
Use Restricted Comments
Administrator
Project PANTEX
Donor
Donor Details
Acquisition Date
Box Count Total 13
Box Count (Archive Half) 13
Box Count (Sample Half) 0
IGSN IEBEG1T0B
Last Edit By
Last Edit On
Physical and Geology
Collection Date
Sample Type WHOLE CORE
Top Depth 0
Well Information
Well Details
API Number
Lease Name T.D.C.J.
Well Number 04 (PANTEX)
Operator Name BUREAU OF ECONOMIC GEOLOGY
Field Name
State Texas
County ANDERSON
Country United States
RRC District
Well Total Depth
Ground Elevation
Kelly Bushing Elevation
Completion Date
Surface Latitude
Surface Longitude
Abstract Number
Section Number
Block Number
Survey Name
Tract Location

That’s, like, a lot of work, man…

- Unambiguous naming – easy to cite in literature
- Enhanced discoverability with central search engine
- Integration with enterprise catalog accelerates registration
- Prevent registration mistakes



Conclusion

- Web service allows integration with catalog
- LibSESAR_CSharp – asbtraction layer on XML/Web Service