CESSDA Research Infrastructure

Technical Framework



John Shepherdson IASSIST 2016, Bergen Norway

cessda

Structure

Introduction to CESSDA

Technical Framework (Phase 1)

What's next?

What is CESSDA?

Consortium of European Social Science Data Archives (http://cessda.net)

Currently 15 members, 1 observer

Aim: Provide a pan-European Research Infrastructure for the Social Sciences

CESSDA Objectives

- coordination of the network of European data service providers and promotion of the results of social sciences
- facilitating researcher access to important resources of relevance to the European social science research agenda regardless of the location of either researcher or data
- work continuously to include further data sources from Europe and beyond, into the infrastructure

CESSDA Objectives

- provide training within CESSDA and beyond on best practices surrounding operational processes and data management
- promote and facilitate wider participation in CESSDA
- the development and coordination of standards, protocols and professional best practice pertaining to the preservation and dissemination of data and associated digital objects

CESSDA Current Activity

Horizon 2020 Projects

- CESSDA <u>Strengthening and Widening</u> (SaW)
- Synergies for Europe's Research Infrastructures in the Social Sciences (SERISS)
- Big Data Europe (BDE)

Work plan Projects

- Expert Seminars
- Metadata Harvester
- Metadata Management
- Data Access Policy
- Technical Framework

Technical Infrastructure Required

In past, relied on members ('Service Providers') to develop and host standalone products

Now a company under Norwegian Law and in process of becoming an ERIC (<u>European</u> Research Infrastructure Consortium)

=> Greenfield site for technical development

Protection of Assets is Vital

Ensure CESSDA has access to

- source code
- configuration files
- technical documentation

that underpin its products and services

CESSDA's common interoperability characteristics

- Loosely coupled but coordinated enable Service
 Providers to retain independence, yet fully interact in an integrated service
- 2. **Sustainable** enable medium and long term investment and business change decisions to be made.
- 3. **Extensible** enable additional services to be built on or around it, including adapting to changing functional requirements over time.
- 4. **Maintainable** enable components to be updated when IT specifications change.
- 5. **Standards based** enable the coordinated and planned change to all the coupled, but coordinated, services.

What is the Technical Framework?

A guide for the development of the various (software) products and services that form part of the CESSDA Research Infrastructure

- promote good practice for software development
- meet common interoperability characteristics

Infrastructure for Development, Staging and Production

harmonise development tool chain for SPs

Partners and Activities

Partners DANS, GESIS, NSD, SND, UKDS (lead)

All partners

Develop and publish Technical Architecture

UKDS and **NSD**

Specify, install and configure development infrastructure

Infrastructure for Development (to date)

- Email Identity c/w Avatar
- Collaborative document creation/management
- Source code management
- Issue Tracker
- Wiki
- File storage
- Chatroom
- Continuous Integration Server
- System Availability Monitor
- Micro-service registry

Quality Control

Software Maturity Levels

- ensure quality of the research infrastructure is maintained
- guidance on minimum, expected and excellent standards
- originally based on <u>NASA's RRLs</u>
- revised in light of 'Capability Development Model' from <u>CESSDA SaW</u> project

Common Interoperability Characteristics

Achieved by:

- REST APIs c/w API design standards
- Architectural standards
- Common development environment
- Adoption of <u>12 Factor App</u> principles
 - methodology for building Software-as-a-Service apps

1. Loosely coupled but coordinated

Adopt (micro) services architecture based on RESTful web service APIs

 provides a mechanism for reusing and combining software artefacts

See also 12 factor app, number 7 (Port binding - Export services via port binding)

2. Sustainable

The provision of common standards

Technical Architecture document

Common development and test environment

via the technical infrastructure

Deployment environment

via extensions to the technical infrastructure

Central source-code repository

See also 12 factor app, number 1 (<u>Codebase - One</u> codebase tracked in revision control, many deploys)

3. Extensible

Service API is key

- Integration point for new services
- Combination point for building new features
 Version and support two versions simultaneously
 - Allows services to evolve, without breaking contract provided to consumers.

See also 12 factor app, number 8 (<u>Concurrency - Scale out via the process model</u>)

See also 12 factor app, number 9 (<u>Disposability - Maximize</u> robustness with fast startup and graceful shutdown)

4. Maintainable

Again, service API is key

- implementation of a service can be changed as required, to take advantage of developments in software technology
- location of services can be changed as required, to take advantage of developments in hardware technology.

See also 12 factor app, number 2 (<u>Dependencies - Explicitly</u> <u>declare and isolate dependencies</u>)

5. Standards Based

- Provision of common architectural standards (via Technical Architecture)
- A consistent (in both the calling and return structures and formats) and versioned API

See also 12 factor app, number 4 (<u>Backing services - Treat</u> <u>backing services as attached resources</u>)

Software Acceptance Criteria

- Documentation
- Intellectual property issues
- Extensibility
- Modularity
- Packaging
- Portability
- Standards compliance
- Support
- Verification and testing
- Internationalisation and Localization

Software Maturity Levels

- 1 Initial usability; software use is not recommended.
- 2 Use is feasible; the software can be used by skilled personnel but with considerable effort, cost and risk.
- 3 Use is possible by most users; with some effort, cost, and risk. A risk assessment should be made before use.
- 4 Software is usable; with little effort, cost, and risk.
- **5 Demonstrable usability**; there is clear evidence that the software is widely used by many users.

Software Maturity Levels

Intellectual Property (CRE)	incombility (CAS)	Modularity (XA4)	Pedraging (CAS)	Potability (CNI)	Standards Congliance (CAP)	SupportitAtt	Numbers and Yesting (CAS)	Security 20,400.	tenantisticostication and tocatamien (CASC)
Software developed have been observed and their requirements have been determined. Redressed policies of developed cognitional ordinal districtions in been entered by a applicability in inhibition to incorpora rights. Their inciples residence of a development and applications in inhibition to incorporate policy. Their inciples residence of a development properties approximately and their applications of the second second from the second properties. The second second companies are second as a second of the second sec	mend, so there is either no ability to extend or mostly program behavior, or 6 is very difficult for	Overgreen on commissional for organizing the code of brining of Fundamental for modularity or som II, may free freem a streamental solar or glidd projess.	One secure code or description monthly, i.e. no packaging Those in this or incomplete, included on decommendation and for accommendation and for accommendation and for accommendation. There are represented one may have difficulties installing the sufference.	The tell-ways as a whole at not pertains, forever, if there is sent-y occurs used provided provided of the sufficient sites red, with real and accident lateral lat	defined standards and book processor. The Userdards may be internally an expensally assertion, but are	there is known-context who waston ameliate for the developer or professional or generalized and there is a wildingered to provide minimal, occupated arguest will have personal occur to one the self-was provided and to one the self-was without come to one the self-was without come to one the self-was without come and one.	polivers appropriet from white and and facility performed.	Security may informated in the production of districts on the production of the continuous districts.	transplanter and transplanter and transplanter and additional
	that responds sent for a historic moviner of use Lanes. Or registers of motivate such as a slight institution of single or other look which annulate looked softenin. Name patientiallity is parallel through continuous hanger; lealation	That are vide general functionally and arrest two subspillants, that each provide multiple, unrelated, functions, unde within teach multiple parties of teach multiple parties from the provides and the provides are the provided and the provided and the provides are the provided and the provided	and harbiting for other oppositing epiterm is associate for an importanced as an	estimate enter and disconsiderates that are not postable, but the robberson cannot be postable without algorithment alranges in the cultivase or the largest equation. Successive allows are providing to insufficient, and though the culture term consideration about the color of their considerations about the color of their experient Consideration and visit and page.	or recognitured standards or patiently serial bend propelliers, last and head or officialism are training and map may be complete. These standards perma deculmentated windstance than place of the control bend to a place of the control bend to a	expair had remains will requisition (parts her final are suppose in the available in a remainship freely facilities. Name supposed is available, but may be below williers, and all hour parameters of surfamilianation. Those is evaluated and as influenced over a commantly when provides remains, for example, you making do no substant.	and helped in a Monthury contains. Turking schedus helping for entire ammellions and grant of handling of politicate tripics.	Security was addressed in the development phase on to and traditions included to the development phase on to and traditions included included to the development have underlanded again, printle for unity tradition.	Software in Incohe person
Aprentismits on development resonantishings, the list of development, a recommended contains, and of inferioral controlled and advantage of the controlled and of inferioral controlled and advantage of the controlled and ad	a made of a range of our cases. The procedures for seconding the collection are believed, whether by accurate modellization or forming the provision of cases type of an involve functionally to a collection transfer or wholey countries. A func-	I disellating the relative consuments from a certain facilities of an exist to option within the consentation of the constitution functions as constant physical relative and constituting excellent than the consented for green's function, that matches have not been consisted for after the cycle of function, such additionable matches contains many interpretated legislations.	Plus following is server non-figure file for different contains; sold in a. Audition of many on (Pfer), device on "Effect on Leafing mobile. All configuration consults, other mobile is user that lead."	The obtaine is indicated postable. The sufficiency is provide will only entitled postablishing recipient to the certain of the sufficient text. Descriptionalities or certain extra will interested the require constituents order and reported toward contributes with and reported toward sufficients antiferiorities of the underlying privace on the topic system may be settleday.	proprietary durabelo, but More is incomplete on fluiden of complete.	Suppose? In conditional in a methods to adjust a product of instances, wherein the PRL offers and instances, wherein the PRL offers and instances, wherein the PRL offers are extended when the first and it is assumed by supposed upon their European drive for a feelf-time or excellence that the bound, There is evidence that the development approximation () in an approximation () in an approximation () is an approximation () in a supposed mentable of extenses of substances (of each other mentable in extenses of substances (of each other mentable in extenses of substances) (in the extenses of the extens	yorket.	Notice was obtained in the Growing and places up to tall including implementation.	Contain to localism in experted to receive godes
Name a androno that all developes organizationals for developed have confirmed that the field of developed, incommended distinct, and mithicularly properly rights distincted. An excluding the case in the california relationship or an excluding the field of the sufficient operation of the case in the california operation and the california operations are large large page. The field has been approved by all patients and their representations, we first matter to be sufficient or their representations, we after matter the case of the large page that a property and impacts a supercentify indicational property, and terration statements in large page that can be understood for page and an a part artifact, are explained to border. And in page case the case of consideration of the page and an artifact, and consideration of the page and an artifact, and consideration of the page and an artifact of the california formation is an acceptance of the page and consideration of the page and conside	defined, found simple of any cases, providing them provide of extraorbidity. A detailed extraorbidity plan is printed or of extraorbidity and its software to allow on any project in a restracted the software for the project in cases and the software of the software for any other than a state of the software for a restricted the software for software for the software for th	convent regist to with considered documentations of all the series in Africa in dipolated until one rice retain faces. Modules, lines become uncontrol for all spicialized formations and represent that these two with accountered fundamental time from laters. Incl. success codes within association fundamental transport of the senior manifest study controls many technique values legical patter.	one capacidad philippin with 21 detection configuration from equilibrium with rate of repression best for philippin-specific to certing. Uninequipoles is reasonably single-forward with the	allocates systems without modification. The discontrollation is samplifier and hisrargin. The changes in this colleges are unassay and the effect in part the software is minimal.	development process contain with recognized and progressing demands. Compilares with these stempt testing for all components, line stempt stempt testing for all components, line stempt conducts for without conducts for without conducts and the widdleston	Where is proposed and clearly defined respect by the developer with an extent helpfork and additional decorrentation.	demonstration (resets requirements), and canadrabelly deference.	Secontly was obtained in the directorment abuses as 50 and 1 multiplier and feature, and feature,	Carinari to Notalius (faet) becont, ginafrici areal endificendia, septiment affanciaris, Rondo, fanule siste em i sharpa ter esta, bashii panameli has lapan lekeresaffanaling di
Figure on a sufficient comments assumed that the pull many contest data strong all wast stood rights and any conditions for puts including contested and put our consequent and part for a summarised contest, and the put of the summarised contest of the put of the summarised contest. The put of the summarised contest of the collection o	Note a solution that the officers for their uncontrol relevants from provide at the original large from the original large from a provide at the original large from a public reproduct for solution and the original large from a public reproduct for solutions and solutions are original large from the original large fro	Signal of incomplete larged each on your standard. There is considered in the large large of the secondard incomplete incomplet	notice:		Mineralizado, ecopicado interchela for the collection and collection of security produces to industry and discontinuous production through to long of all companiones dessignation or discontinuous dessignation dessign	charty defined with frequently one from packets, response to the packets of the p	orbitation for our constitution of constitution conjust		contracts that have to man with provide that of them.

Meaning				
Minimum standard				
Expected standard				
Excellent standard				

cessda

Maturity Modelling - More Info

Mike Priddy, CESSDA SaW project

Wednesday 15:45

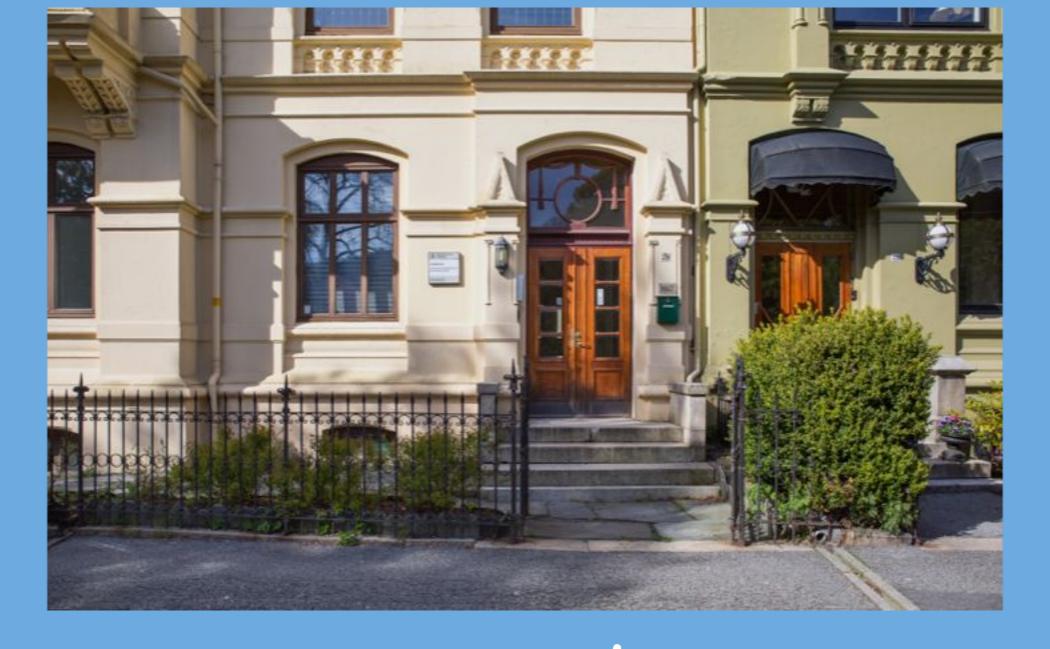
Session 1G - Data Services: Setting up and evaluating

'Maturity Model For Assessing Data Infrastructures – Cessda As Example'

What's Next?

Proposed Phase 2

- more dissemination understanding and acceptance
- Staging and Production environments
- extend 'how to' documentation for developers and admins



cessda

A pan-European Research Infrastructure for the Social Sciences