

The consortium is working on tools, services and methodologies that can be exploited by organizations like software vendors, integrators, and communities building, maintaining and using data-intensive systems like Wolters Kluwer JURION (specialised information systems), DBpedia (web data), Seshat (e-Research for humanities and social science) and SWC PoolParty (commercial application development).

This covers challenges like:

- Thesaurus and taxonomy building
- Support for Linked Data consistency
- Metadata and Linked Data strategies
- Developing digital services for publishers
- New Business models and the Web of Data
- Foundations for a sustainable metadata strategy
- Introducing efficient content processing workflows
- Accessing large scale historical and archaeological data sets
- Processing much larger amounts of data with the same effort
- Archaeological foundations for the study of the evolution of human societies
- Supporting a highly-iterative agile development methodology through model-driven development
- Constructing and managing precise models documenting datasets and data sources

The technologies used are based on standards, so that adoption is easy.

In order to show the potential of the ALIGNED approach, the consortium is offering consulting services that interested parties can apply for during the course of the project. Calls on the website are open each year.

Stay tuned and apply!

### Dr. Rob Brennan

Project Coordinator  
Tel: +353 1 896 2181 (direct)  
Email: rob.brennan@cs.tcd.ie

### Address

Knowledge and Data Engineering Group  
School of Computer Science and Statistics  
Trinity College Dublin  
Dublin 2, Ireland

### Christian Dirschl

Dissemination Leader  
Tel: +49 (89) 36007-3158  
Email: CDirschl@wolterskluwer.de

### Address

Wolters Kluwer Deutschland GmbH  
Freisinger Straße 3  
85716 Unterschleißheim, Germany

### Project

info@aligned-project.eu  
twitter: <https://twitter.com/alignedproject>  
Website: <http://aligned-project.eu/>

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 644055. This communication reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

## Aligned, Quality-centric, Software and Data Engineering

### Focus areas:

- **Increased agility and productivity for data-intensive applications**
- **Reducing the effort to produce and maintain high quality data**
- **Enabling fully integrated software and data engineering teams**
- **New methods to integrate Linked Data in applications**

Enterprise Linked Data  
Model-Driven Engineering  
Data Quality



# Project Summary

ALIGNED is a research and innovation project within the Software and Services, Cloud Unit (E2) of the European Commission. It started in February 2015 and completes in January 2018. The focus of ALIGNED is on developing tools and methods to increase productivity in building and maintaining IT systems that leverage web-data.

ALIGNED technology will help organisations address these questions:

## 1. How can I integrate my software engineering and data engineering teams and processes?

The ALIGNED software and data engineering methodology will provide a basis for unified management views of co-evolving software and data systems. Tool and process interoperability will be increased by publishing and consuming enterprise Linked Data using ALIGNED vocabularies that describe the lifecycle of the software and data engineering processes.

## 2. How can I assure application quality for systems built on web data?

ALIGNED is developing model-driven development techniques for Linked Data applications that provide quality guarantees through automated consistency and integrity checks and methods for integrated software and data testing that address the gaps in current, siloed approaches to software and data testing.

## 3. How can I easily integrate Linked Data into my applications?

Linked Data sources are typically inconsistent, incomplete and highly heterogeneous in structure. ALIGNED tools and methods will address these barriers to their reuse in applications.

ALIGNED system specifications can natively include or reference external Linked Data sources enabling automatic generation of application workflows, data manipulation user interfaces and specification of application level data quality constraints. ALIGNED tools will allow developers to browse and select external datasets, automatically test and characterise these datasets in terms of data quality metrics and to support semi-automated data curation processes aimed at increasing data quality to the level required for a specific application.

## 4. How can I efficiently produce and maintain high quality web-scale datasets and applications?

Collection of high quality datasets requires huge effort and data quality degrades over time as data sources change or new data becomes available and old data becomes stale or outdated. ALIGNED will build semi-automated data curation methods and systems that manage data through the entire lifecycle and both minimise the human effort required of domain experts and maximise the ability of non-experts to increase the productivity and quality of the data curation process.

# ALIGNED Use Cases

## DBpedia+ Quality Tool Chain

DBpedia is the centre of the current web of data. It publishes authoritative RDF-based datasets that are used as a common point of reference for interlinking and enriching most of the structured data on the web today. It is therefore an important basis for future linked data developments. ALIGNED will provide methods and tools to combine the diverse tool development and data lifecycle management tasks that are already underway within DBpedia for increased productivity and higher data quality.

## Wolters Kluwer JURION

JURION is an innovative legal information platform developed by Wolters Kluwer Germany that merges and interlinks over 1 million documents of content and data from diverse sources. By adopting the ALIGNED methodology and tools, software development and data processing pipeline maintenance will gain integrated governance mechanisms. ALIGNED will enable JURION to address more complex business requirements that rely on tighter coupling of software and data. The valuable results of this use case will be applicable to all kinds of knowledge systems.

## Seshat: Global History Databank

The Seshat global history databank is an international initiative of humanities and social science scholars to build an open repository of expert-curated historical time-series data, extracted from a combination of databases, Linked Data, web sites, academic publications and human experts. ALIGNED will develop new techniques for data curation that build on data quality metric frameworks for data integrity assurance, runtime quality analytics to automate and prioritise data curation tasks and curation workflows that link together data harvesters, domain experts and data consumers to improve data quality. These developments will not only enhance the sustainable use of the Seshat dataset but can also be applied to other ambitious dataset curation initiatives.

## PoolParty Enterprise Application Development

Since its first release in 2009, the Poolparty Enterprise Application has evolved continuously. The requirements engineering system for this process was mainly based on several loosely coupled tools – a situation that is encountered by many software companies. ALIGNED technology will help to generate unified views on distributed business objects and entities relevant for the development processes and to coordinate the data management and development workflows required to deliver new versions of the evolving PoolParty product.

# The ALIGNED Consortium

	Trinity College Dublin Coláiste na Tríonóide, Baile Átha Cliath The University of Dublin	Trinity College Dublin (Coordinator) <b>Ireland</b>
	UNIVERSITY OF OXFORD	University of Oxford <b>United Kingdom</b>
	UNIVERSITÄT LEIPZIG	Universität Leipzig <b>Germany</b>
	SEMANTIC WEB COMPANY linking data to knowledge	Semantic Web Company <b>Austria</b>
	Wolters Kluwer	Wolters Kluwer Deutschland <b>Germany</b>
	UAM	Universytet im. Adama Mickiewicza w Poznaniu <b>Poland</b>

# Tools & Platforms

	Seshat		DL-Learner
	PoolParty		ORE
	Dacura		Jurion
	DBpedia		qSKOS
	RDFUnit		Booster

# ALIGNED

Software and Data Engineering