



SOCRAT

**Platform Design: A Web Architecture for Interactive
Visual Analytics Applications**

Alexandr A. Kalinin, Selvam Palanimalai, Ivo D. Dinov
University of Michigan
SOCR.umich.edu, 2017

@alxndrkalinin

Motivation

- Visual Analytics (VA) – emerging field focused on analytical reasoning facilitated by interactive visual interfaces
- Web became successful platform for interactive apps
- SOCR.umich.edu – over 100 Java-based web-apps for modeling, analysis, and visualization
- Modern web tools: interactive visualizations (D³, Vega), analysis (jStat, Science.js), data management (Datavore)

How can we efficiently combine existing tools and new developments to build better VA apps?

Designing VA platform

What is important specifically for VA applications?

data preparation, representation and rendering

- interactivity – central piece of any VA app
- flexibility – all VA apps are different
- efficient use of existing tools – everyone uses D³
- automated data management (loading, storing, querying)

Design requirements

Integrative web VA system should:

- be flexible
- provide component interoperability
- allow easy extension
- be robust in runtime
- allow component reuse for various purposes
- emphasize expressiveness (allow low-level control)
- accessibility (simple protocols for creating new modules)

Design requirements

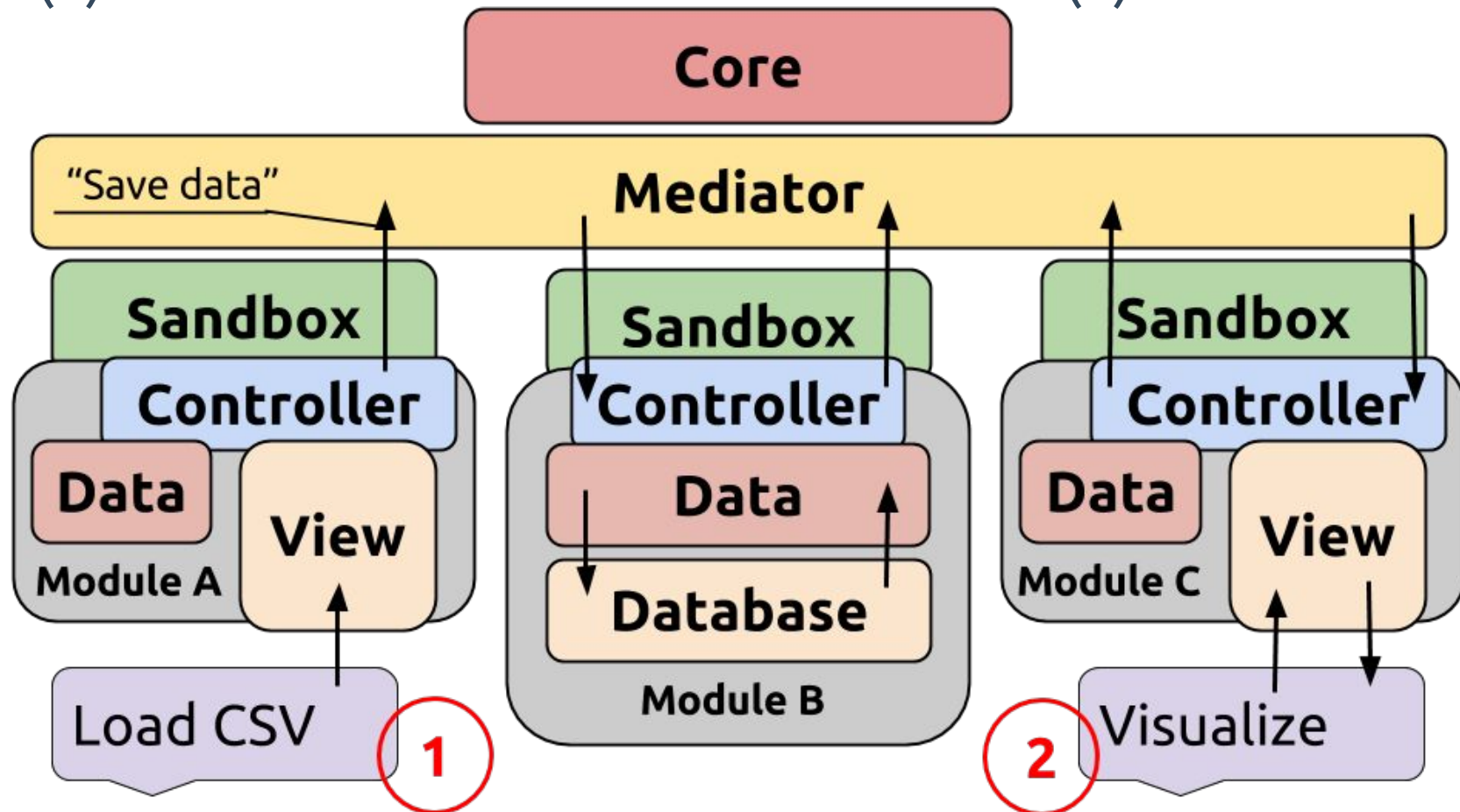
Suggested design:

- multi-level modularity:
 - system is modular
 - system components are modular
 - component modules are modular
- components are decoupled, isolated, and centrally managed (Core)
- optimized module interaction (messaging system)
- extensibility (automated integration)
- common web techniques

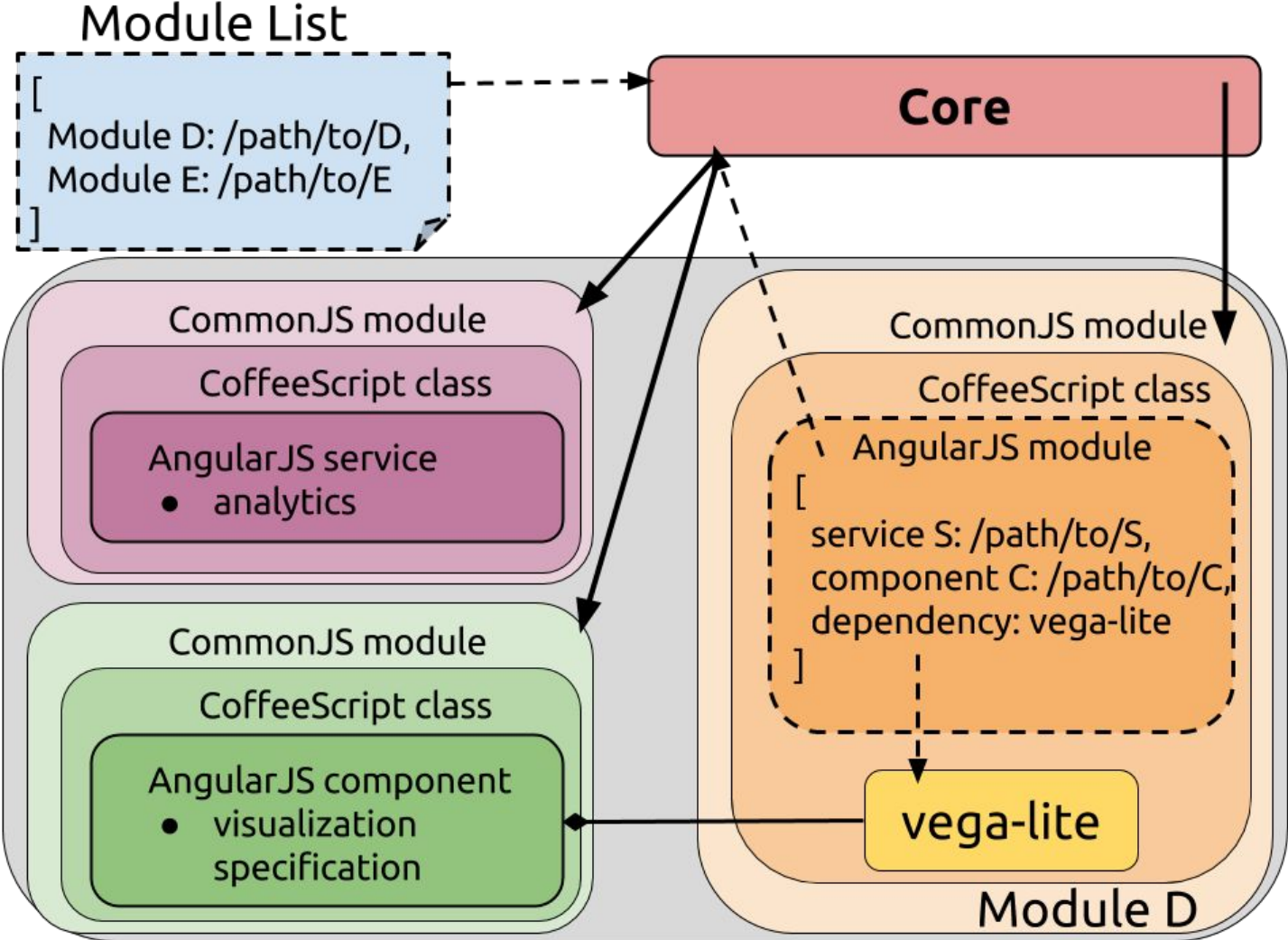
General modular platform architecture

(1) CSV to DB

(2) DB to screen



Example of module implementation



SOCR Analytical Toolbox

Load data

CVS/JSON
Drag-n-drop
APIs

Transform

Integrated
Data Wrangler

Visualize

10 different
charts

Analyse

Interactive
K-Means
Clustering

The screenshot shows the SOCR Analytical Toolbox interface. At the top, there is a navigation bar with 'SOCRAT', 'SOCR', and 'Contact' links. Below this is a main menu with 'Analysis', 'Raw Data', 'Data Wrangler', 'Charts', and 'Tools' (with a dropdown arrow). The 'Analysis' menu is open, showing options like 'Data inputs', 'Interactive spreadsheet', 'Load SOCR dataset' (highlighted in blue), 'Load data from World Bank', 'Use data generator', 'Parse JSON file by URL', and 'Parse SOCR Data page by URL'. The main content area is titled 'Get Data' and 'SOCR Datasets', featuring a search bar with 'Simulated SOCR Knee Pain Centroid Location Data' and a 'Load' button. Below the search bar, there is a text prompt 'Copy-paste data or drag-and-drop CSV file' and a table with data. The table has columns for 'x', 'Y', and 'View'. The footer of the interface reads 'The Statistics Online Computational Resource (SOCR), 2016'. Green arrows point from the text labels on the left to the 'Load SOCR dataset' menu item, the 'Data Wrangler' menu item, the 'Charts' menu item, and the 'Tools' menu item.

	x	Y	View
1	11	73	RF
2	20	88	RF
3	19	73	RF
4	15	65	RF
5	21	57	RF
6	26	101	RF
7	24	117	RF

Clustering example

Parameters

Algorithm
K-means

Number of clusters 4 Detect k

Distance
Euclidean

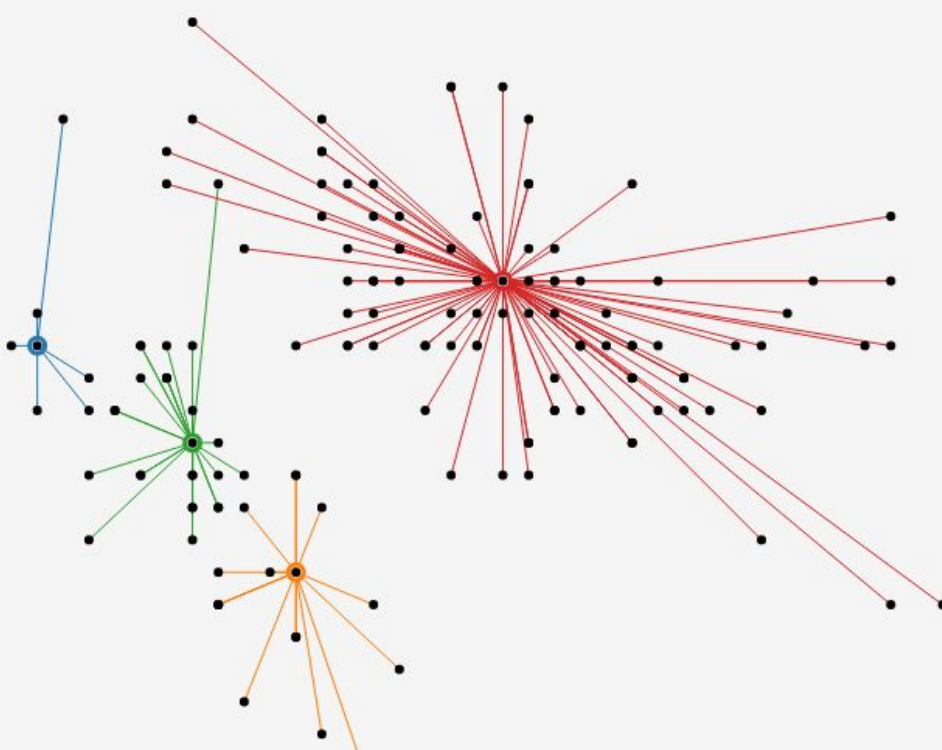
Initialization
Forgy

Dataset

Sepal_Length
Sepal_Width
Petal_Length
Petal_Width

Clustering module

k-means clustering aims to partition n observations into k clusters in which each observation cluster with the nearest mean, serving as a prototype of the cluster.



The Statistics Online Computational Resource (SOCR), 2016



Try it online

<http://socr.umich.edu/HTML5/SOCRAT>

+ Poster later today (4:50pm)

**Feedback, bugs, requests:
<http://github.com/SOCR/SOCRAT>
statistics@umich.edu**

@alxndrkalinin