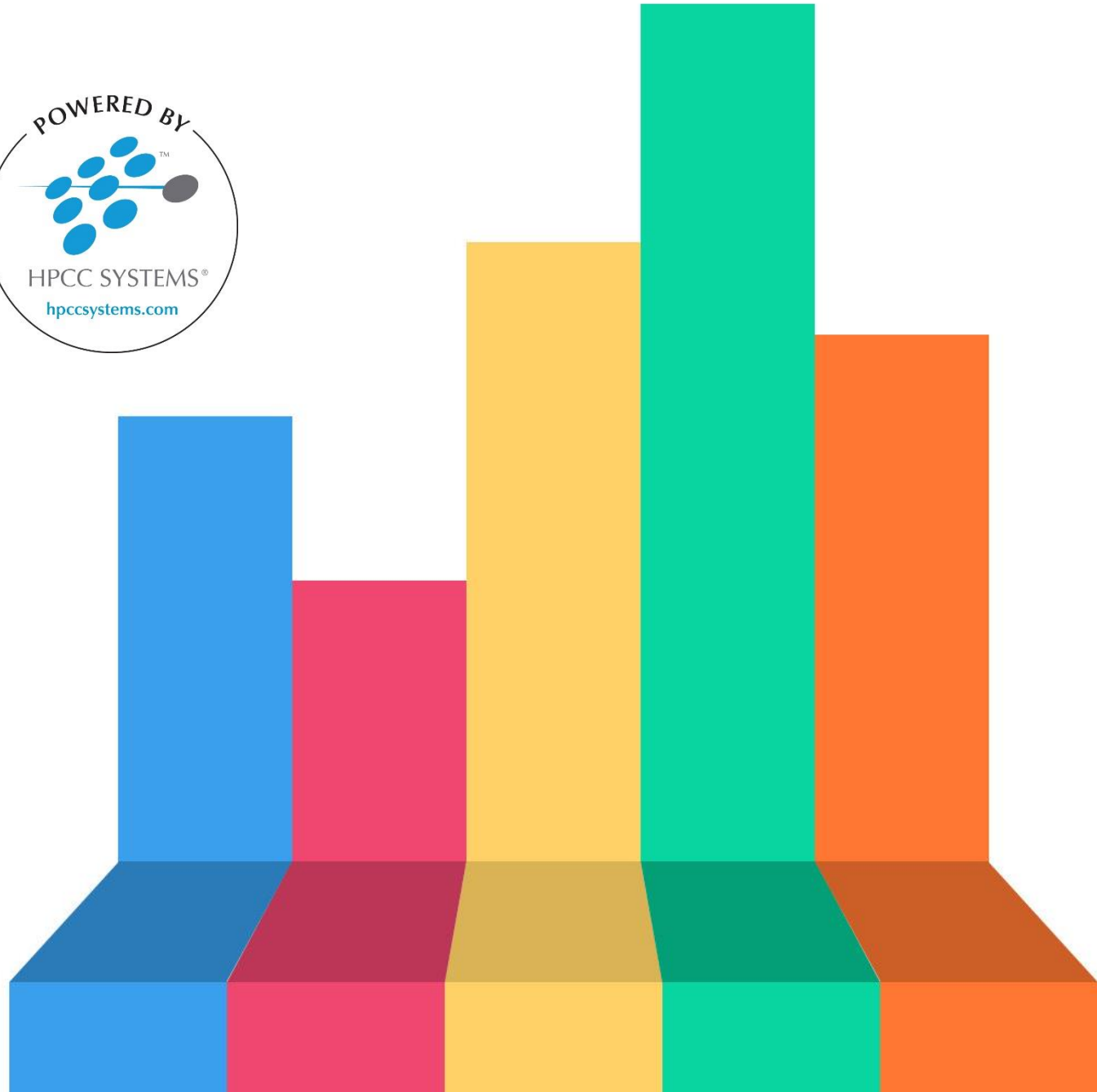


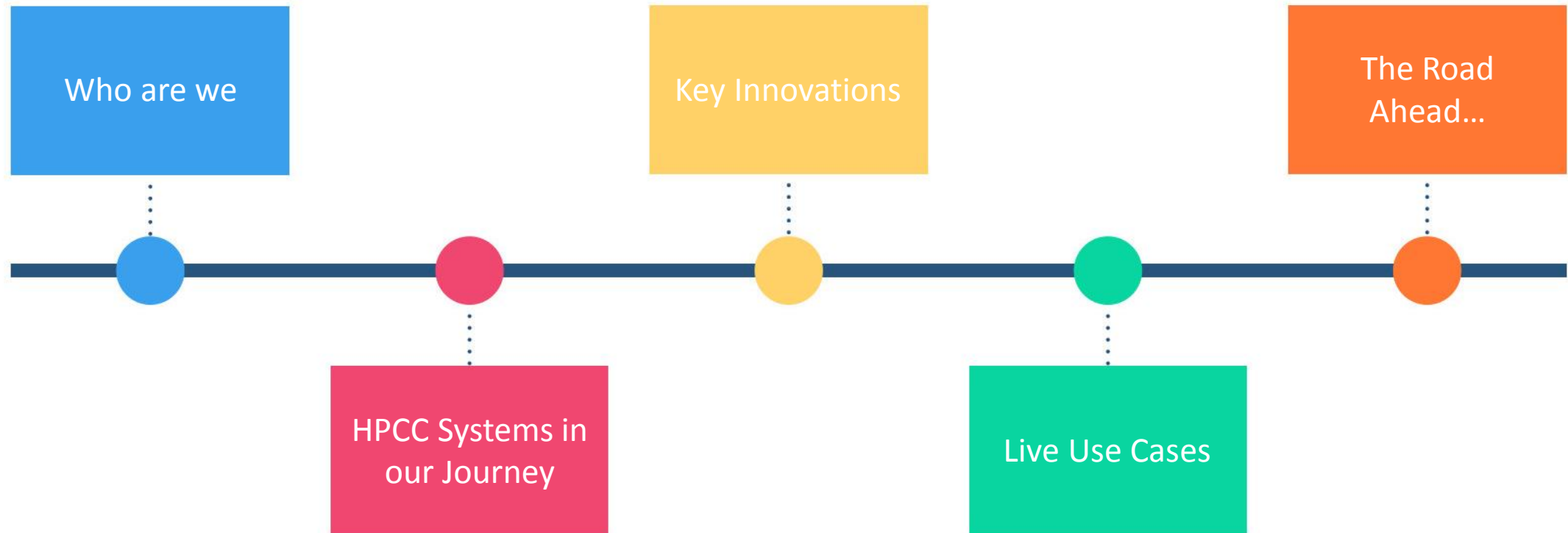
Boot-strapping a Big Data Solutions Startup: HPCC Systems on Amazon

Raj Chandrasekaran
CTO & Co-Founder

Oct 12, 2016

<https://clearfunnel.com>





About Us: Offering an innovative approach to Big Data use cases

ClearFunnel provides custom Big Data Solutions 'as-a-Service' on Subscription Fee

What we do



We provide custom solutions 'as-a-Service' for Big Data and Data Science use cases



Clients only pay low monthly subscription fees for using the solutions



We extensively use HPCC Systems (on a busy day, our clusters grow up to 3,300 slaves)



We host and operate our solutions on AWS



We thrive in solving complex Machine Learning and Big Data use cases

What we are not

Not a Systems Integration firm

Not a technology consulting firm

Not a staff augmentation or resource provider

Not a technology reseller

Unique Benefits

Cost follows Revenue

Low Subscription (Opex) and
Pay-for-Performance

No Capex in Big Data Talent,
Technology, and Infrastructure

Zero Risk, Try-before-You-Pay

Accelerated Product Launches
(~ 5 - 6 weeks)

Seamless and Turnkey Solution as-a-
Service

HPCC Systems in our Start-up journey



Fail-fast, fail-often

Re-writing our platform 'n' number of times with ECL proved to be a game-changer



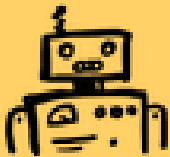
Dev-ops on auto-pilot

True one-click automation of cluster provisioning, HPCC Systems setup, ECL execution, and job monitoring



Simplified data-pipeline

Tech stack: HPCC Systems - Bash/Py - AWS micro-services



Complex Machine Learning, NLP

ECL helped us implement complex text analytics, ML, and custom NLP algorithms



Near real-time with HPCC Systems!

We deliver a 30-second index refresh frequency for a complex IOT analytics use case

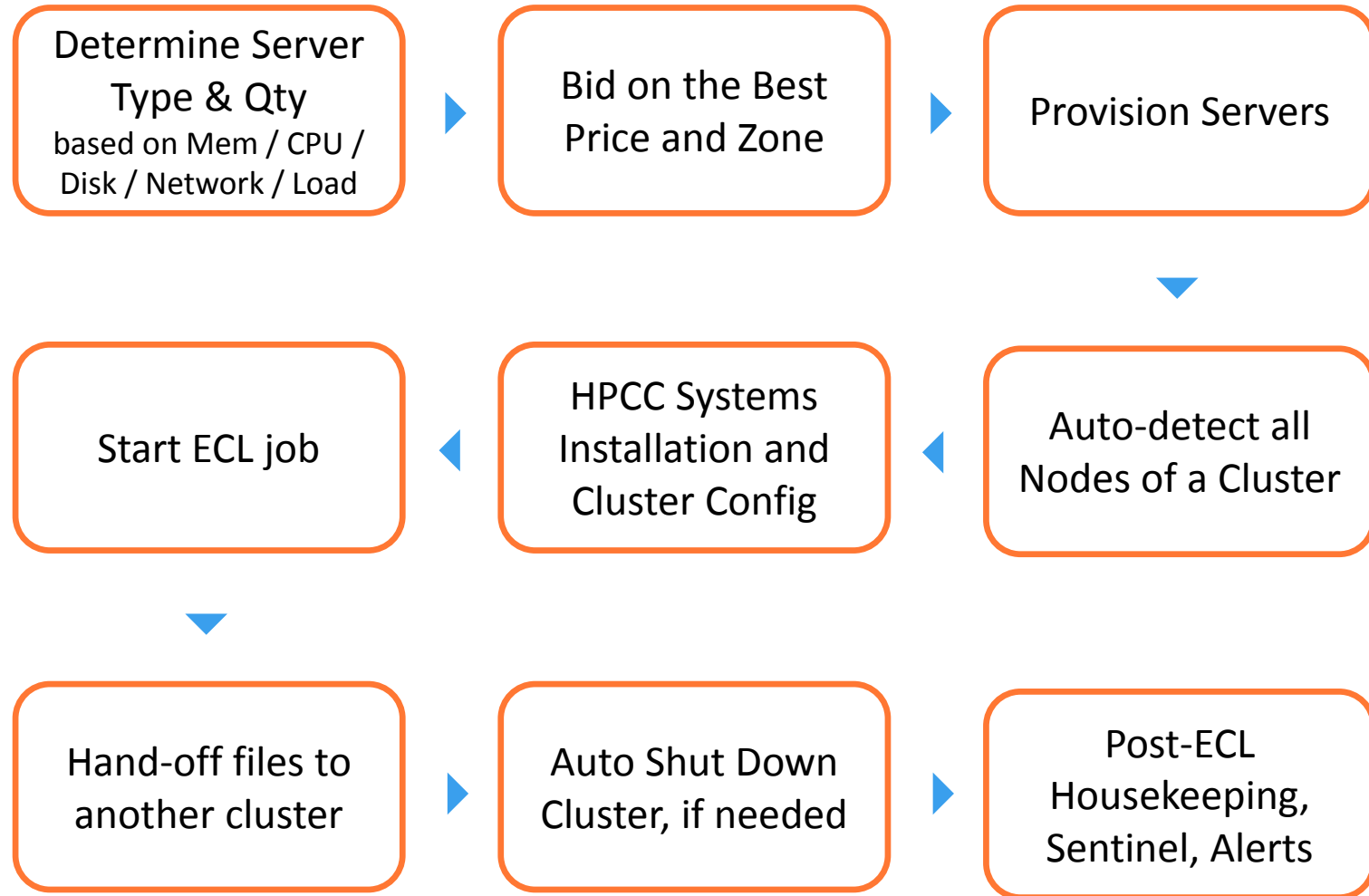


Zone of Excellence

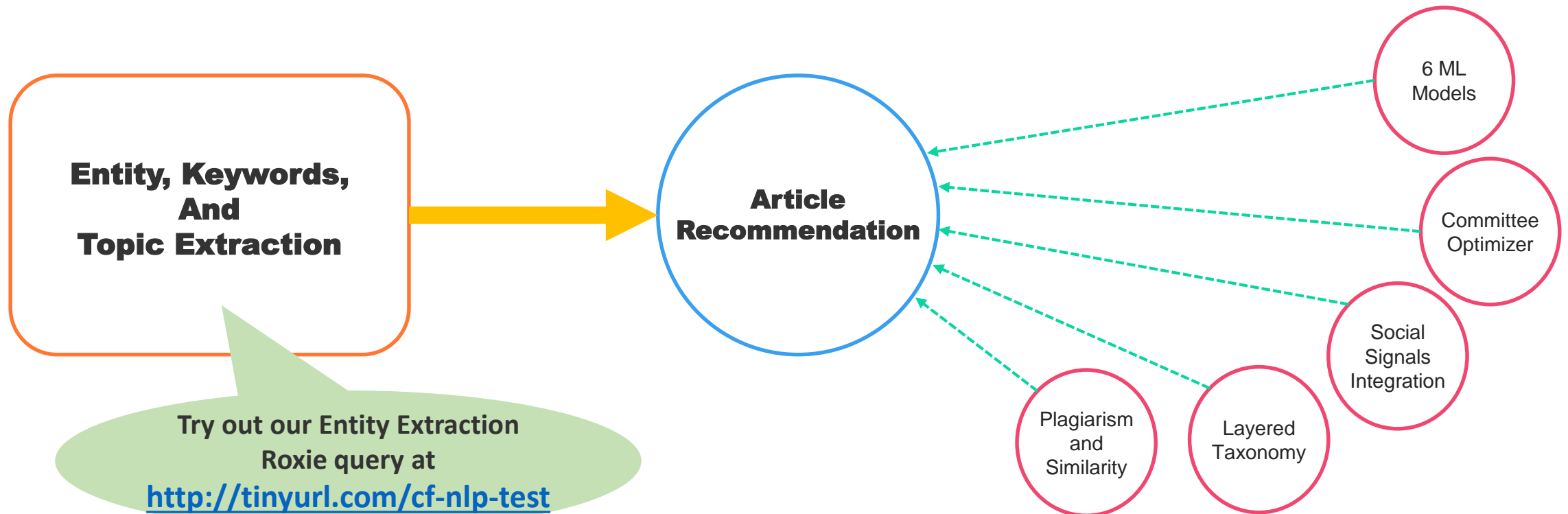
ECL re-usable libraries allow us to deliver end-to-end solutions within 5-6 weeks

Key Innovations: Cloud DevOps Automation

All steps executed with just
ONE CLICK



Key Innovations: NLP



Key Innovations: Prediction Engine

Key ask of businesses:

Timely identification of genuine prospects

Our Client's IP:

- Collect global B2B web traffic data
- Algorithms to analyze buying behavior
- Data triangulation to identify companies and contacts currently in buying cycle

HPCC Systems Implementation:

- End-to-end, automated data processing pipeline
- ETL
- Prediction, Text Analytics and Scoring Engine
- Reporting
- Data Feeds and API

Some of our other live Use Cases powered by HPCC Systems



IOT and Satellite-signal based global Maritime Domain Awareness and Supply Chain solution



Advanced **Text Analytics** and **NLP** based medical information solution



Streaming Topic analysis and signals to drive Content Marketing of Brands



Wrapper to **embed client's existing app** to provide scale-out and parallelization capabilities

The Road Ahead...

01

Consolidate existing capabilities...

- Advanced Text Analytics and NLP
- IOT and Stream processing
- **Partnering with new businesses / startups** to launch their analytics use cases faster than any alternative

02

Build expertise in...

- Image analytics at Big Data scale
- Real-time acoustic analytics
- Operational analytics (IOTs / Sensors in energy and utilities infrastructure)

03

Operate the largest deployment of HPCC Systems outside of Reed...

- End-to-end automation driven management of our growing HPCC Systems footprint (scale to thousands of servers)
- Continue innovating on and extending HPCC Systems capabilities (using in-line/embedded coding support, web services interface, command line integration, homogeneous processing and reporting stack, Cassandra option for Dali, MySQL support, etc.)

Let's Stay in Touch



raj@clearfunnel.com



@clearfunnel



<https://clearfunnel.com>

