Enda Farrell

OBJECTIVE

Business outcomes-driven software architect, data-science engineer and team lead with a repeatedly-proven ability to decompose large systems into manageable and discrete components. Requirements gathering, business design, technical design, estimation, label gathering and verification, project communication and team-leading skills have been honed over many years of delivering to both fixed-time and fixed-price engagements and being an in-house expert. Understands the operations world and the very different criteria under which they work. Understands the data science world and how to enable their work to be deployed to production and release their value to the business. An excellent communicator constantly adding business value by bridging business, data science and technology.

WORK EXPERIENCE

Senior Software Architect, Core Maps Engineering and previously Search & Places
HERE (formally Nokia Gate5 GmbH), Berlin. 2010-present

Facebook Places 2016-

The Facebook Places project is a large, strategic to the company programme of work that builds on my earlier "External Content Ingestion" and "Online Place Engagement" endeavours. My role is data modelling lead, I combine being a subject matter expert, a Data Scientist and Data Science Engineer, Systems Architect and team lead given the complexities of the programme.

The business objectives are clear: extract places from the Facebook feed which are missing from HERE's PDS system, and, give supporting evidence for other places in our datasets.

The Data Science continues to evolve. Many different models were tried: SVM (not a binary problem), logistic regression (problem isn't really linear) & eventually decision trees. There continue to be problems with over-fitting due to inaccurate labels. These labels come from in-house experts (spreadsheet), Mechanical Turk and CrowdFlower jobs: but label veracity is still a concern. I am the owner of the (still) champion model which is a high-precision, rules-oriented, multi-class classifier - its success is largely down to deep domain knowledge and intuition.

Data-science engineering has been made "production grade": the Data Science team (based in Seattle) used email to share jupyter-based notebook models. We educated them in the use of git, issue tracking systems and codes reviews. Initially their models "worked on my machine" but not on the shared infrastructure that the ops team provided us: more tutoring on coercing code to work with existing tooling and convincing ops teams to deploy anaconda distributions of python on the spark infrastructure.

Engineering and operations aspects are still being addressed: 10TB of data, now growing at 36GB per day. 200M entities, with 170M updates daily: data retention policies and storage are operational issues.

The business value being brought is significant though. In 2014 we were at 77% coverage against Google's 91% and as of Jan we're both at 93% - a huge gain and a strategically important improvement as HERE's competitive landscape continues to change.

Online Place Engagement 2015

There are considerable costs to licensing external content - which providers are most important? By combining Place Usage Analysis with External Content Ingestion & Volumetrics details, we could begin to educate the business about the importance of different providers and to democratise this knowledge (through interactive reports), allowing new cross-business unit partnerships to form. Our analysis shows that external content accounted for more than 40% of engagement: this lead to the redoubling of the Strategic Licensing Team's scope and budget. It also alerted our sales teams to a recent - but huge (>700%) year-on-year increase in traffic from south-east Asia, resulting in new revenue steams, and was a significant contribution to the setting up of the "Facebook Places" project.

External Content Ingestion & Volumetrics, 2014

The speed at which we could update the online indices had become a business issue and in 2014 I took what had become an ad-hoc, unformalised process and rebuilt it. With a dedicated team of 7 updates took just 20% of their former time; we had details about the amount of adds/changes/removals per source, and quantitative numbers about the updates. Making changes though is of no

Schönhaueser Allee 168, 10435 Berlin, Germany.
Tel: +49 151 551 553 31 enda.farrell@gmail.com

value unless we know that it improves the end-user experience, so we incorporated A/B tests into this process and got qualitative results - with their statistical significance p-values.

In parallel, we developed a framework for building detailed volumetrics about our place details: in which countries were the places, what kinds of places are they, did they have phone numbers, websites and opening-hours? The main business value of this was to have categorical records of our progress.

Place Usage Analytics, 2012 & 2013

By late 2012 the Local Search services were becoming part of a new Location-Based Services platform and new questions of the service were being asked: which customers, in which countries, what kinds of places and how often? I co-lead the complete rebuilding of our Hadoop-based log-processing and analytics platform, uniquely combining Cascalog/Clojure with our Jenkins build-automation infrastructure to answer these key business metrics on a daily basis via email. A large part of this was anomaly detection, platform misuse filtering and test/heart-beat/operational request removal. This work allowed our business leaders to understand customers' usage of the Place Services for the first time.

Universal Search, early 2012

Joined the search team for a Linux-based phone, Meltemi, and built the server-side components for a federated multi-source"Universal Search" engine. No business value was realised: Nokia abandoned the Meltemi platform, though two patents were applied for.

Places Registry, 2010 & 2011

Joined what was then Nokia's Services/Maps division to be the technical lead and architect for the central "Places Registry" (14 devs) which held the maps' POIs. Within 12 months, I helped save Nokia $10M/year by decommissioning "DataOS" systems which had marginal value, replacing them by building a Lucene-backed indexing service in the Places Registry. 6 months later, I successfully lobbied the business to increase functionality in Core Map Engineering' PDS system such that we could begin decommissioning the Places Registry in its favour, saving a further $1M/year.

Software Architect, Forge Platform Engineering
BBC, London. 2008-2010

One of the technical leads for building an entirely-new highly scalable web-facing delivery platform. Platform built for operational supportability, "shared nothing", RESTful Java JSON/XML service layer supplying data for a PHP page assembly layer. Mutual certification from our own X509 certificate CA allows the infrastructure to live on the internet - securely accessible from anywhere - enabling new hardware to be spun-up quickly when the need arise.

SVN/Hudson CI/Maven build process, made Tomcat/Java/Spring the default stack for service development.

Primary responsibility is for the Key Value Store - a multi-datacentre, 100% availability, RESTful API to securely store JSON documents in sharded CouchDB instances. It is now the de facto multi-master service for the BBC, having made it operationally stable and robust - even though at the time CouchDB was alpha code. Other significant responsibilities were to plan and track resources on assigned delivery tracks, be a technical authority to the wider platform community, to mentor and advise engineers on building systems which would fit within our scaling model and platform design.

Software Engineering Team Leader, Content Management Culture
BBC, London. 2006-2008

Joined the BBC to lead the technical team of 14 for what was at the time the 2nd biggest content management system at the BBC, behind the system that manages News and Sport. The team wasn't running smoothly, the productivity was poor and over the first year I introduced Agile development practices, automated unit, functional & load testing, and refactored the user's interface to creating content. However, this did not overcome the fact that, due to EU procurement regulations for publicly-funded organisations, the underlying EMC/Documentum enterprise CMS was not a suitable engine for "write once, rarely edit" web content. I called for the system to be decommissioned, a view which gained pan-BBC business support, approval and ratification in the pan-BBC Content Management Strategy paper chaired by Matthew Karas. This led to the decommissioning of the system, and my team of XSLT, Java and test engineers being transferred with myself to the Forge Engineering team.

Developer later Manager Technology
Sapient; Chicago, Australia, London. 1997-2006

Started as a developer in Chicago, spent two years in Australia building financial software on Oracle, often writing PL/SQL. First played senior engineer role in Mar 2000 as we designed and built a trading platform using what would now be known as AJAX. In Sep 2002 first played architect role for the UK Cabinet Office on a pan-government content management system. Over the following 3 years became on of Sapient's experts in content management - systems, technology, implementation and business change needed to make them successful.

SKILLS & QUALIFICATIONS

Presentation - ability to convey information in an engaging way and informative. Comfortable presenting to C-level executives.

Cross-business prioritisation and alignment building, building technical requirements from business needs and allowing business to realise some of the untapped potential of technical assets.

Team member mentoring and growth.

Java - web applications, Java concurrency daemons.

Cascalog/Clojure for MapReduce in Hadoop

Python - multi-threading, scripts and utilities. pyspark, numpy, pandas, maplotlib, sklearn.

Javascript & D3 for detailed, interactive presentations.

Oracle - DDL, PL/SQL, replication, PostgresSQL windowing functions.

CouchDB - lead for what was the world's largest production installation.

EDUCATION

Bachelors of Engineering (Mechanical)
University College Dublin, Dublin, Ireland. 1993-1997

REFERENCES

Available upon request.