Novo Nordisk’s Journey to a “Data Ecosystem”

Kate Mulroney
Novo Nordisk Inc.
9 October 2017
Caveats

- Reflects Novo Nordisk’s experience only
- Any opinions reflected are my own
- Our journey is still in progress...
Novo Nordisk at a Glance

A global healthcare company with more than 90 years of innovation and leadership in diabetes care.

We also help patients with haemophilia, growth disorders and obesity.
Based on turnover, Novo Nordisk is the 15th largest pharmaceutical company in the world

Source: Novo Nordisk internal benchmark report based on individual annual reports published by each pharmaceutical company. Reported sales figures are converted into DKK. Figures represent net turnover for the first nine months of 2016.
We have a global presence

I work here (most of the time)

Headquarters (and many stakeholders) is here
Our Analytics Journey
Stages of Novo Nordisk’s Analytics Development

- **Pre-history (<2006)**
  - Limited internal BI capability (SAP BW?)
  - Complete dependence on vendors

- **Fragmented BI 2006-2012**
  - Development of US Data Mart & MDM
  - US ad hoc analytics: OBIEE, SAS, QlikView

- **Global BI 2012-2014**
  - Global Business Intelligence Strategy
  - Global growth in ad hoc capability

- **“Big Data” 2014-2016**
  - US Data Warehouse appliance

- **Analytics 2016→**
  - Awakening! & “The Modern Era”
Data Mart Growth in the “Early Years”
“The Modern Era”

**Big Data Strategy developed**
- Partnership with R&D/Epi initiated around RWD
- Big Data & Analytics Centre of Excellence established & initiated pilot program

2014

2015
- Two pilots completed with R&D (Epi)
- R&D endorses continuing activities based on results
- Activities initiated with other functions

2016
- Analytics Strategy developed
- Epidemiology staff expansion – Data Science roles created

2017
- Partnership with Accenture & AWS to create OASIS
Conclusions of Big Data Maturity Assessment

- Novo Nordisk was “immature” with respect to Big Data technologies and analytics
  - “Pre-Adoption” in all dimensions in TDWI model
  - Scientific Computing is a possible exception
- Much data exists in silos throughout the company, with little ability to share
- Increasing volumes of data make it harder to process, access and analyze
- Users are expressing interest in predictive (and prescriptive) analytics, but do not know where to begin

* According to Deloitte, Novo Nordisk is on par with most Pharma companies in most areas except Infrastructure, where we are below average.
Big Data Strategy: Guiding Principles

1. Big Data initiatives should address specific business needs: “start with the business question”
2. Encourage (sponsor) experimentation to demonstrate value and gain learnings
3. Leverage “pockets of experience” within and across Novo Nordisk
4. “Agility” mindset is key, especially in early phases; do not get locked too early into standards
5. Be “open” to Open Source and Cloud options, and to challenging our “fast follower” mindset, but...
6. Initiatives must align to Novo Nordisk governance & security guidelines
7. Invest in developing competencies required by Big Data (inside and outside IT)
8. Emphasize sharing of data and knowledge across the organization (break down silos)
Big Data Strategy: Vision & Goals as of 2014

- Within ~2 years, Novo Nordisk’s maturity regarding Big Data will be at the mature end of the TDWI “Early Adopter” stage:
  - **Executive sponsorship** for Big Data technologies and tools will be established.
  - Several **successful Pilot applications** will have been run, with one or two of them ready to be expanded or productionized.
  - A core team of employees both from IT and LoB will be **trained in Big Data** tools and techniques, and knowledge is being shared across functions & geographies.
  - The **business value** of advanced analytics & big data will be understood by both IT and LoB Partners.
  - Some **technical options** will have been proven viable and candidates for standards.
  - Advanced Analytics tools will be incorporated into revised **BI Strategy**.
  - **Data Governance** processes are robust and global in reach.

We recommend that the Big Data Strategy should be refreshed in ~2 years, to refine plan based on knowledge gained and create a more concrete plan defining standards, architecture (global or local) and governance.
“The Modern Era”

- **2014**: Big Data Strategy developed
- **2015**: Partnership with R&D/Epi initiated around RWD, Two pilots completed with R&D (Epi), R&D endorses continuing activities based on results, Activities initiated with other functions
- **2016**: Big Data & Analytics Centre of Excellence established & initiated pilot program, Partnership with R&D (Epi) endorses continuing activities based on results
- **2017**: Analytics Strategy developed, Epidemiology staff expansion – Data Science roles created, Partnership with Accenture & AWS to create OASIS

Epidemiology staff expansion – Data Science roles created.
Global IT Analytics Strategy

Endorsed 2016

**Vision** To be a trusted and agile partner for analytics that drive business value

**Mission** “Global IT will foster a data-driven culture globally by providing capabilities to access, manage and analyze internal and external data assets”

**ANALYTICS** refers to the process of COLLECTING, ORGANIZING and ANALYZING data to discover patterns and other useful information.

“Business insights require data – lots of data – in real time”

Global IT Strategy 2020
Shifting from “What happened and why?” to “What do I need to know?”
Opportunities for Novo Nordisk create a compelling case for investment in data management & analytics

**Operational Excellence**
Optimize operations, and improve quality and decision-making across the value chain, from R&D to product supply to sales & marketing

**Product & Service Innovation**
Identify and create new product and service offerings, from enhancing current offerings with data and insight, to developing new business models and capabilities

**Analytics opportunities for Novo Nordisk**
Integrate data and analytics with disparate technologies into healthcare to prevent, diagnose, treat and manage diseases and to support wellness

**Risk Management**
Leverage structured and unstructured data to proactively manage and respond to possible compliance and regulatory issues

**Customer Engagement**
Differentiate products by supplementing with services that improve outcomes and/or experience; prove value vs cost of treatments to prescribers, HCPs and payers

**Digital Health**
Integrate data and analytics with disparate technologies into healthcare to prevent, diagnose, treat and manage diseases and to support wellness
The Five Guiding Principles for Novo Nordisk’s Global IT Analytics Strategy

1. Be an INNOVATIVE, EFFICIENT and VALUE-BASED business partner by blending traditional in-house/on-premises IT capabilities with external partnerships.

2. Offer a BREADTH of CAPABILITIES and SERVICES which enable users to collaborate, discover and analyze data throughout the organization.

3. Balance AGILITY with STRUCTURE in IT service-delivery models to promote active collaboration, and to meet growing and changing needs of the organization.

4. Provide FLEXIBLE and ACCESSIBLE mechanisms to capture, manage and consume diverse data from structured/unstructured & internal/external sources.

5. Ensure data QUALITY, PRIVACY and SECURITY by establishing policies in line with corporate and industry standards.
### Four Engagement Models
to serve varying needs of the business

<table>
<thead>
<tr>
<th>Illustrative Scenario</th>
<th>Run operational reports from production applications such as SAP or Model N to see latest trends and metrics</th>
<th>Run ad hoc queries, reports and dashboards from integrated repositories such as SELAS or Greenplum</th>
<th>Build predictive or prescriptive models in SAS or R, leveraging statistical techniques, etc.</th>
<th>Offer platform to load data – irrespective of size, source and type – for use in one of other models (↩)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Offering / Business Need</td>
<td>Standardized, trusted sources of operational reporting, often directly from underlying production systems</td>
<td>Dashboards, scorecards and ad hoc queries and reports on a variety of integrated and curated data assets</td>
<td>Capabilities and tools to enable innovative application of data and analytical methods in support of key business questions &amp; opportunities</td>
<td>Agile and flexible mechanism(s) for loading, storing, evaluating, and accessing data sets of all sizes and types</td>
</tr>
<tr>
<td>Governance</td>
<td>Highest level of governance for accuracy in data</td>
<td>High level of governance to ensure consistency and quality</td>
<td>Medium level, data available in raw form with self-service enrichment, but usage rights governed</td>
<td>Medium to low level, data available in raw form but access rights and related policies governed</td>
</tr>
<tr>
<td>Trust vs. Agility</td>
<td>Highly reliable and trusted for operational reporting</td>
<td>High trust based on business rules and load timing schedules</td>
<td>Agile – requires users to have good understanding of data with limited curation</td>
<td>Highly agile, users have full ownership of data within defined limits</td>
</tr>
<tr>
<td>Time to Deliver</td>
<td>Weeks to Months</td>
<td>Weeks to Months</td>
<td>Days to weeks</td>
<td>Hours to Days</td>
</tr>
<tr>
<td>Data Access</td>
<td>Highly restricted access based on user role</td>
<td>Selective – query and dashboard access based on functional area and business rules</td>
<td>Broadest access to correlate disparate data sets, perform studies, etc.; limited restriction applied</td>
<td>Defined by data ownership or other privileges</td>
</tr>
</tbody>
</table>
OASIS for Data Management & Analytics
An “Open Analytics & InsightS” Platform for Novo Nordisk

- Cloud-based infrastructure & software (AWS)
- Managed service (Accenture)
- Data management & analytics capabilities
- Support & enablement of analytics initiatives
OASIS Architecture: Building to Demand

Current Data Sources
- Real World Data:
  - CPRD
  - GE Centricity
  - Truven
  - Pharmetrics Plus
  - IMS LaaD
- Commercial Data:
  - IMS Subnational
  - IMS National Audits
  - Field Activity
  - Other CDW sources
- Product Supply Data:
  - POInT

Future Data Sources
- Pilots/In discussion
  - Bioinformatics/Lab Informatics
  - PADAWA, other I&F data
  - Multichannel Marketing data (NNI)
  - Other TBD

Current Data Sources (Informatica, SQL)

Kinesis Streams, Informatica, SQL

AWS S3 (+ Yarkon browser)

AWS RedShift

Alteryx

QlikSense, OBIEE, other

Established

In progress

R, Spark, SAS, SQL (& CodeCommit)

3 products being evaluated

Data Sources: Current

Data Sources: Future

Data Sources: Current & Future

Data Sources: Current, Future & Beyond

Data Sources: Current, Future & Beyond (including POInT and other TBD)
Current (conceptual) View of OASIS “data lake”

Cross-functional offering but primary use by Epidemiology for RWD
What’s next?

Establish the Global Enterprise Data Hub

- Enhance OASIS contract with Accenture
- US Commercial Data Warehouse migration to OASIS with completion planned 1Q18
- Preparing Manufacturing streaming data for production

In progress

- Re-evaluate technical architecture – are we still on the right path?
- Grow capabilities as projects demand and generalized to future use cases

Develop the Analytics Service Catalogue

- Tool evaluation and selection for:
  - Self-service data prep
  - **Data cataloguing capability**
  - Code repository
- Developing roadmap for real-world data factory (for analytic datasets)

Coming soon?

- Text mining
- Artificial intelligence (beyond basic ML) experimentation/PoC
- Even larger datasets (bioinformatics & proteomic) and sharing with research partners

Mature the Global Analytics Organisation

- Engage in analytics projects across the organization
- Enhance security for PHI and other critical data
- **Formalize data lake processes & operating model (ingestion, governance, etc.)**
- Grow data engineering & science capabilities to support business-initiated analytics projects
Data Engineering and Data Science is a natural extension of IT capabilities

A **Data Scientist** combines **programming** and **mathematical** skills to mine data, visually represent data, and develop advanced algorithms, leveraging statistical & machine learning techniques.
Lessons Learned

- An appetite for experimentation and willingness to evolve is essential – “AGILITY”
  - Technologies are still evolving
  - Trial & error: How big? What works best? ...
  - Small successes can be big wins
  - The cloud is your friend
- Collaboration and partnership is critical
  - Both internally and externally
  - Build strategic skills across the board!
- IT can and should play a leadership role
  - Platform selection
  - Data engineering (and data science?)
  - User training and support to develop Citizen Data Scientists