IBM Watson for Medical Coding

Martina Viell
Head Medical Coding, Bayer AG

April, 2018

Agenda

- **Medical Coding Services at Bayer**
  - **Organization**
  - **Metrics**
- Proof of Concept for IBM Watson
  - Objectives
- Watson Implementation

Medical Coding Services (1)

- Process and standards management for medical coding to ensure high quality medical coding standards and best medical coding practices
  - Decisions about coding strategies, philosophies, processes and tools in cooperation with relevant functions
  - Development and maintenance of global coding guidelines for e.g. MedDRA, WHO-DD
  - Global MedDRA and WHO-DD synonym list maintenance
  - Provision of efficient auto-encoding algorithms

Medical Coding Services (2)

- Performance of all medical coding tasks (Drug Safety and Clinical studies)
  - Tracking and coding of all internal clinical studies (phase 1 to 4), NIS and all outsourced studies
  - Medical coding of relevant data for cases processed by Global Pharmacovigilance
- Legacy data re-coding
- Version updates of both coding thesauri (MedDRA, WHO-DD) in accordance with regulatory requirements
- Organization of recoding of all clinical and drug safety data after MedDRA version updates

Medical Coding Services (3)

- Implementation, maintenance and versioning of standards for data retrieval
  - Standardised MedDRA Queries provided by CIOMS working group
  - Bayer MedDRA Queries which are developed in cooperation with expert functions in Global Pharmacovigilance, Global Clinical Development and Global Regulatory Affairs/ Labelling
  - Standardised Drug Groupings provided by the UMC
  - Bayer Drug Groupings

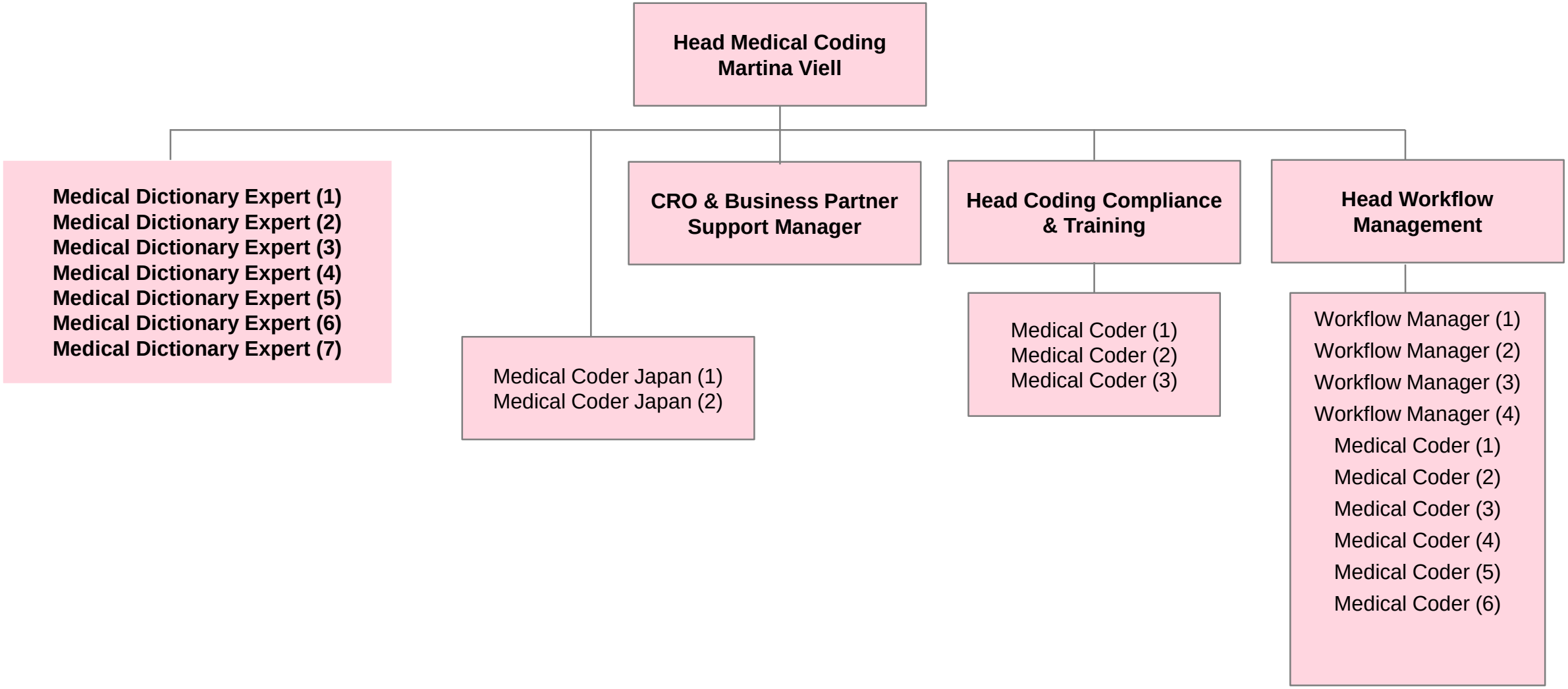
Medical Coding Services (4)

- Support of data aggregation for product safety labeling purposes (MedDRA Labeling Groupings/ MLGs)
  - GMC services in the MLG area
    - Provision of compilation rules that were reconciled with the relevant stakeholders in Labeling and Drugs Safety
    - Review and adaptation of content of existing MLGs according to the compilation rules and MedDRA content (more than 350 MLGs)
    - Provision of documentation with version history
    - Provision of new MLGs on request
    - Regular MedDRA maintenance of MLG content

Medical Coding Services (5)

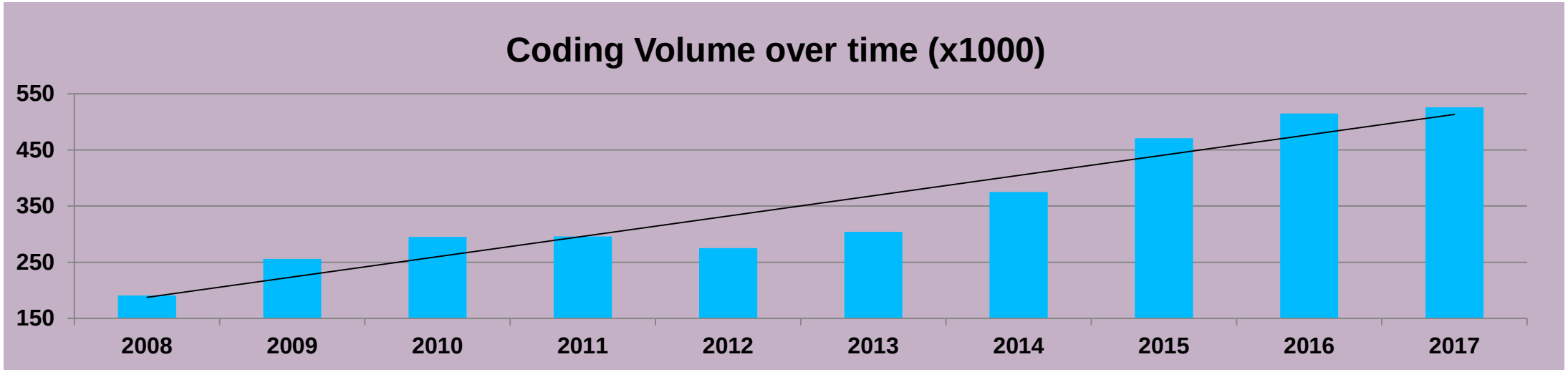
- Support and advice of relevant functions in the use of medical coding dictionaries
  - Term Specification Rules
- Company-wide provision and maintenance of role-specific MedDRA training material
  - Baylearn MedDRA training modules
  - MedDRA update training
- QA and metrics for medical coding

Medical Coding Organization



Coding Volume

- 2500 omissions (manually to be coded terms) for Pharmacovigilance on a daily basis
- Actually 140 ongoing studies
- Overall approx. 60 000 terms per month to be manually coded via ‘4-eye concept’ (proposing/accepting)

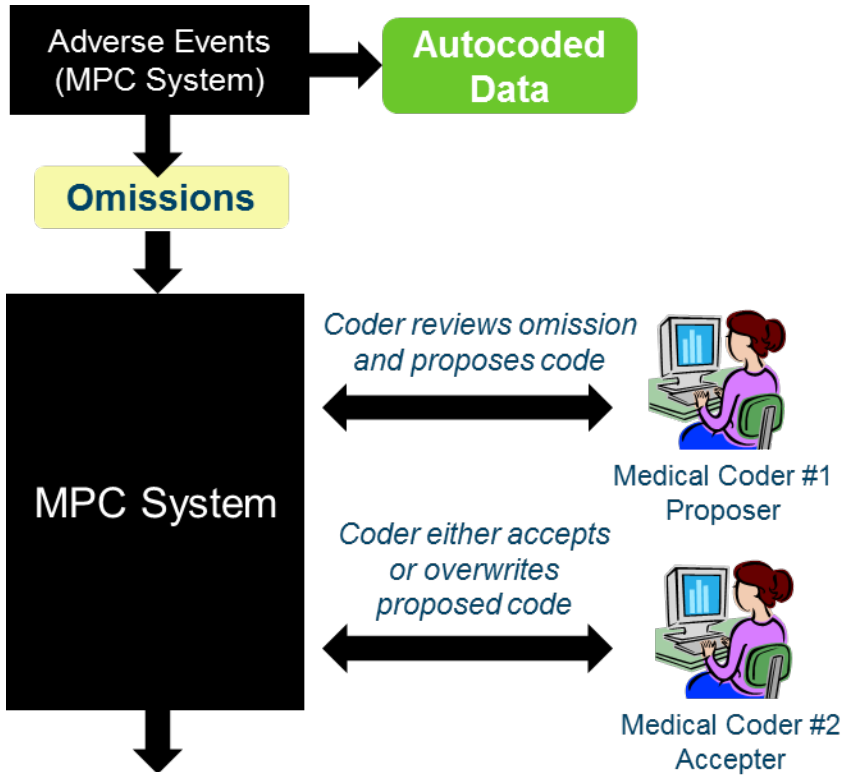


Agenda

- Medical Coding Services at Bayer
  - Organization
  - Metrics
- **Proof of Concept for IBM Watson**
  - **Objectives**
- Watson Implementation

Medical Coding Process

Before IBM Watson implementation



- Medical coding is a labor-intensive, repetitive work
- 17 FTEs
- Increasing coding volume expected caused by
  - Acquisitions
  - New therapeutic areas
  - Increase of indications
  - More clinical trials
  - Higher sales figures

  → leads to higher workload for Medical Coding team
- Skilled employees are difficult to hire
- Requires at least one year of training

MPC…MatchPoint Coder (central coding system)

Proof of Concept Objectives

Key Challenges in Bayer's Medical Coding Process for Watson to Address

Anticipated growth in volumes of omissions/day	• Increasing resource needs due to more cases predicted for 2015 in clinical studies and GPV (i.e. Merck Consumer Care acquisition) • Higher omissions rates for GPV (~80% of the overall terms to be coded for clinical and GPV) based on higher volume and lower autocoding rate
Maintain quality and consistency	• Significant training effort required for medical coding team (i.e. weekly meetings, unique training plans for each coder) • Difficulties finding and retaining qualified staff (high turnover)
Capture, internalize, and scale expertise	• Expertise and knowledge lost due to turnover • Difficult to make individual's expertise available to all

*IBM Watson can use machine learning and natural language processing to bring **consistency** to the omission coding process with a **scalable** **capacity** while **maintaining** **expertise** internal to Bayer*

Proof of Concept - Objectives

- Prove Watson's abilities to accurately propose MedDRA codes for omission data, specifically for Pharmacovigilance
  - Watson to identify and recommend appropriate synonyms or LLT terms representative of the reported data
- Watson to mimic the process of a Bayer Medical Coder today
  - Subject to the same conditions as a the human Medical Coder
  - Access to the same information (i.e. verbatim terms and alternative information, not full narratives) as the human medical coder

Proof of Concept - Process

Input:

Pharmacovigilance omissions data provided by Bayer (100K)

Output:

Proposed MedDRA code to the LLT level

in spreadsheet format

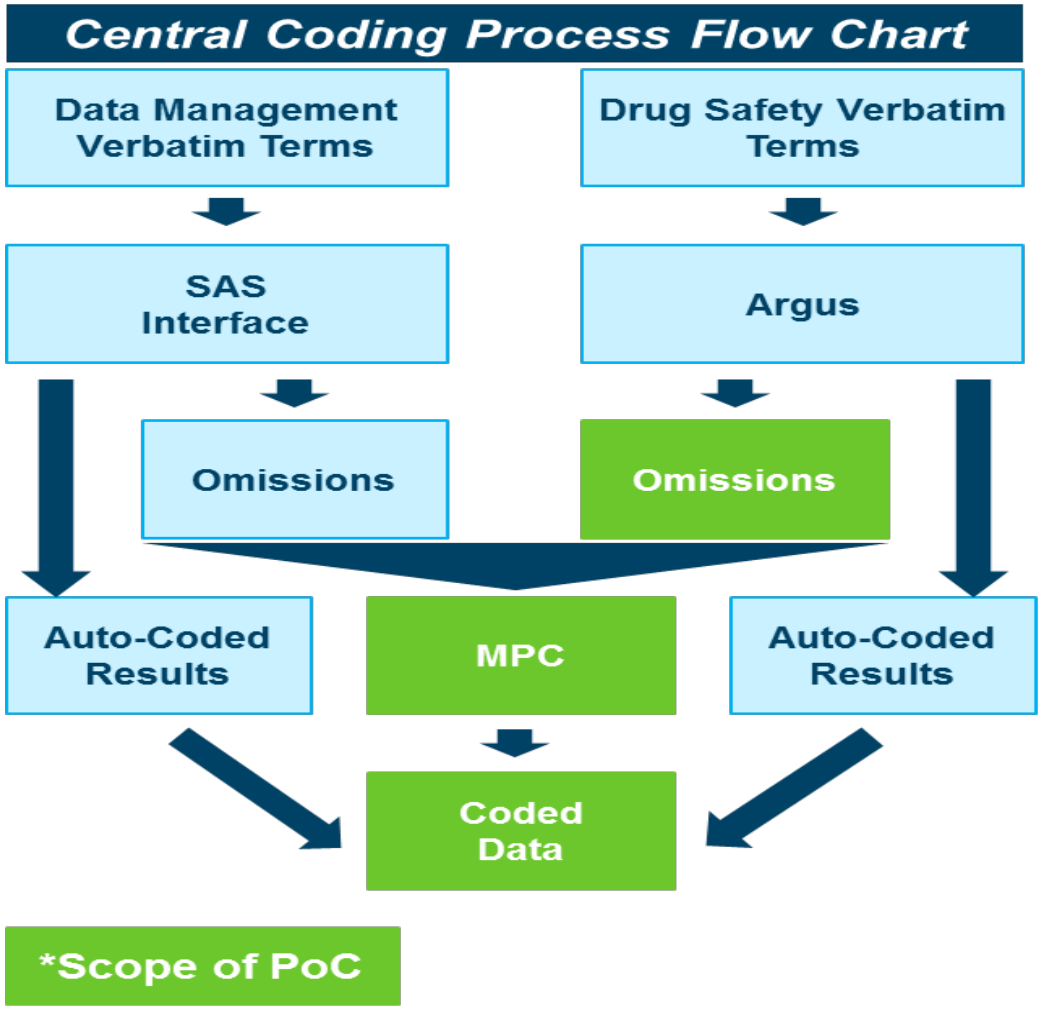
Data Sources:

Public sources:

- MedDRA (v17.1)

Bayer proprietary sources:

- Synonym List
- “Blocked List” of Terms
- Coding Convention Document



Watson's learning approach

Example #1

Omission #: 13015510

Verbatim Terms:

Gained another 10 before it was removed 6 months later

Solution Code:

Weight gain
[MedDRA Code: 10047896]

IBM Watson Proposed Term

Top 5 Recommended Codes

1) **Weight gain**
2) Drug effect delayed
3) Weight loss
4) Miscarriage
5) Cancer

Coding Decision

- Learns a pattern made up of the Stem of the word "Gained" and the numbers occurring within a window of 2 words
- Learns a correlation between this pattern and the code for "Weight Gain"

Watson's learning approach

Example #2

Omission #: 67279510

Verbatim Terms:	Solution Code:
Erroneous interruption of minisiston intake for more than 14 days	*Drug administration duration too short [MedDRA Code: 10064313]*

IBM Watson Proposed Term

Top 5 Recommended Codes

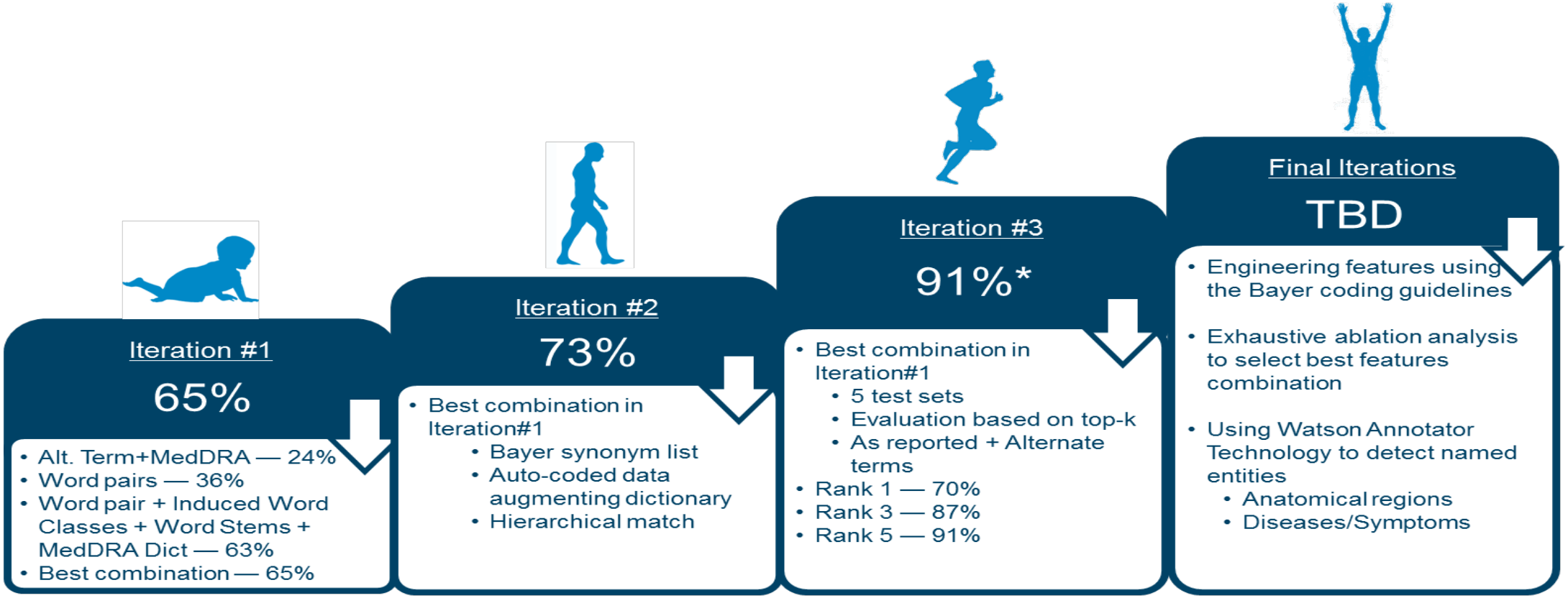
1) **Drug administration duration too short**
2) Therapy interrupted
3) Out of medication
4) Delayed period
5) Maternal exposure during pregnancy, second trimester

Coding Decision

- Learns a pattern composed of a word such as interruption or stopped etc. and a number followed by the word days
- Learns to correlate this pattern with the code for "Drug administration too short"

Watson's learning path and hit rates

Training Watson on Bayer's AE Coding process has been performed with an iterative approach, increasing its accuracy with each iteration



What could be a cost – benefit case ?

Global Medical Coding Team

Increase:

- Omissions coded per day while maintaining high success rate
- Consistency and quality of codes proposed

Reduce:

- Need to hire additional coders with the projected growth of omissions
- Cost and time needed to train new medical coders

What could be a cost – benefit case ?

Today

Internal Coders

8
FTEs

External Coders

9
FTEs

50,000 Omissions / Month
to be manually coded

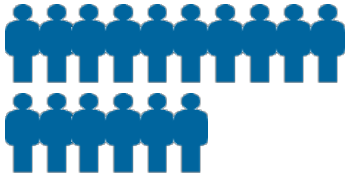
400 Omissions per Day
average per medical coder

Year 3 Projections (Without Watson)

Internal Coders*

10
FTEs

External Coders*



16
FTEs

97,656 Omissions / Month
projected to be manually coded

400 Omissions per Day average per medical coder

9 additional FTE coders needed
to handle projected increase in omission volumes

Training time and costs will increase

What could be a cost – benefit case ?

Today

Internal Coders

8 FTEs

External Coders

9 FTEs

50,000 Omissions / Month
to be manually coded

400 Omissions per Day
average per medical coder

Year 3 Projections (With Watson)

Internal Coders

2 FTEs

With Watson in place, only ~2 medical coders are required to allocate time to proposing codes while remaining staff can devote more time to other valuable activities

2 Million Omissions/Day
can be processed by Watson
assuming > 90% success rate

Watson helps drive efficiency in the medical coding process while maintaining overall quality. Medical coders can re-allocate their time to focus on other responsibilities and coding activities. The medical coding team will also save the time required to train new medical coders.

Agenda

- Medical Coding Services at Bayer
  - Organization
  - Metrics
- Proof of Concept for IBM Watson
  - Objectives
- **Watson Implementation**

Project Timelines

- Proof of concept Dec 2014 to April 2015
- Proof of concept was nominated as one of top 3 candidates of the Bayer IT Innovation Award 2015
- Start of project phase 1A in Feb 2016 (delay due to internal approval processes)
- Implementation and go-live of phase 1A in March 2017

Project Scope (1/2)

Approach

Split program into different projects: Phase 1A, 1B, 2, 3

- Phase 1A
  - Implementation of a scalable interface between MPC and Watson
    - Will cover complete program requirements
  - Enables coding of adverse events, cause of death and autopsy results from Pharmacovigilance incl. MedDRA dictionary updates
  - Implementation of mechanisms for knowledge source updates and feedback loops

Project Scope (2/2)

- Phase 1B
  - Coding of all other MedDRA-coded term types from Pharmacovigilance cases (go-live in 02/2018)

- Phase 2
  - Coding of all MedDRA term types from clinical studies

- Phase 3
  - Coding of all WHO-Drug Dictionary term types for clinical studies and PV cases

Phase 1A

Training of the Model

To train Watson we gave IBM:

- MedDRA 18.1 dictionary (as implemented by Bayer)
- Bayer MedDRA Synonyms
- Every distinct AE term in Argus autocoded under MedDRA 18.1 – approx. 200,000 terms
- Every distinct Argus AE term which created an omission in MPC
  - Approx. 200,000 terms
  - 60% with solution codes unblinded for training Watson
  - 40% with solution codes blinded for testing training outcomes

- IBM also provided other knowledge sources e.g. UMLS as a medical dictionary rather than Stedmans

Phase 1A

Training Objective and Outcomes

Training objective:

- For any batch of omissions Watson must be able to propose the correct solution at the MedDRA PT level with an accuracy no less than -5% of the accuracy achieved by human coders
  - e.g. if coders achieve an accuracy rate of 90% a rate of ≥ 85% achieved by Watson is acceptable

Training outcomes:

- Over 6 training cycles Watson achieved an accuracy of approximately 80% at the PT level when compared with the unblinded solution taken from production data
- The accuracy of human coders on the same set of omissions was estimated to be >90% (this was verified by further comparative analysis during User Acceptance Testing)

Phase 1A

Improvement Areas

// Localization of conditions

Reported term	HEART UNCOMFORTABLE, CHEST PAIN	
	LLT/Synonym	**PT**
Human	CARDIAC CHEST PAIN	Angina pectoris
Watson	Chest pain	Chest pain

// Selection of closest matching text

Reported term	ANEMIA (HAEMORRHAGIC)	
	LLT/Synonym	**PT**
Human	Hemorrhagic anemia	Haemorragic anaemia
Watson	Anemia	Anaemia

How Watson learns to code

For initial learning Watson

- Ingests a large volume of ‘facts’, primarily all terms ever seen in the coding system and the codes assigned to them
- Analyses the frequency of occurrence of words in the facts, individually and in combination
- Correlates the established frequencies with the codes assigned
- Uses this to build a knowledge base considered to be the ‘ground truth’ of coding

When presented with a term to code Watson

- Evaluates the term and its constituent words against the ground truth
- Proposes 1 – *n* codes based on statistical probability derived from the ground truth

How Watson learns to code

On-going learning is implemented to keep the ground truth correct and current

- A correct proposal is fed back to reinforce the ground truth, an incorrect proposal is fed back to adjust it
- Ground truth is also adjusted by dictionary changes e.g. stop proposing a newly-retired term or deleted synonym

Learning challenges

- Watson does not ‘understand’ the terms presented to it in the same way a human does
- To teach it how to code new and changed concepts it needs to be given lots of similar examples
- The more data it has, the higher the chance it will make correct proposals
- Coding data is characterised by very uneven frequency distributions so it may take a long time for Watson to improve accuracy of proposals for less frequently occurring terms

Forward-Looking Statements

This website/release/presentation may contain forward-looking statements based on current assumptions and forecasts made by Bayer management.

Various known and unknown risks, uncertainties and other factors could lead to material differences between the actual future results, financial situation, development or performance of the company and the estimates given here. These factors include those discussed in Bayer’s public reports which are available on the Bayer website at http://www.bayer.com/.

The company assumes no liability whatsoever to update these forward-looking statements or to conform them to future events or developments.

Thank you!

Bye-Bye