

# Data Science

## Data Science - ESA 1- Sandra Tieben

### A). Read the PWC Report. What areas are covered to have a huge impact by data science? Welche Bereiche weisen ein besonders hohes Potential auf?

Link: <https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf>

#### Geographisch:

Laut dem PWC-Bericht wird erwartet, dass durch die Einführung und Weiterentwicklung von KI das **globale BIP in 2030 um bis zu 14%** steigt - das entspricht zusätzlichen 15,7 Billionen USD. Besonders groß ist die Steigerung in **China** (geschätzte Steigerung des BIP von **26%**) und **Nordamerika (+14 %)**.

#### Felder mit großem Potential (S. 4):

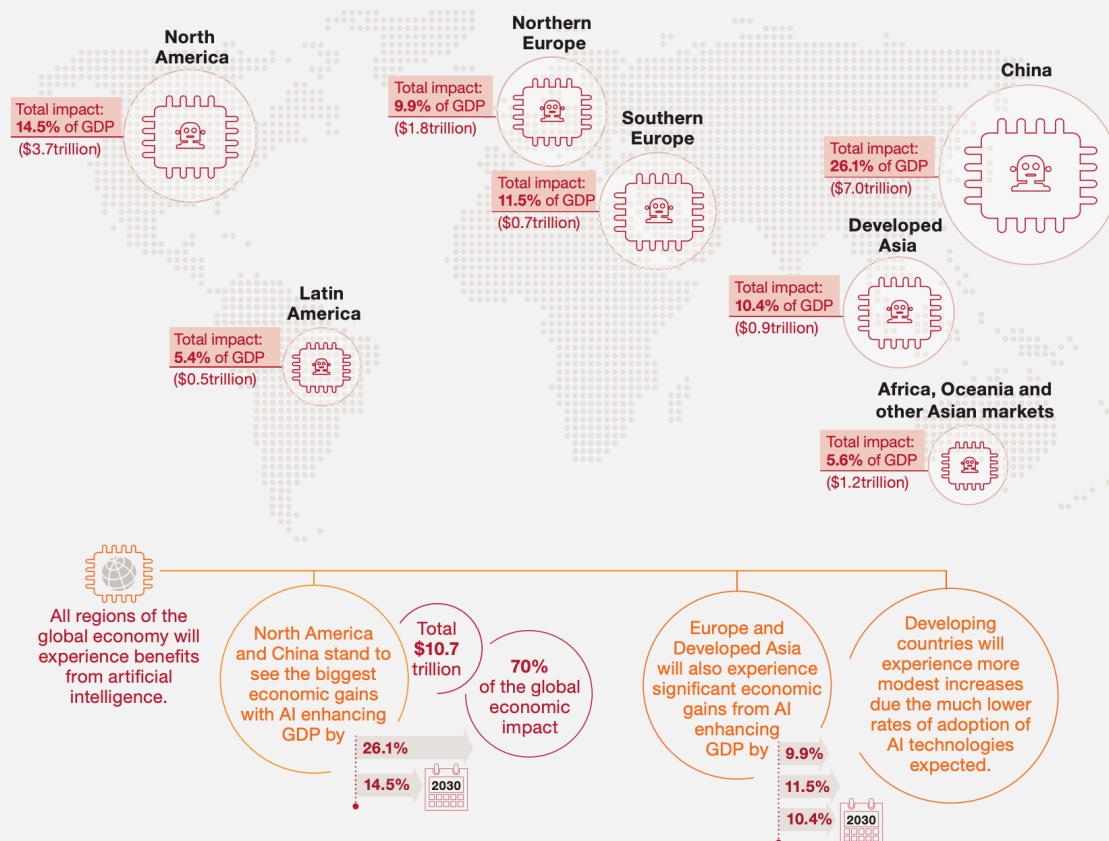
1. Automatisierung von Prozessen mittels KI: Die zunehmende Automatisierung steigert die Produktivität.
2. Durchführung von Arbeitsvorgängen mit KI-Unterstützung (Assistenz): Steigende Produktivität der Mitarbeiter.
3. Optimierung von Produkten und Erstellung von personalisierten Produkten mittels KI: Steigerung der Nachfrage.

#### Branchen mit hohem Potential:

Neben der Geschäftsfelder und Gebieten, die besonders von KI profitieren, werden auch acht Branchen mit hohem Potential aufgeführt:

1. Gesundheitswesen
2. Automobilsektor
3. Finanzsektor
4. Verkauf
5. Fertigung
6. Technologie-, Kommunikation- und Unterhaltsbereich
7. Energiesektor
8. Transportwesen/Logistik

**Figure 2: Which regions will gain the most from AI?**

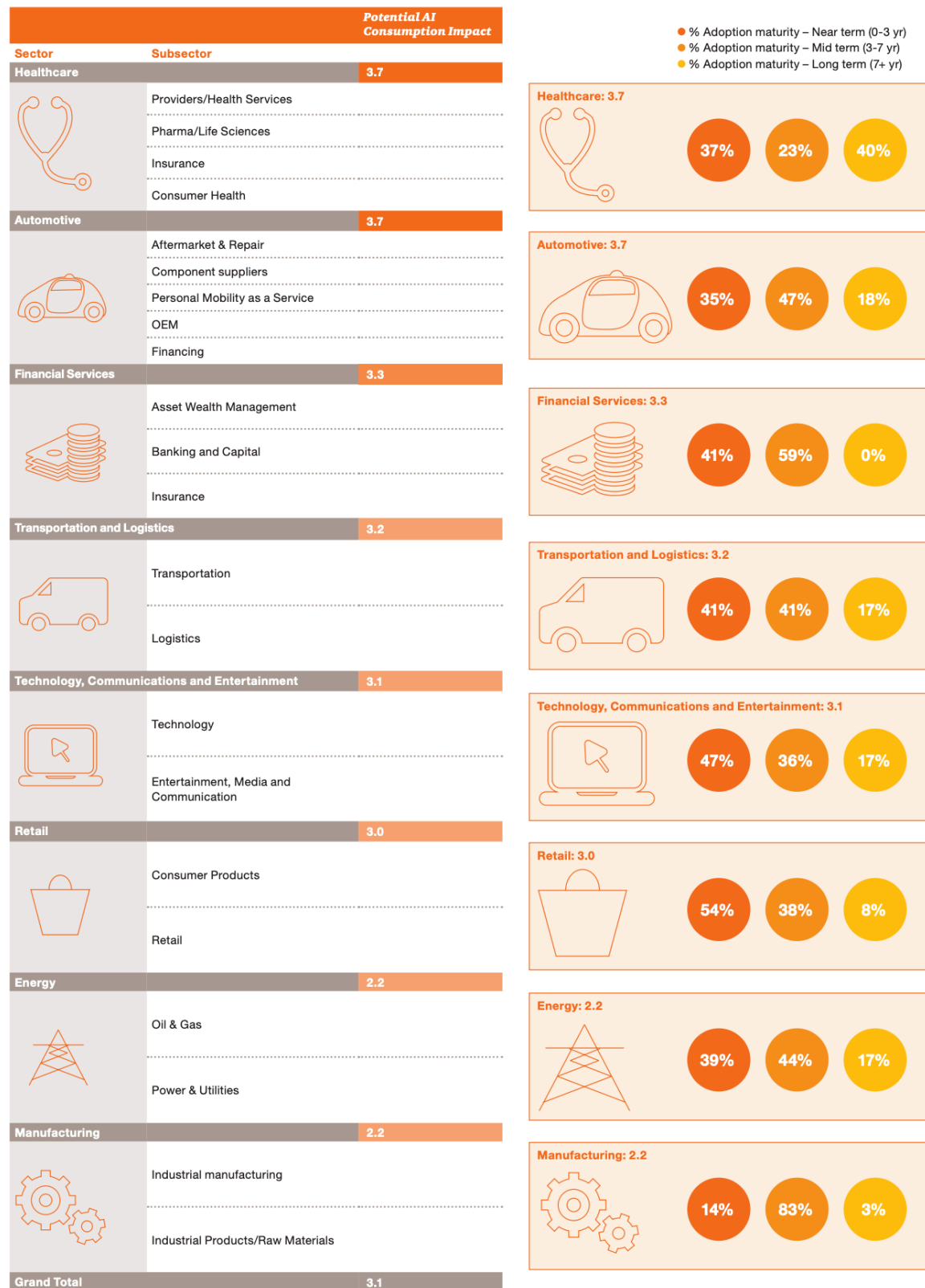


All GDP figures are reported in market exchange rate terms

All GDP figures are reported in real 2016 prices, GDP baseline based on Market Exchange Rate Basis

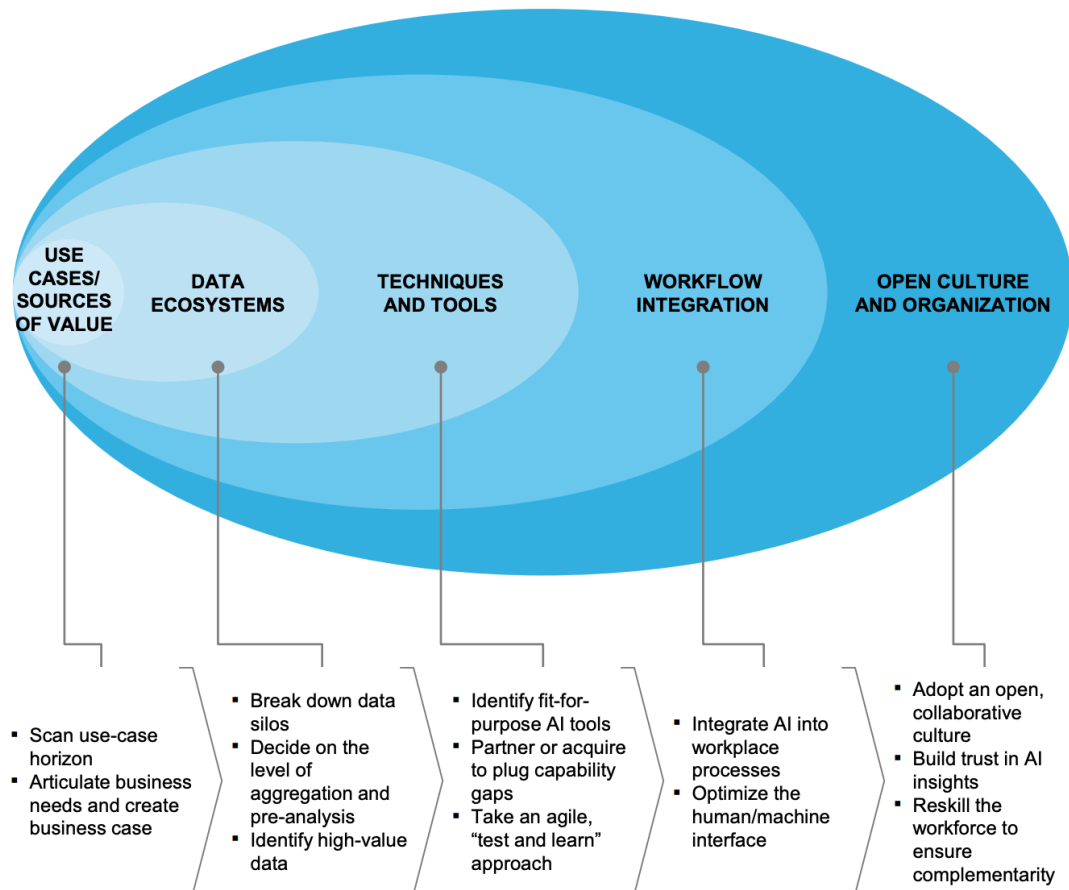
Source: PwC analysis

**Figure 4: What's the potential impact for your sector?**



## B) Read the McKinsey Report on AI. Do you agree and understand exhibit 8? Are there steps missing?

Successful AI transformations require elements similar to those found in successful digital and analytics transformations



SOURCE: *The age of analytics: Competing in a data-driven world*, McKinsey Global Institute, December 2016; McKinsey Global Institute analysis

Prinzipiell stimme ich der Aussage zu. Allerdings kommt meiner Meinung nach die Einbeziehung von Menschen, Datenschutz und Regeln/Gesetze zum ethischen Umgang mit den Daten zu kurz. Diese werden zwar zum Teil im Feld „Open Culture and Organization“ angeschnitten („Adopt an open, collaborative culture“, „Build trust in AI insights“), aber sind für die praktische Anwendung von immenser Bedeutung.

Dies gilt insbesondere bei der Nutzung von AI zur Überwachung oder im Gesundheitssystem. Dies sieht man aktuell besonders anschaulich an Diskussionen zur Handyüberwachung zur Eindämmung des Coronavirus’.

## DSI-02

**Make up your own example of a machine learning task with a binary classification with your own numbers. Like e.g.**

## **recognising skin cancer. Calculate precision and recall.**

Wettererkennung - Regen/trocken

Beispiel:

Im Februar hat es an 21 Tagen geregnet und an 7 Tagen war es trocken.

Der Test erkennt 19 der 21 Regentage. Zusätzlich werden 3 trockene Tage als Regentage erkannt.

$$\rightarrow \text{Precision} = 19/(19+3) = 0.864$$

$$\rightarrow \text{Recall} = 19/(19+2) = 0.905$$

### **EXERCISE DSI-03**

**What is the F1 measure of your example? (real calculation needed here).**

$$F1 = (2 * \text{precision} * \text{recall}) / (\text{precision} + \text{recall}) = 2 * 0.864 * 0.905 / (0.864 + 0.905) = 0.884$$