

# SPARBUCH

Vor langer Zeit gab es **exponentielles Wachstum** ... auf dem Sparbuch



Bildquelle: Konstantin Evdokimov/Unsplash



**The sexy job  
in the next ten years  
will be statisticians.**



**BETTER READY THAN JUST AWARE**

**Wie ein erfolgreicher Umgang mit  
Daten und KI gelingt**

**Katharina Schüller, CEO STAT-UP**

**[katharina.schueller@stat-up.com](mailto:katharina.schueller@stat-up.com)**



Bildquelle: Konstantin Evdokimov/Unsplash



## AGENDA

[1] Smart Data & Analytics =  
Erfolgsfaktoren

[2] Data & AI Literacy =  
Kompetenzen

[3] Nachhaltige & integrierte  
Datenstrategie



# [1] Smart Data & Analytics = Erfolgsfaktoren

- **Daten verdoppeln sich alle 2-3 Jahre**
- **Begriff „exponentielles Wachstum“ weithin bekannt und kaum verstanden**
- **„Mehr Daten“  $\neq$  „mehr Wissen“**
- **Problem: Mangel an guten Daten**
- **Daten- & KI-Experten: schlechte Daten „reparieren“, z.B. Lücken schließen**

**→ DATEN = GRUNDLAGE FÜR ENTSCHEIDUNGEN, DARIN LIEGT WERT**



Bildquelle: Clay Banks/Unsplash

## **Geschäftsmodelle auf dem Prüfstand → datengetrieben?**

**THESE: Daten = Öl des 21. Jahrhunderts**

- 1. Daten sprudeln wie Öl**
- 2. Daten & Öl in zahlreichen Formen und Anwendungsmöglichkeiten**
- 3. ABER: Nicht Rohmaterial zählt, sondern Wertpotenzial**

**Wertpotenzial ≠ Produkt (Benzin), sondern Bedürfniserfüllung (Mobilität)**

**→ DATEN SIND NICHT DAS ZIEL, GENAUSO WENIG WIE STATISTIKEN**





→ **ZIEL = WERTBEITRAG VON DATEN FÜR BESSERE ENTSCHEIDUNGEN**

- **Daten-Wertschöpfung misslingt häufig**  
**3/4** aller Analytics-Projekte scheitern  
**90%** Daten in Unternehmen liegen brach  
Kosten schlechter Datenqualität mehrere **100 Mrd USD**, allein in USA
- **SMART DATA** = qualitativ hochwertige Daten, direkt verarbeitbar zu entscheidungsrelevanter Information
- **ANALYTICS/AI** = Werkzeuge, um Information aus Daten zu gewinnen



## [2] Data & AI Literacy: Kompetenzen

## Datenkompetenz = Voraussetzung für Relevanz / Nutzen von Analytics

### 1. Nachhaltige Ausrichtung an **Kundenbedürfnissen**

Kundenverständnis, passgenaue, datenbasierte Produkte/ Services

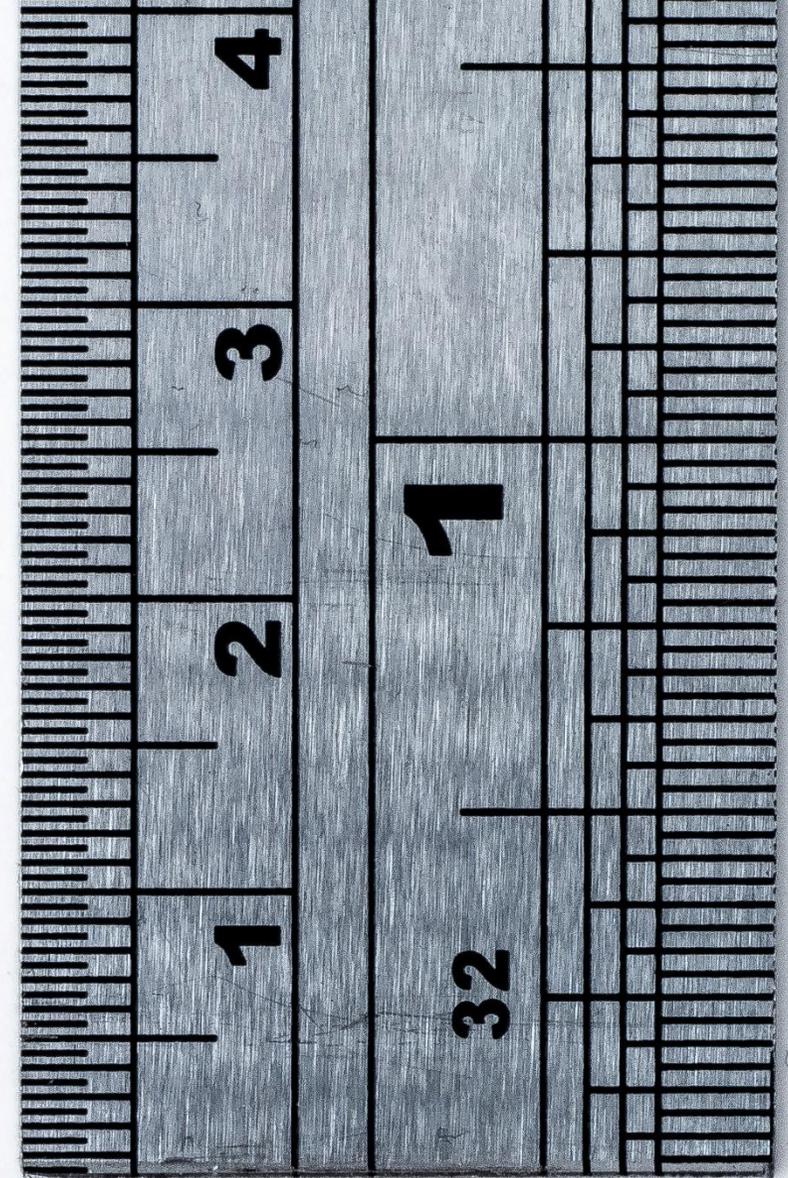
### 2. Strategische **Entscheidungen**

Neue Märkte/Kundengruppen

### 3. Verbesserung **Wirtschaftlichkeit**

Optimiertes Angebot, Risikobewertung, operative Leistung

→ **DATENKOMPETENZ = ZENTRALER BAUSTEIN EINER DATENSTRATEGIE**



**Datenkompetenz (DATA LITERACY) wird  
Schlüsselkompetenz für alle**

**→ 4 Fragen in Zusammenhang mit der  
Geschäftstätigkeit einer Bank:**

- 1. Was WILL ich mit Daten machen?**
  - 2. Was KANN ich mit Daten machen?**
  - 3. Was DARF ich mit Daten machen?**
  - 4. Was SOLL ich mit Daten machen?**
- Datenethik wird vom bloßen Schlagwort  
zu konkreter Roadmap für CDR**
  - Auch KI-Kompetenzen werden eine  
immer größere Rolle spielen**



**Unerlässlich: VERSTÄNDNIS** von

- Einfluss der Technologien auf Interaktion von Menschen
- Chancen & Grenzen von Daten/Algorithmen

**Größtes Problem:**

- Blindes **VERTRAUEN** von Menschen in Black-Box-Algorithmen
- Mangelnde menschliche Kompetenz (in doppelter Hinsicht), **Entscheidungen** von Maschinen zu **hinterfragen**

**“People try to twist and turn and use every tool to get sense out of these data sets. But we’re flying a little bit blind at the moment.”**

## **GOOGLE FLU TRENDS**

- **Grippewellen vorhersagen aus Suchanfragen nach Symptomen**
- **Ergebnis: Grippewelle 50% überschätzt**

## **CHURN PREVENTION**

- **Korr. Vertragslaufzeit – Kündigung**
- **Kündiger: kürzere Laufzeiten, weil sie gekündigt hatten**
- **Nicht-Kündiger waren zum Zeitpunkt der Analyse noch Kunden**

**→ FALSCHES SICHT AUF DIE DATEN FÜHRT ZU TRUGSCHLÜSSEN**

## **INTERNET-NUTZUNG**

- **Online-Befragung (2016): Können Sie sich Leben ohne Internet vorstellen?**
- **Indien 82%, Japan nur 62%**
- **Aber: Internetnutzung JAP 91%, IND 28%**
- **D.h. 77% der Inder können/müssen sich Leben ohne Internet vorstellen**

## **SCORING AUS BESTANDSKUNDEN**

- **Zusammenhänge mit best. Merkmalen > Rekonstruktion früherer Scorings?**
- MAN LERNT, WAS MAN SCHON WEISS, bzw. dass Kampagne funktioniert hat**

## TARGET

- **Klage: Werbung verführe 17-jährige Tochter zum Schwangerwerden**
- **Aber Tochter war schon schwanger**
- **Target: Algorithmen identifizieren Schwangere durch Kaufverhalten**
- **Aufschrei > Privatsphäre?!?**

## → LAW OF TRULY LARGE NUMBERS

- **Daten von >100 Mio Amerikanern**
- **Selektion aller Frauen im gebärfähigen Alter → 100.000+ jetzt / in 1-2 Wochen schwanger**

## FACEBOOK/TRUMP-KAMPAGNE

Charakterzug zu **88%** aus Likes  
vorhersagbar – was heißt das?

Ann: 10%, d.h. 1.000 / 10.000 Personen

880 richtig, 1.080 falsch positive

- Treffer  $880 / (880 + 1.080) = 45\%$

## PROBLEM: VIELE KOMBIS VIELER DATEN

- Microtargeting kurzfristig billig
- 100K Tests/Tag,  $\alpha=5\%$ : 5.000 falsch pos.
- → **LANGFRISTIGER SCHADEN**  
(Wirkungsverlust von Kampagnen)



**“There are a lot of small data problems that occur in big data. They don’t disappear because you’ve got lots of the stuff. They get worse.”**



## **DATENBEZOGENE HERAUSFORDERUNG**

- **Datenbeschaffung, Datenerfassung, Datencharakteristika, Datenqualität, Datenbereinigung**

## **PROZESSBEZOGENE HERAUSFORDERUNG**

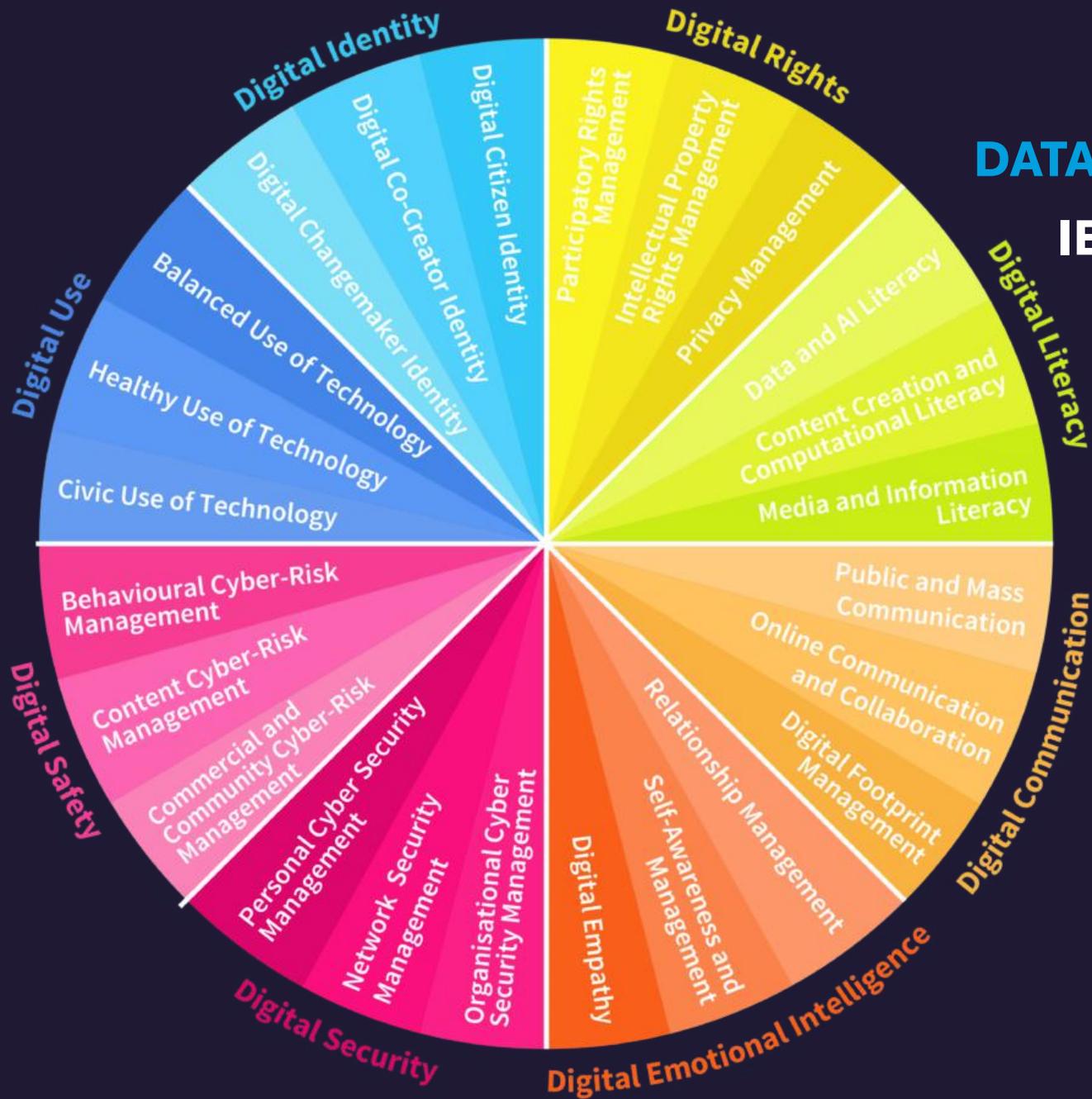
- **Speicherung, Integration, Transformation, Modellwahl, Implementierung, Interpretation, Kommunikation**

## **MANAGEMENT-HERAUSFORDERUNG**

- **Privatsphäre, Sicherheit, Steuerung, Ethische Aspekte**



**“Big data  
has arrived,  
but big insights  
have not.”**



## DATA & AI LITERACY

### IEEE 3527.1: DIGITAL INTELLIGENCE DQ

#### Ability to

- generate, process, analyze, present **MEANINGFUL INFORMATION** from data
- develop, use, and apply AI and related algorithmic tools and **STRATEGIES**
- to guide informed, optimized, and contextually relevant **DECISION-MAKING** processes

KODIEREN

DEKODIEREN



Reale Welt



BEOBACHTEN



HANDELN



Daten



Wissen

MODELLIEREN

INTERPRETIEREN



Information



Entfernen von Kontext

Kontext hinzufügen



## DER WERTSCHÖPFUNGSPROZESS

[1] Rohmaterial → **DATEN**

Edelstein + Edelmetall

[2] Aufbereiten → **INFORMATIONEN**

Schmieden/Schleifen

[3] Kombinieren → **WISSEN**

Ring

[4] Bedeutung verleihen/ interpretieren →  
**HANDELN/MACHT**

Tiffany's Stempel

→ Größter Teil der Wertschöpfung  
passiert in [4] und steckt nicht in [1]

**"Bewusstsein für Notwendigkeit zentraler Datenhaltung und Datenqualität."**

**"Use Cases: Was ist sinnvoll und plausibel?"**

**"Wir müssen Storytelling machen."**

**"Wir müssen Daten konsistent und hochqualitativ kommunizieren."**

**"Prozessdaten von Personenbezug trennen, damit Daten nutzbar bleiben."**

**"Wir brauchen hohe Transparenz bei den Daten."**



Reale Welt

**"Wie schaffen wir Verständnis, was wir mit den Daten machen?"**

**"Nicht jedem, der Daten bereitstellt, sind Nutzung und Mehrwert klar."**

**"Welche Daten sind valide, welche nicht? Besser eine Lücke als falsche Daten!"**



Daten

Wissen



**"Wir versuchen Modelle mit genaueren Daten besser zu machen, aber sind uns der Modellrisiken nicht bewusst."**

MODELLIEREN

INTERPRETIEREN



Information

**"Als Führungskraft muss man eine Chance haben, das Ergebnis zu plausibilisieren."**

**"Plausibilisierung fällt vielen schwer; Mitarbeiter können nicht einschätzen, was best/worst/real case ist."**

**"Es braucht mehr Kompetenz, um einzuschätzen, was relevant ist."**

**"Prognosen machen wir oft in Excel, sehr individuell und nicht standardisiert."**

**"Wie messe ich Prognosefehler, welche Stellgrößen habe ich?"**



## [3] Nachhaltige & integrierte Datenstrategie

## **„Gläsernes Unternehmen“ = Illusion**

- **In die Daten „hineinschauen“ heißt nicht, das Richtige zu sehen und die richtigen Schlüsse zu ziehen**
- ES GEHT NICHT UM DATEN, SONDERN UM MUSTER IN DEN DATEN**
- **Gezielter Einsatz von Data Analytics & selbstlernender Algorithmen (Machine Learning)**  
**Z.B. Anwendungen entwickeln, die automatisierte, datengetriebene Risikoanalysen durchführen und Warnmeldungen aussteuern**



## Ziele der integrierten Datenstrategie:

1. **Datenkultur** zur Entwicklung & Implementierung von Anwendungsfällen schaffen
2. **Kompetenzen** dauerhaft auf konkurrenzfähige Produkte und Services ausrichten

Fähigkeit der Organisation, Daten systematisch in Wert zu verwandeln.

Betrifft Analyseprozess (Coding) & Ergebnis-Interpretation (De-Coding)

→ Interpretation ist nicht Teil der Daten –  
**DATEN UND BEDEUTUNG VON DATEN  
SIND WIE RING UND STEMPEL**

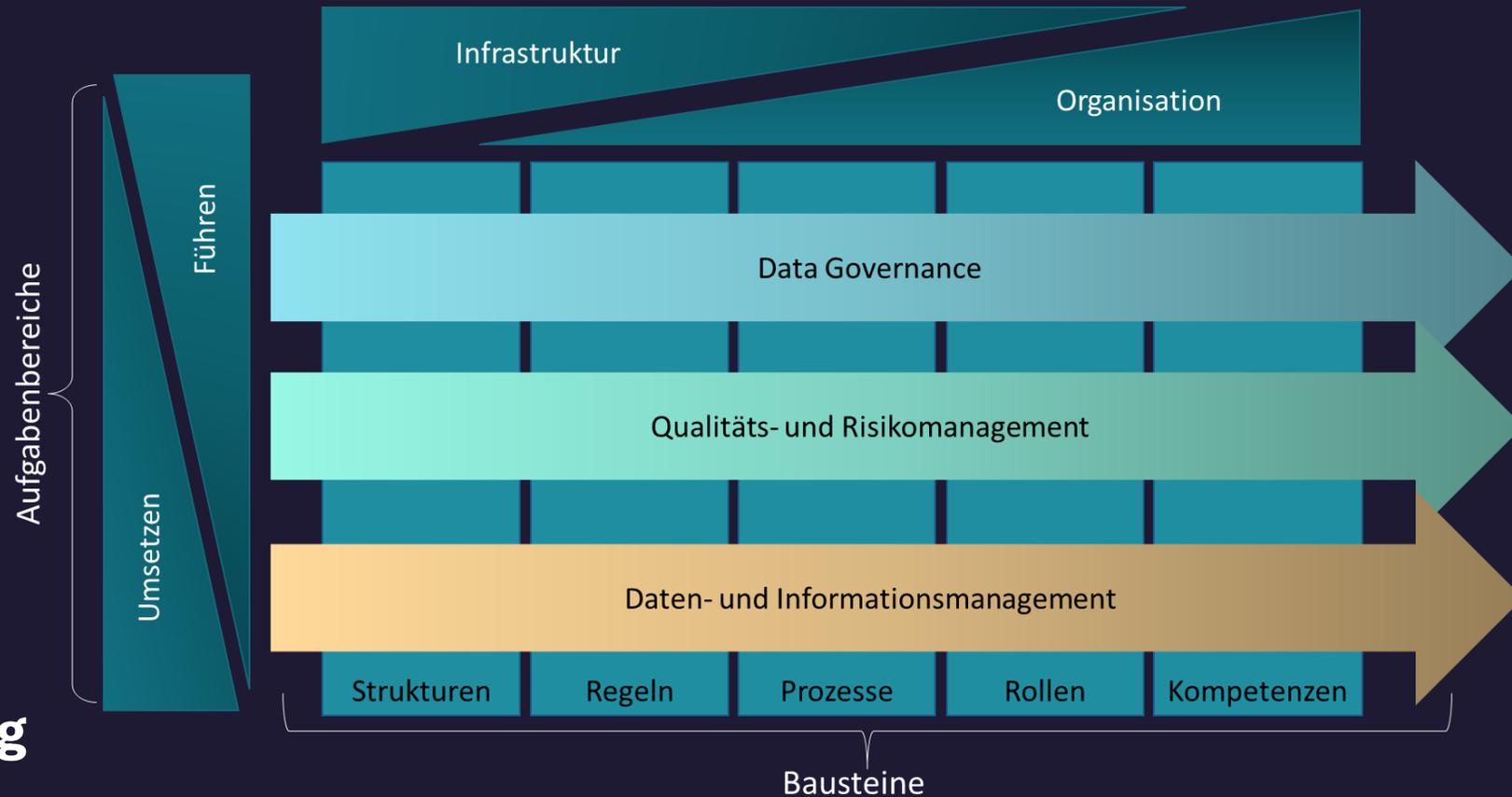
## „Daten-DNA“ durchdringt

- **Strukturen**
- **Regeln**
- **Prozesse**
- **Rollen**
- **Kompetenzen**

## Kultur-Change betrifft

- **Organisation & Infrastruktur**
- **Führung & Umsetzung**

→ **Rahmenbedingungen schaffen, um DATEN AUS SILOS HERAUSZUHOLEN**





## FAZIT

**Datenstrategie muss Organisation wie ein Kompass an 3 Fragen ausrichten:**

- 1. Was wollen und können wir zukünftig mit Daten machen?**
- 2. Wie machen wir es richtig?**
- 3. Und wie machen wir das Richtige?**

→ **Ethische Parallele** von Daten und Öl:

**Missbrauch führt zu erheblichen Schäden**

**Öl: Verkehrskollaps, Feinstaub, Plastikmüll in den Meeren**

**Daten: Verlust von Selbstbestimmung, Diskriminierung, Fehlentscheidungen**

# SPARBUCH

**GELD** verdient man heute mit **DATEN**.

Darauf muss sich die Finanzbranche vorbereiten - heute, nicht morgen.

**CHANCE:** Den Menschen nicht aus dem Blick zu verlieren

→ Wie wollen wir zukünftig **WERT** schaffen, ohne **WERTE** zu verletzen?

Ethik & (Wert-)Haltung garantieren kein exponentielles Wachstum, aber ermöglichen nachhaltiges

# SPARBUCH

**We have a new resource here.  
But nobody wants „data“.  
What they want are the answers.**