



**Springer Vieweg**

ISBN 978-3-658-31938-0 | eBook

ISBN 978-3-658-31937-3 | Softcover

<https://link.springer.com/book/10.1007/978-3-658-31938-0>

The book is currently only available in German.

[www.ai-one.com](http://www.ai-one.com)

[www.braindogs.ai](http://www.braindogs.ai)

[www.bii-institute.com](http://www.bii-institute.com)

[www.diwa-capital.com](http://www.diwa-capital.com)



## Fragen an Walt Diggelmann zum Thema: Semantischer Fingerprint

Im Kapitel 11 vom Buch: «Semantische Daten Intelligenz im Einsatz» sprechen Sie vom semantischen Fingerprint, können Sie diesen etwas genauer erklären?

Jeder kennt den forensischen Fingerprint und die Eigenschaft, dass alle Fingerprints unterschiedlich sind, selbst bei eineiigen Zwillingen.

Das ist so, was hat das jetzt mit Semantik zu tun?

Beim forensischen Fingerprint besteht die Datei aus sogenannten «Minutiaen». Das sind charakteristische Punkte beim Linienverlauf der Rillen in den Händen und Fingerkuppen.

Genau, das sind die sogenannten Templates, welche aus dem gescannten Fingerprint-Bild konvertiert wurden, diese können mit Computerprogrammen blitzartig mit Hunderten Millionen anderer Templates verglichen werden.

Betrachtet man die Worte nun auch als Datenpunkte in einem Text und die Abfolge der Worte als Kette und Verlauf, so kann man daraus eine Charakteristik und ein Muster erkennen, wie bei den Minutiaen.

Aha, ich verstehe, Sie betrachten einen Satz also nicht als ASCII Worte, sondern als eine Art Mustergrafik.

Jeder Text setzt sich aus Buchstaben, Wörtern und Sätzen zusammen. Daraus leiten wir Menschen eine Botschaft ab. Betrachtet man die Buchstaben, Worte und Satzstellung abstrahiert als Datenpunkte, so kann man aus der Anordnung und Abfolge der Worte ein Muster ableiten. Dabei ist ein Wort ein Datenpunkt und kann aus einem oder mehreren Buchstaben bestehen. (vereinfacht erklärt).

Wie verstehen wir Menschen ein solches Text-Muster?

Damit Menschen Texte verstehen, haben sie sich auf einen gemeinsamen Buchstabensatz und auf eine Rechtschreibung (Orthografie) und Satzstellung (Grammatik) geeinigt. So ist es möglich, dass Sender und Empfänger die Botschaft grundsätzlich verstehen.

Damit die Botschaft möglichst im Sinne des Autors verstanden wird, sollte bei der Textanalyse neben der Syntaktik und Semantik auch die Pragmatik mit analysiert werden. Damit wird es zum Semiotischen System. Die Semiotik definiert sich aus dem Zusammenspiel der 3 Basis-Elemente:

**Semiotik: Syntax | Semantik | Pragmatik**

Das ist so weit klar, aber das «Zusammenspiel» ist noch unklar! Vielleicht erklären Sie zuerst die Begriffe.

## Questions for Walt Diggelmann on the topic: Semantic Fingerprint

In chapter 11 of the book: "Semantic Data Intelligence in Use" you talk about the semantic fingerprint, can you explain it a bit more?

Everyone knows the forensic fingerprint and the property that all fingerprints are different, even in identical twins.

Yes, I would agree, now what does that have to do with semantics?

In the forensic fingerprint, the file consists of so-called "minutiae". These are characteristic points in the lines of the grooves in the hands and fingertips.

Exactly, these are the so-called templates, which were converted from the scanned fingerprint image. Now these vector data can be compared with a computer program in a flash with hundreds of millions of other templates.

If you also look at the words as data points in a text and the sequence of words as a chain and progression, you can also see a characteristic and a pattern from this, as with minutiae.

Aha, I see, so you don't think of a sentence as ASCII words, but as a kind of pattern graphic.

Every text is composed of letters, words, and sentences. From these, we humans derive a message. So, if you look at the letters, words and sentence order abstracted as data points, you can derive a pattern from the order and sequence of the words. Thereby a word is a data point and can consist of one or more letters. (Simplified explained)

How do we humans understand such a text pattern?

For us humans to understand texts, we have agreed not only on a common set of letters, but also on spelling (orthography) and sentence order (grammar). This ensures that the recipient understands the same message as the sender/author.

OK, we need to know this a little more precisely.

The meaning and significance of a message, also called semiotics, is formed from the interaction of 3 basic elements:

**Semiotics: Syntax | Semantics | Pragmatics.**

This is clear so far, but the "interaction" is still unclear! Perhaps explain the terms first.

**Semiotics:** is the science of sign processes in culture and nature. So, the connection between syntax, semantics, and pragmatics.

**Semiotik:** ist die Wissenschaft von den Zeichenprozessen in Kultur und Natur. Dabei werden Zeichen wie Buchstaben, Wörter, Hieroglyphen, aber auch Bilder, Skizzen, Zeichnungen, Piktogramme sowie Gesten verstanden. Sie alle vermitteln Informationen aller Art über die Zeit und den Ort bei der Entstehung einer Botschaft. Semiotik, Syntax, Semantik und Pragmatik sind nicht nur auf die Linguistik beschränkt, sondern beziehen sich auf Zeichen und Botschaften im Allgemeinen.

Die Semiotik ist also das Resultat des Zusammenspiels zwischen Syntax, Semantik und Pragmatik im grösseren Allgemeinen gesprochen.

Genau, vor allem bringt die Semiotik die Pragmatik mit in die Betrachtung. Erst wenn die Pragmatik mit in die Analyse einer Botschaft einbezogen wird, spricht man von Semiotik und dadurch kann eine Botschaft genauer ausgewertet werden, weil dann oftmals auch die Motivation zur Botschaft klar wird.

Ok, dann müssen sie noch die anderen drei Begriffe erläutern.

**Syntax:** Ist die Lehre vom Bau des Satzes als Teilgebiet der Grammatik; der Satzlehre. Diese zeigt die in einer Sprache übliche Verbindung von Wörtern zu Wortgruppen und Sätzen; die korrekte Verknüpfung sprachlicher Einheiten im Satz.

**Semantik:** ist die Lehre im Teilgebiet der Linguistik, die sich mit den Bedeutungen sprachlicher Zeichen und Zeichenfolgen befasst. Also Bedeutung, Inhalt (eines Wortes, Satzes oder Textes).

**Pragmatik:** beschäftigt sich in der Linguistik mit der Beschreibung von kontextabhängigen und nicht-wörtlichen Bedeutungen bei der Verwendung von sprachlichen Ausdrücken in jeweils konkreten Situationen und mit den Bedingungen für ihr Entstehen. Vereinfacht: der Bezug zur Realität

Wie kommt jetzt der semantische Fingerprint ins Spiel?

Wenn wir uns also einig darüber sind, dass alle korrekt aufgebauten Texte, egal ob sie gelesen werden oder nicht, inhärenten Regeln der Semiotik folgen, dann haben Texte immer eine inhärente Botschaft, selbst wenn sie weder gelesen noch interpretiert werden. Die Kunst ist jetzt, diese inhärente Botschaft, oder intrinsische Semantik wie wir diese bezeichnen, in eine maschinenlesbare Form zu überführen.

Wozu ist das wichtig?

Wenn es gelingt, die inhärente Botschaft automatisch maschinenlesbar zu konvertieren, und dabei keine Veränderung oder Deutung am Text vorzunehmen, wird wie in der Forensik, ein generischer Fingerprint der Botschaft gebildet.

Der semantische Fingerprint ist ein multidimensionaler Vektorraum, in dem das semantische Muster der Botschaft gespeichert ist.

Es ist ein Faksimile der Botschaft. Nachfolgende Prozesse können den Fingerprint weiter auswerten und in Kontext stellen, mit anderen Botschaften vergleichen und matchen etc. Der semantische Fingerprint ist ein optimales Daten-Transport-Format, so gesehen der optimale Input für nachfolgende ML- (maschinelles Lernen) und KI- (Künstliche Intelligenz) Systeme.

I understand, semiotics is the result of the interaction between syntax, semantics, and pragmatics.

Exactly, if pragmatics (relation to reality) is included in the analysis of a message, a message can be evaluated more accurately.

Ok, then they still need to explain the other three terms.

**Syntax:** Is the study of the construction of the sentence as a branch of grammar; sentence theory. In a language usual connection of words to word groups and sentences; correct connection of linguistic units in the sentence.

**Semantics:** is the branch of linguistics that deals with the meanings of linguistic signs and sequences of signs. Thus meaning, content (of a word, sentence, or text).

**Pragmatics:** in linguistics deals with the description of contextual and non-literal meanings in the use of linguistic expressions in concrete situations and with the conditions for their emergence. Simplified: the relation to reality

Now how does the semantic fingerprint come into play?

If we agree that all correctly constructed texts, whether they are read or not, follow an inherent rule of semiotics, then texts always have an inherent message, even if they are neither read nor interpreted.

The art now is to convert this inherent message, or intrinsic semantics as we call it, into a machine-readable form.

Why is this important?

If we succeed in automatically converting the intrinsic message into a machine-readable form, without changing or interpreting the text, we create a generic fingerprint of the message, as in forensics.

The semantic fingerprint is a multidimensional vector space in which the semantic pattern of the message is stored.

It is an facsimile of the message. Subsequent processes can further evaluate and contextualize the fingerprint, compare, and match it with other messages, etc.

The semantic fingerprint is an optimal data transport format, so seen as the optimal input for downstream ML (machine learning) and AI (artificial intelligence) systems.

What is the advantage of making messages machine-readable?

The advantage of ML and AI lies in the fast processing of large amounts of text by computer systems. Once the message is machine-readable, computers can process massively more information and miss nothing. They don't get tired and are ready 7/24.

ML can compare, match, classify and sort any messages without understanding, thanks to the semantic fingerprint. This ensures that messages with similar content are found without prior interpretation. ML and AI can then be used again for the interpretation.

I see, ML and AI can search for similarities, contradictions, supporting or negating messages in millions of texts at lightning speed.

## Worin liegt der Vorteil, wenn Botschaften maschinenlesbar sind?

Der Vorteil von ML und KI liegt in der schnellen Verarbeitung von grossen Textmengen durch die Computer Systeme. Sobald die Botschaft maschinenlesbar ist, können Computer massiv mehr Information verarbeiten und übersehen dabei nichts. Sie werden nicht müde und sind 7/24 bereit.

ML kann dank dem semantischen Fingerprint irgendwelche Botschaften, ohne diese zu verstehen vergleichen, matchen, klassifizieren und sortieren. Damit ist gewährleistet, dass inhaltlich ähnliche Botschaften ohne vorherige Interpretation gefunden werden. Für die Interpretation kann in der Folge dann nochmals ML und KI eingesetzt werden.

## Ich verstehe, ML und KI können so blitzschnell in Millionen von Texten nach Ähnlichkeiten, Widersprüchen, unterstützenden oder negierenden Botschaften suchen.

Alles mit dem Ziel, dem Menschen eine möglichst ausgewogene Informations-Basis zu geben, damit er die bestmögliche Entscheidung treffen kann. Es gibt nichts Gefährlicheres, als wenn sich Menschen bewusst oder unbewusst, in einer geprägten Informations-Blase bewegen und trotzdem glauben, frei und ausgewogen informiert zu sein, um ihre Entscheidungen zu treffen.

## Dann wären wir wieder beim Thema Fake News oder alternativen Tatsachen. Einfach gesagt: Lügengeschichten.

Menschen, welche sich bewusst oder unbewusst in einer Informations-Blase bewegen, bemerken gar nicht, wie einseitig sie sich informieren. Paradoxerweise ist es so, dass diese Menschen das selbst nicht tolerieren, was sie von anderen fordern, alternative Ansichten! Solche stark fokussierte Personen verweigern jegliche andere Meinungen und taxieren jeden Andersdenkenden als Feind. Sie bezeichnen sich als tolerant und fordern Toleranz, sind aber zutiefst ignorant und intolerant.

Es ist wie bei an Bulimie erkrankten Menschen, diese können spindeldürr vor einem Spiegel stehen und sich selbst trotzdem als fett wahrnehmen.

Stark beeinflusste Menschen folgen bedingungslos ihrem Idol. Sie bemerken nicht, dass sie missbraucht werden, damit das Idol seine Partikular Interessen umsetzen kann.

## Wir kann man solche Menschen aus diesem Tunnelblick befreien?

Es gibt immer einen Anteil von Minderheiten, welche man nicht aus ihrem Tunnel befreien kann. Da würde nur ein dramatisches Ereignis, oder der Sinneswandel des ideologischen Führers helfen. Diese Menschen sind nicht an anderen Meinungen interessiert, diese könnten ihr fixes Weltbild und ihre Glaubenssätze zerstören.

## Das ist ein tragisches Bild, was sie skizzieren.

Das stimmt, aber es ist leider so.

© ai-one | wd | 25. Oktober 2021

All with the goal of giving humans the most balanced information base possible so they can make the best decision possible.

There is nothing more dangerous than when people move, consciously or unconsciously, in an imprinted bubble (filter cloud) and still believe to be freely and well-informed to make their decisions.

## Then we would be back to the topic of fake news or alternative facts. Simply put: Lying stories.

People who move consciously or unconsciously in an information bubble don't even notice how one-sided they inform themselves. Paradoxically it is so that these people do not tolerate themselves what they demand from others, alternative views! Such strongly focused persons refuse any other opinions and rate every dissenter as an enemy. They call themselves tolerant and demand tolerance but are deeply ignorant and intolerant.

It is like people suffering from bulimia, they can stand spindly in front of a mirror and still perceive themselves as fat.

Strongly influenced people follow their idol unconditionally. They do not notice that they are abused so that the idol can implement its particular interests.

## How can such people be freed from the tunnel?

There is always a percentage of minorities that cannot be freed from their tunnel. Only a dramatic event or the ideological leader's change of heart would help. These people are not interested in other opinions, which could destroy their fixed worldview and beliefs.

## This is a tragic picture they sketch.

It is true, but this is unfortunately the case.

© ai-one | wd | 25. Oktober 2021

