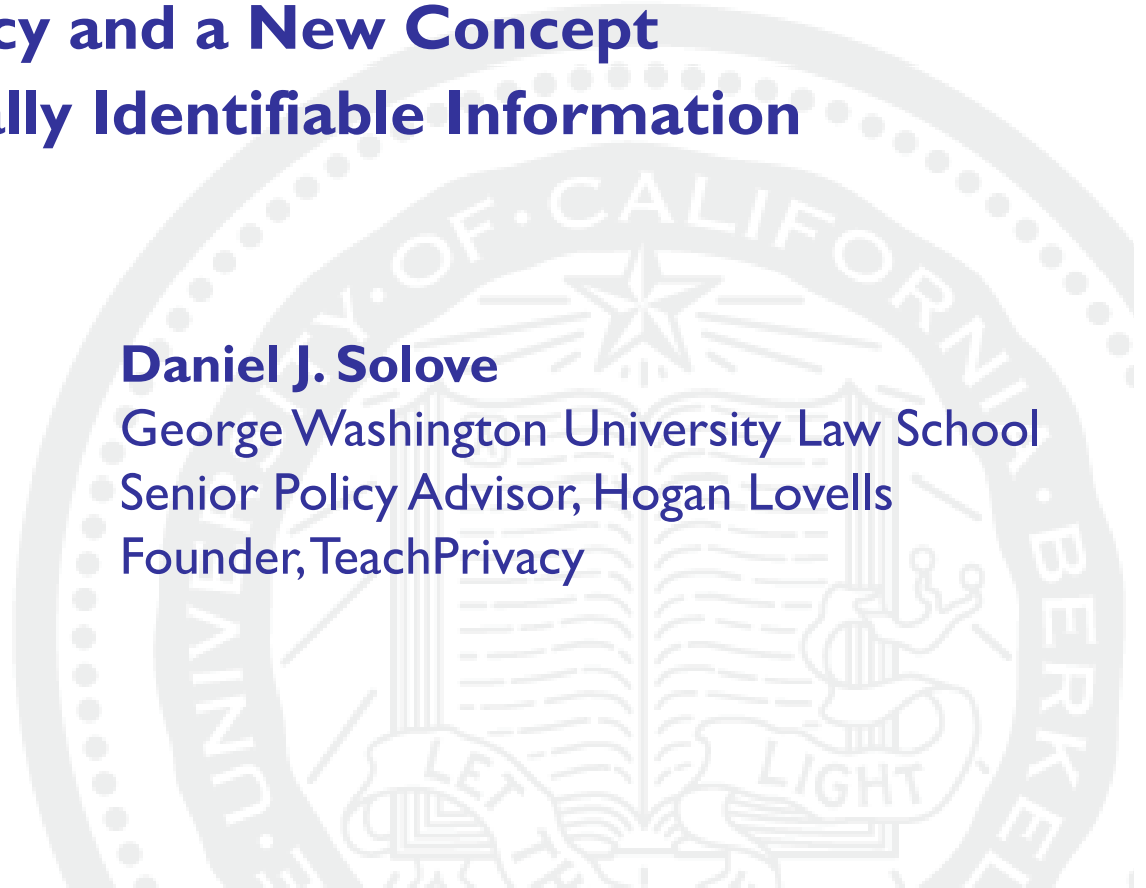




# **The PII Problem: Privacy and a New Concept of Personally Identifiable Information**

**Paul M. Schwartz**  
Berkeley Law School

**Daniel J. Solove**  
George Washington University Law School  
Senior Policy Advisor, Hogan Lovells  
Founder, TeachPrivacy



# THE PII PROBLEM: PRIVACY AND A NEW CONCEPT OF PERSONALLY IDENTIFIABLE INFORMATION

PAUL M. SCHWARTZ<sup>†</sup> & DANIEL J. SOLOVE<sup>‡</sup>

*Personally identifiable information (PII) is one of the most central concepts in information privacy regulation. The scope of privacy laws typically turns on whether PII is involved. The basic assumption behind the applicable laws is that if PII is not involved, then there can be no privacy harm. At the same time, there is no uniform definition of PII in information privacy law. Moreover, computer science has shown that in many circumstances non-PII can be linked to individuals, and that de-identified data can be re-identified. PII and non-PII are thus not immutable categories, and there is a risk that information deemed non-PII at one time can be transformed into PII at a later juncture. Due to the malleable nature of what constitutes PII, some commentators have even suggested that PII be abandoned as the mechanism by which to define the boundaries of privacy law.*

*In this Article, we argue that although the current approaches to PII are flawed, the concept of PII should not be abandoned. We develop a new approach called “PII 2.0,” which accounts for PII’s malleability. Based upon a standard rather than a rule, PII 2.0 utilizes a continuum of risk of identification. PII 2.0 regulates information that relates to either an “identified” or “identifiable” individual and it estab*

# „Personenbezogene Daten“ aus internationaler Perspektive

Das Datenschutzrecht beruht in erster Linie auf dem Konzept der „personenbezogenen Daten.“ Informationen innerhalb dieser Kategorie sind geschützt, Informationen außerhalb sind es nicht. Angesichts der zentralen Bedeutung des Konzepts der personenbezogenen Daten erscheint es überraschend, dass es keine weltweit übereinstimmende Definition dieses Begriffs gibt.

Der Mangel an Einheitlichkeit hinsichtlich eines so grundsätzlichen Konzepts hat erhebliche Auswirkung im Zeitalter der globalen Datenübertragung. Informationen, die in Deutschland als personenbezogene Daten angesehen werden, unterfallen möglicherweise keinerlei Datenschutzbestimmungen in den Vereinigten Staaten.

Darüber hinaus führt die weltweite Entwicklung der Informationstechnologie zu einer Verschiebung der Grenze zwischen personenbezogenen und sonstigen Daten. In vielen Fällen ist es daher schwierig bereits im Vorfeld zu entscheiden, ob bestimmte Informationen personenbezogene Daten sind oder nicht. Aus diesen Gründen sollte den unterschiedlichen nationalen Definitionen des Begriffs der personenbezogenen Daten sowie dem Erfordernis der Bemühungen um eine globale Harmonisierung in Zukunft größere Bedeutung beigemessen werden.

## „Personenbezogene Daten“ im Recht der Europäischen Union und nach deutschem Recht

Sowohl das Recht der Europäischen Union als auch das deutsche Recht folgen bei der Definition des Konzepts der persönlichen Daten einem extensiven Ansatz. Demzufolge definiert die EU Da-

Als Auslegungshilfe nennt die Richtlinie in Erwägungsgrund Nr. 26 außerdem, dass „(b)ei der Entscheidung, ob eine Person bestimmbar ist, ... alle Mittel berücksichtigt werden (sollten), die vernünftigerweise entweder von dem Verantwortlichen für die Verarbeitung oder von einem Dritten eingesetzt werden könnten, um die betreffende Person zu bestimmen“ (a.a.O., Erwägungsgrund 26).

Mithin werden nach EU-Recht Informationen über eine „bestimmbare“ Person genauso behandelt wie Informationen, die sich auf eine „bestimmte“ Person beziehen. Dieses Konzept der Gleichstellung von „bestimmten“ und „bestimmbaren“ Personen ist als deutsche Innovation im Bereich des Datenschutzrechts anzusehen. Die Berücksichtigung potenzieller Gefahren

ausgehend von Daten über „bestimmbare“ Personen fand bereits in der ersten Fassung des Bundesdatenschutzgesetz (BDSG) von 1977 Ausdruck. Zu dieser Zeit bestand die Befürchtung, dass die zunehmende Verwendung von Computern in Kombination mit den neuen Möglichkeiten des Datenverarbeitungsmanagements zu einer Re-Individualisierung von Daten führen könnte.

Entsprechend definiert das BDSG auch heute „personenbezogene Daten“ sowohl als Informationen über „bestimmte“ als auch über „bestimmbare“ Personen. Die entscheidende Regelung findet sich in § 3 Abs. 1 BDSG, in welchem „personenbezogene Daten“ als „Einzelangaben über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmbaren natürlichen Person (Betroffener)“ definiert sind.

Die Gleichstellung von „bestimmten“ und „bestimmbaren“ Personen



**Prof. Paul M. Schwartz** lehrt an der Berkeley Law School, ist Direktor des Berkeley Center for Law & Technology und Mitglied des Wissenschaftsbeirats der ZD.

---

Reproduced with permission from Privacy & Security Law Report, 11 PVLR 142, 11/23/2012. Copyright © 2012 by The Bureau of National Affairs, Inc. (800-372-1033) <http://www.bna.com>

---

**Consumer Privacy****Personally Identifiable Information**

There is no uniform definition of “Personally Identifiable Information,” or PII, in the United States, the authors say. For privacy law to remain effective in the future, they have reconceptualized PII as “PII 2.0,” a model that categorizes PII as “identified” or “identifiable” and further recognizes a category of “non-identifiable” information.

**PII 2.0: Privacy and a New Approach to Personal Information**

statutes turn on this distinction.<sup>1</sup> Similarly, many state statutes also rely on PII as a jurisdictional trigger.<sup>2</sup> These laws all share the same basic assumption—that in the absence of PII, there is no privacy harm. Thus, privacy regulation focuses on the collection, use, and disclosure of PII, and leaves non-PII unregulated.

Given PII’s importance, it is surprising that information privacy law in the United States lacks a uniform definition of the term. In addition, computer science has shown that the very concept of PII is far from straightforward. Increasingly, technologists can take informa-

# Schwartz and Solove

---



# Changes in Technology and the Meaning of PII

---



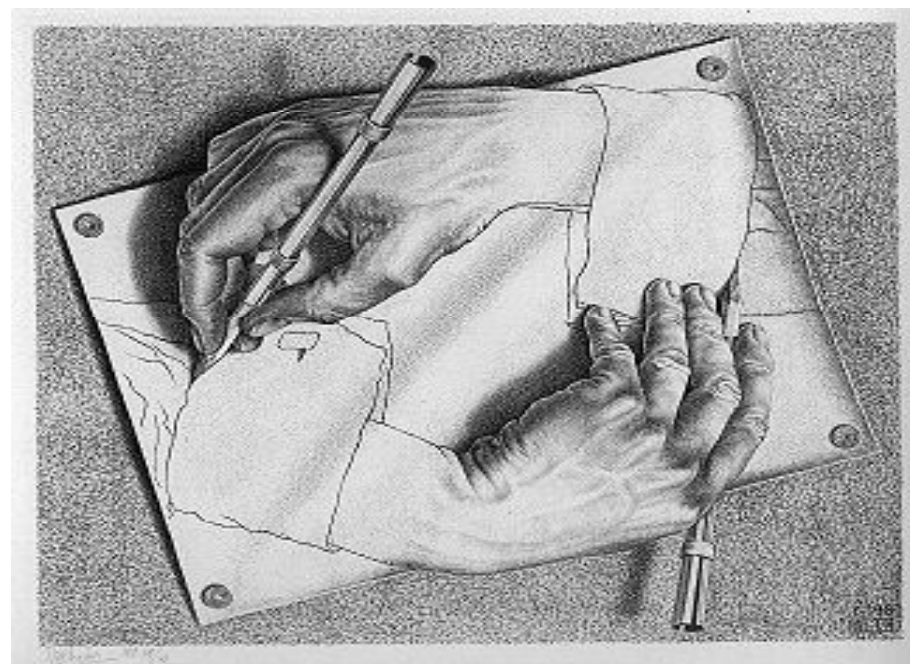
# Three Approaches to PII in US Law

---

1. Tautological Approach

2. Non-Public Approach

3. Specific Types Approach



## No uniform international definition of PII

---

- PIPEDA uses term “identifiable” data
- Tendency is for broad definition of PII: PIPEDA reflects EU perspective





# EU approach to PII

---

Broad definition:

“information relating to an identified or identifiable person”



Identifiable = identified

---

Personal data if “the reference person is identifiable”

Dammann, Kommentar zum BDSG, (Simitis, ed., 2011)



# Problems of De-Identification



The screenshot shows the Netflix Prize website interface. At the top, the Netflix logo is on the left, and a large yellow banner with the text "Netflix Prize" and a red "COMPLETED" stamp is on the right. Below the banner is a navigation bar with links for "Home", "Rules", "Leaderboard", and "Update". The main content area features a "Movies For You" section with a list of recommended titles and a "Congratulations!" message box on the right. The message box contains the following text:

**Congratulations!**

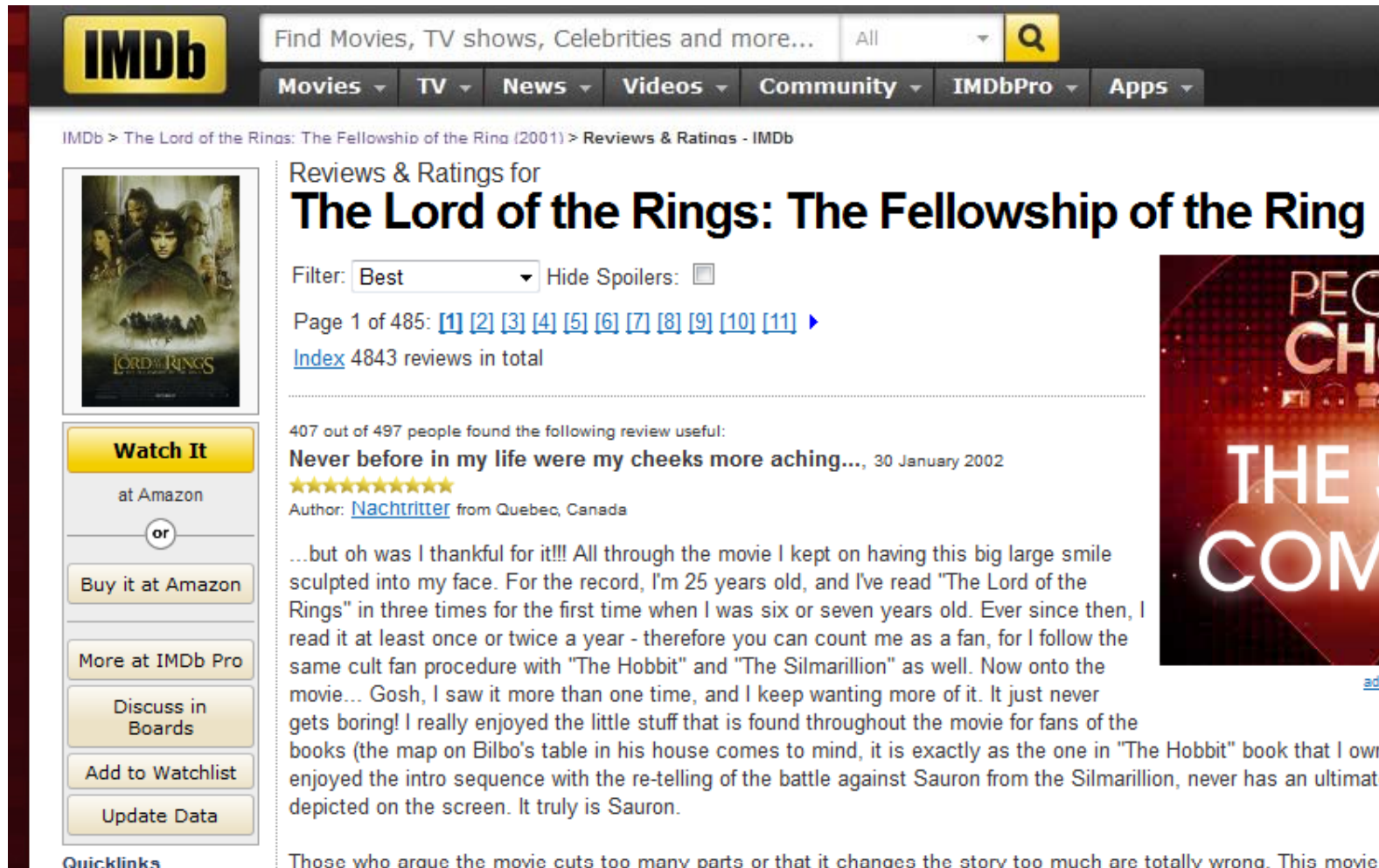
The Netflix Prize sought to substantially improve the accuracy of predictions about how much someone is going to enjoy a movie based on their movie preferences.

On September 21, 2009 we awarded the \$1M Grand Prize to team "BellKor's Pragmatic Chaos". Read about [their algorithm](#), checkout team scores on the [Leaderboard](#), and join the discussions on the [Forum](#).

We applaud all the contributors to this quest, which improves our ability to connect people to the movies they love.

At the bottom of the page, there are links for "FAQ", "Forum", and "Netflix Home", and a copyright notice: "© 1997-2009 Netflix, Inc. All rights reserved."

# Internet Movie Database



The screenshot shows the IMDb website interface. At the top, there is a search bar with the text "Find Movies, TV shows, Celebrities and more..." and a dropdown menu set to "All". Below the search bar are navigation tabs for "Movies", "TV", "News", "Videos", "Community", "IMDbPro", and "Apps". The main content area is titled "IMDb > The Lord of the Rings: The Fellowship of the Ring (2001) > Reviews & Ratings - IMDb". The page features a movie poster on the left, a "Watch It" button, and a "Buy it at Amazon" button. The central review is titled "Never before in my life were my cheeks more aching..." and is rated with 10 stars. The author is "Nachtritter" from Quebec, Canada. The review text discusses the author's long-standing fandom and their appreciation for the movie's details. On the right side, there is a vertical banner for "PEOPLE'S CHOICE AWARDS THE COMING OF AGE".

IMDb > The Lord of the Rings: The Fellowship of the Ring (2001) > Reviews & Ratings - IMDb

Reviews & Ratings for  
**The Lord of the Rings: The Fellowship of the Ring**

Filter: Best Hide Spoilers:

Page 1 of 485: [\[1\]](#) [\[2\]](#) [\[3\]](#) [\[4\]](#) [\[5\]](#) [\[6\]](#) [\[7\]](#) [\[8\]](#) [\[9\]](#) [\[10\]](#) [\[11\]](#) ▶

[Index](#) 4843 reviews in total

407 out of 497 people found the following review useful:

**Never before in my life were my cheeks more aching...**, 30 January 2002  
★★★★★★★★★★  
Author: [Nachtritter](#) from Quebec, Canada

...but oh was I thankful for it!!! All through the movie I kept on having this big large smile sculpted into my face. For the record, I'm 25 years old, and I've read "The Lord of the Rings" in three times for the first time when I was six or seven years old. Ever since then, I read it at least once or twice a year - therefore you can count me as a fan, for I follow the same cult fan procedure with "The Hobbit" and "The Silmarillion" as well. Now onto the movie... Gosh, I saw it more than one time, and I keep wanting more of it. It just never gets boring! I really enjoyed the little stuff that is found throughout the movie for fans of the books (the map on Bilbo's table in his house comes to mind, it is exactly as the one in "The Hobbit" book that I own enjoyed the intro sequence with the re-telling of the battle against Sauron from the Silmarillion, never has an ultimate depicted on the screen. It truly is Sauron.

Those who argue the movie cuts too many parts or that it changes the story too much are totally wrong. This movie

PII and non-PII-- not a fixed line

---

# Impact of technology developments and social practices



# Abandon PII?

---

PII



# Abandon PII?

---



# Keep PII? Abandon PII as Regulatory Concept?

---



Just  
regulate  
data?



## PII 2.0

---

- Identifiability is a continuum of risk.
- A standard not a rule
- Not a hard “off-on” switch, but tailored Fair Information Practices



PII 2.0: Three categories

---

Identified

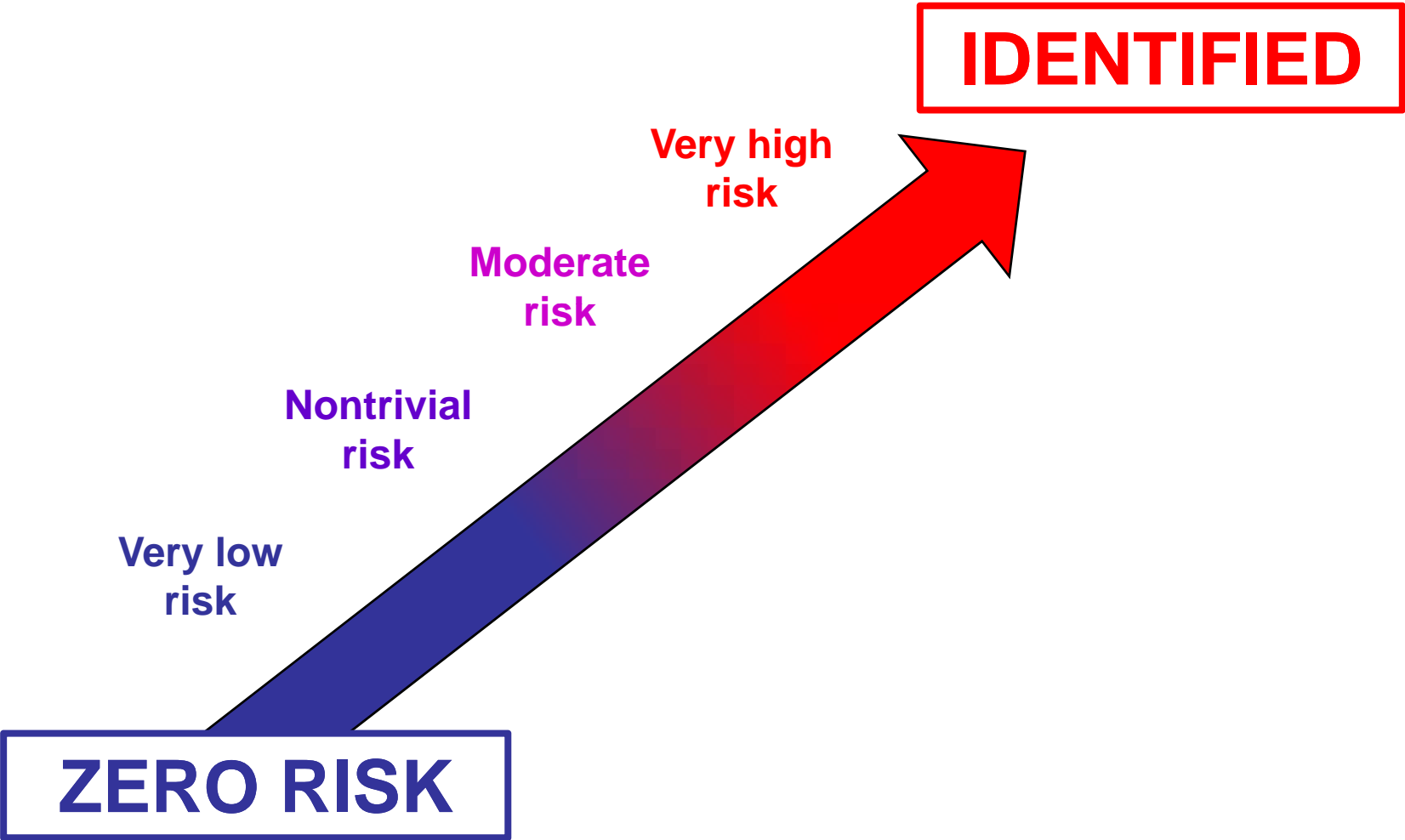
Identifiable

Non-Identifiable



# Risk of Identification

---



## PII 2.0: Three categories

---

# Identified

- plus identifiable data when significant probability of linkage to specific person

# Identifiable

# Non-Identifiable



# PII 2.0 -- Dangers of “Release and Forget”

---

Need for:

Track-and-audit approach

Risk assessments



PII 2.0 = compatible with “privacy by design”

---



Privacy protection  
embedded in  
technological  
design and business  
practices

## Takeaway

---

- Great legal uncertainty about concept of PII– and on worldwide basis
  - Hard to predict impact of privacy law on businesses: a source of risk
- 



Thank you!

---

