淺談汽車領域的安全三連：Safety 、Security 和 SOTIF (2)

以下原文源於 20211229 EDN TAIWAN 電子報

在功能安全(FuSa)、資訊安全(Cybersecurity)和預期功能安全(SOTIF)這個三連的「指導」下，對內促進汽車電子電氣系統的高效可靠，對外提高汽車智慧化、抗干擾、抗攻擊的能力，才能確保人身安全和資料隱私這塊汽車領域的 346.6 高地…

編按：*本文是「淺談汽車領域的安全三連：Safety 、Security 和 SOTIF」一文的後續報導，在針對汽車安全密不可分的三個主題加以定義與初步介紹後，本文將深入剖析攸關汽車安全的場景與標準流程…*

最近忙裡偷閒在追劇《功勳-能文能武李延年》，非常敬佩劇中指導員李延年。對內每個戰士的背景特長如數家珍，對外敵人的戰術和心理也瞭若指掌。其實汽車安全的要求也是這樣。在功能安全(FuSa)、資訊安全(Cybersecurity)和預期功能安全(SOTIF)這個三連的「指導」下，對內促進汽車電子電氣系統的高效可靠，對外提高汽車智慧化、抗干擾、抗攻擊的能力，這樣才能確保人身安全和資料隱私這塊汽車領域的 346.6 高地。

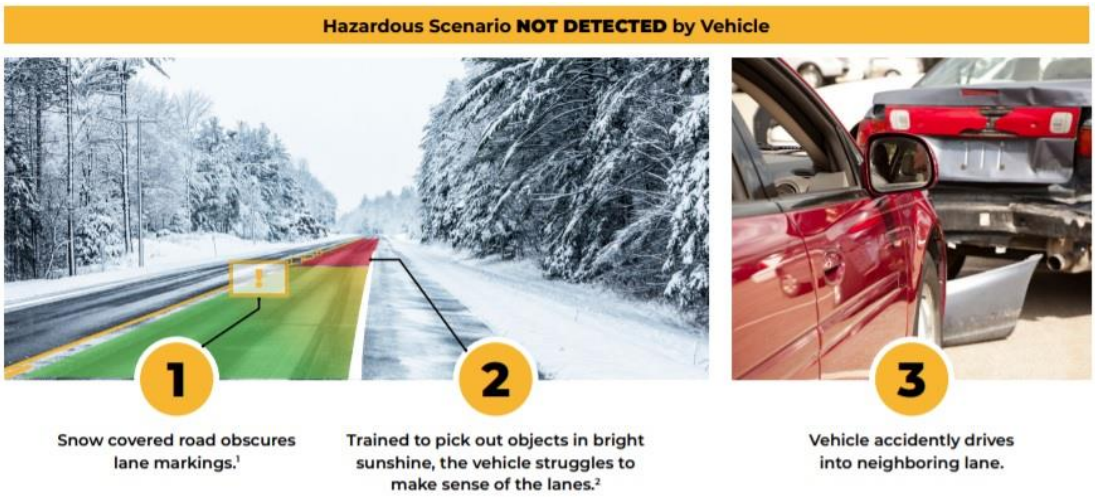

圖 1：自動駕駛系統功能不足造成的事故。(來源：https://www.ansys.com/content/dam/product/systems-embedded-and-integrated/medini/sotif-infographic.pdf)

在汽車安全領域，如何提高安全也要能文能武，要能結合實際場景，配合專業知識經驗，甚至向「敵人」學習，來制定不同的策略。文能結合理論、總結日誌、制定標準，武能親自實踐、靈活變通、百戰不殆，這樣才能真正做到知行合一，「打得贏」汽車安全這場「仗」。

上一篇淺談了這三方各自的定義和範疇，說了它們的相互關聯和承繼，也提到作為車廠，在面對安全這個話題上，有的「結硬寨，打呆仗」，穩紮穩打；有的有亮劍的精神，在失敗中總結反覆運算，浴火重生。作為利益相關者，也是個局中迷者，我不想講黑白對錯，想做的只是儘量不站隊地繼續淺談汽車安全。大家有興趣的話，建議可以先看看第一篇簡單瞭解大背景。

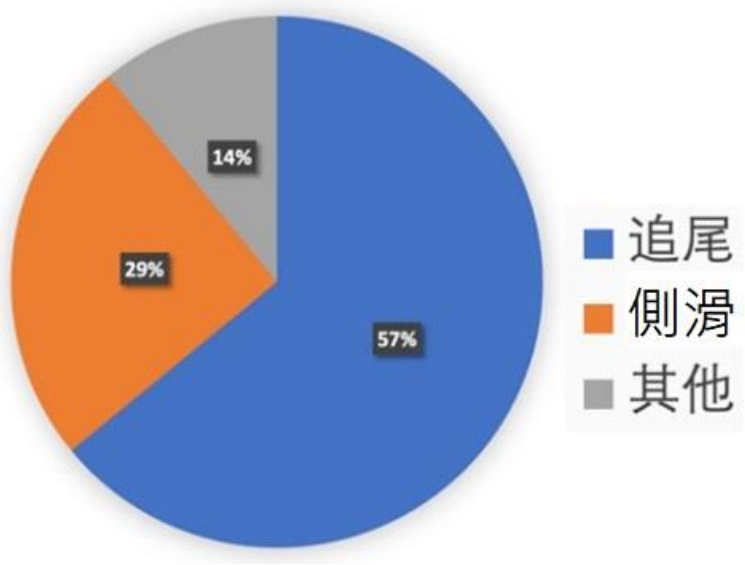

圖 2：加州自動駕駛事故調查。

隨著自動駕駛和先進駕駛輔助系統(ADAS)的到來，越來越多的自動駕駛交通事故開始出現(**圖 1**)。加州的一項統計指出 57%的自動駕駛相關事故是幽靈剎車導致的追尾，29%來源於脫離車道(**圖 2**)。SOTIF 標準旨在提高自動駕駛系統的能力，保證自動駕駛汽車的安全。本篇將主要圍繞 SOTIF 和一些相關實例展開。

安全三連的金字塔

說到安全三連，這三個 S 看起來似乎是平等的地位，但現實是這樣嗎？

FuSa 面對的是汽車自己的系統，是對內的，要解決的是自己內在系統失效和隨機失效(Failure)。SOTIF 則是應汽車智慧化趨勢而生，彌補了 FuSa 在人工智慧(AI)領域、自動駕駛領域的不足，涵蓋了自動駕駛等級 L1-5。它更多面對的是

性能失效、系統預期功能不足和人員誤操作。這就開始向外延伸，更多地涉及到汽車與環境的互動，汽車和駕駛操作人員的互動。資訊防禦安全 Cybersecurity 則不是因「系統」而生，而是因「人」而生，針對來自於外部環境、外來惡意者甚至內在惡意者的威脅(Threat)，加強的手段有全狀態防火牆、對稱非對稱加密、金鑰管理和入侵防禦和檢測系統等。

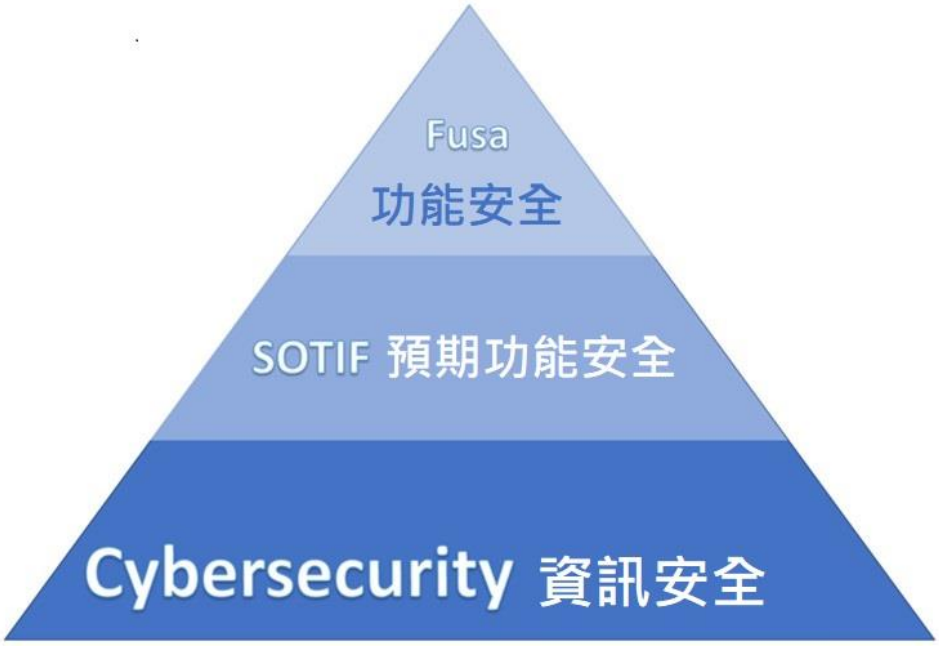

圖 3：汽車安全三連的金字塔地位。

比較三者的物件，不難看出，主動的「人」是最難對付的，其次是複雜多變的「環境」。其實最好面對的反倒是自己，要麼為什麼說「達則兼濟天下，窮則獨善其身」哪。從這個角度，不如把這三者想成是一個金字塔結構(**圖 3**)，FuSa 是塔尖，SOTIF 居中，Cybersecurity 是最下方的根基，先保證了下層，

才能更保證上層。沒有可靠的資訊安全，僅僅保證上層的功能安全就是徒勞。SOTIF 就是這樣一個承上啟下、攘外安內的新角色。

系統性能限制和人員誤用

說 SOTIF，必然要說汽車智慧化，說汽車智慧化就必然要說 AI。AI 的核心大致可以說是基於複雜不可量化的輸入，經過神經網路的運算，給出辨識結果。這個結果不可能做到 100%的準確，只能盡可能地貼近現實，這是一個統計學機率問題。而在汽車領域，駕駛環境複雜多變，各種靜態和動態的物件共存，現實世界的多維更加劇了 AI 運算的量化計算複雜度。只有保證了 AI 演算法，自動駕駛系統的決策才能得到保證。

從更高的高度看，決策還僅僅是自動駕駛系統功能中的一環。除此之外，向前端需要通過感測器系統對環境因素採集輸入，是否能夠得到客觀正確的資訊資料是一切的前提；往後得到 AI 演算法決策後，汽車系統能否合理的回應，採取正確的輸出控制也同樣至關重要。這些步驟不論哪裡出現問題都可能會導致自動駕駛系統出現功能安全問題。SOTIF 主要是和這些與自動駕駛相關的新功能適配，比如目標感知定位和路徑規劃等等。

SOTIF 的另一個重要的適用方向是合理合乎邏輯可預見的人員誤用(misuse)。汽車的設計人員瞭解汽車和人之間的人機介面，也瞭解人們對自動駕駛系統的操作習慣，要做的就從這裡出發，發揮想像，預先找到各種人們濫用誤用車載

系統的情形，因為當人數和駕駛次數達到一定量級，總有駕駛人次對車載系統產生錯誤的期望而作出匪夷所思的錯誤操作。

SOTIF 和 FuSa 類似是針對整車層面，而不是離散的元件層面，不同的是 FuSa 的系統安全問題是歸因於失效，而 SOTIF 擴大到了由設計缺陷或者誤操作帶來的功能安全問題。

SOTIF 和場景

SOTIF 處理的物件是場景(Scenario)，比如早晨和傍晚迎著太陽自動駕駛，風雪或者落葉遮擋了公路的標誌線，暴風暴雨天氣產生的惡劣可見度等等。如果感測器系統不足以正確辨識公路和前方障礙，AI 系統不能分別炫光產生的大面積白色和一輛橫亙路中的白色卡車，自動緊急制動系統(Automated Emergency Braking System；AEB)就會錯誤進行急剎車，或者該剎車的時候不剎車。也可能車道保持系統沿著落葉或者積雪的邊際而不是白線進行了轉向，這些都是巨大的安全隱患。

這些場景下，車載系統按照設計開發時的功能都在工作，並沒出現問題，只是出現了不足。類似這類對周圍環境的感知辨識不足的問題並不包括在 FuSa 的範圍內。作為補充的 SOTIF 在這些領域需要發揮自己的用武之地。

圖 4：加拿大馬路上逼真的頑童貼紙。(來源：

https://www.engadget.com/2010-09-08-optical-illusion-lets-you-safely-run-over-fake-children.html)

當然，我也不認為用戶應該僅僅把目光聚焦在這些極端安全問題上，而面對自動駕駛躊躇不前。其實在更多場景下，一個完善的自動駕駛系統的感知辨識反應能力都是超過人類的。這是一個平衡的問題，也是一個機率問題，汽車產業的從業人員要做的就是最大範圍地降低風險的機率，就像坐飛機也有風險，但如果這個風險降到了小數點後有 9 個零，那麼它也會成為最安全的旅行方式之一。

回到 SOTIF 的場景，關注 SOTIF 的人肯定對**圖 4** 不陌生。這是加拿大的一個叫 Preventable (可避)的組織在某個街區的路上貼了一張高模擬的頑童貼紙，30 米以為看起來非常逼真，目的是測試經過司機的反應。可是僅貼了一個禮拜

就撤掉了，估計是怕引起後車司機的震怒甚至追尾。那麼一輛自動駕駛的汽車開到這裡會如何反應呢？

SOTIF 和真假陰陽

圖 5：自動駕駛的混淆矩陣。

圖 5 是機器學習(ML)裡的混淆矩陣。橫行是預測表現，也就是自動駕駛車對場景的辨識，在頑童貼紙的場景下，就是汽車決策這是不是需要剎車的場景，1 為剎車，0 為不剎車。縱列是實際情況，也就是真實世界是不是需要剎車的緊要場景，1 為緊要應剎車，0 為正常勿剎車。當兩者相同的時候結果就是決策正確的，為「真」；如果兩者不同，就是決策失誤了，結果為「假」。四種結果分別為：真陽(True Positive；TP)、假陽(False Positive；FP)、假陰(False Negative；FN)和真陰(True Negative；TN)。我們追求的是「真」(True)，不論是真陽性還是真陰性，都是好的結果，而要避開「假」。

在**圖 4** 中這種情況下，馬路上的小朋友是個假像，車載系統應該透過感測器系統和 AI 決策系統得出這並非是一個緊急剎車的場景，而讓 AEB 不採取任何動作地開過去。也就是說應該得到一個真陰性的結果，不剎車。如果車載系統判斷錯了，在圖片面前剎車了，那就是假陽性了，是一個錯誤的決策，可能造成後車追尾的安全問題。

但如果這個小朋友是一個真的小朋友，那麼理想的車載系統就應該得到一個真陽性的結果，而通知 AEB 緊急剎車制動。如果這種情況下感測器沒有辨識出小朋友而未剎車，那就是假陰性了，出的事情可能更大。所以只要結果為「假」，就證明系統辨識錯了，就會出現大家不想看到的結果。

那麼可不可以設計出一個完美純「真」的系統呢？可恨的是魚與熊掌往往不可兼得，真陽和真陰就像某天屠龍記說的一樣不能雙修。如果你想透過系統設計提高真陽性，比如更多的實現 AEB 的剎車行為，那麼也就勢必會增加假陽性的機率，在不應該剎車的情況下，採取了 AEB 制動，這也就相應的減少真陰性的部份。在系統設計中往往只能上下移動圖中藍線的位置而改變系統的行為，要做的就是要根據具體問題取捨平衡。

同時，值得一提的還有查準率(TP/(TP+FP))和查全率(TP/(TP+FN))。在一些非常要求準確性的場景下，比如氣囊(雖然氣囊作為一個很成熟的技術並不是 SOTIF 的範疇，但是作為例子卻能很好地解釋查準率的概念)，必須在準確的時

候打開，不該開的時候萬不能開，這就需要很高的查準率。而另一邊查全率則誇張點說有「寧可錯查一千，不可放過一個」的意思(其他應用，比如洪水預測或者皮膚癌監測)，比如道路上小朋友圖片這種場景，如果自動駕駛系統智慧化程度不夠，就應該適當調高查全率，降低撞上真正小朋友的事故率。

SOTIF 和場景分區

圖 6：SOTIF 的四個場景區域。

前段時間和部門裡做 FuSa 的同事聊天，她說車其實並不是公司的核心，人才是，一切都是為了人開心，也就是為了用戶體驗。對汽車特別是自動駕駛車而言，無法實現最大可能的保證安全，就會失去最基本的用戶體驗。真實世界複雜多變，我們並不能要求世界消滅所有的危險場景。但是作為車廠，透過規範流程，提高技術來增強系統的能力，將風險降到最低是可以做到的。

如**圖 6**，SOTIF 在標準 ISO PAS 21448 中將場景按是否已知，是否安全分為四類：區域 1 已知安全、區域 2 已知不安全、區域 3 未知安全和區域 4 未知不安全。已知未知這點很有意思，說的是針對汽車開發而言，而不是對駕駛人員。有些開發人員覺得，一旦未知場景被想到，不是就變成已知的了麼？可以把它想成是針對系統設計開發的時間點來界定已知未知的考量。開發人員可以變未知為已知，但是過了這個發佈時間點，系統不升級的話已知和未知的界限就定下來了。

如何設計一個能夠最小化不安全區域 2 和 3，同時最大化已知安全區域 1 的汽車系統是 SOTIF 的目標所在。**圖 6** 中可以看出區域 4 本來就是最大的，汽車的駕駛場景其實有無數種，大部份都是未知且安全的，沒有必要在流程中全部列出，安全就好。區域 1 是已知安全的場景，這些場景出現在車廠的流程中，可能還被測試過，證明是安全的，所以也是好的，多多益善。

剩下的兩個是不好的區域了，區域 2 已知不安全的場景，比如前面說的眩光、暴雨、積雪等，要麼透過提升系統解決，要麼通知駕駛員結束自動駕駛來避免，系統至少需要辨識這些已知不安全的場景。這個區在開發流程中要盡可能地探索測試，轉化為區域 1，在需求分析中這塊區域是已知的。而最為棘手的是區域 3，就像我們知道黑暗中有危險，但是我們不知道或者沒想過這危險是什麼。例如**圖 7** 中所示如果汽車突然在公路上碰到了「恐龍」，到底會辨識成

什麼，動物？非機動車？在某些國家的法律上，公路上碰到有些動物是不能亂剎車的，符合法規的自動駕駛車會撞過去嗎？SOTIF 的目的就是上下求索的將區域 2 和 3 降低到一個可以接受的程度。

圖 7：未知不安全的馬路「恐龍」。

SOTIF 的流程

ISO PAS 21448 中對 SOTIF 的方法論和流程定義如**圖 8** 所示，主要分為三大部份。首先，透過分析評估確定 SOTIF 相關的危害和風險，當風險不被接受時進一步分析辨識風險的觸發誘因。然後，在定義驗證和確認策略(Verification & Validation)的基礎上評估已知危害場景區域 2。最後，更進一步對未知危害場景區域 3 進行評估、驗證確認，直到風險被控制在一個可以接受的範圍後結束 SOTIF 流程。之前講過 SOTIF 是作為車輛智慧化時代對 FuSa 的補強而產生

的，SOTIF 的流程也可以嵌合至 ISO26262 的流程中，兩種標準協作實現更好的功能安全。SOTIF 在方法論上很多地方也和 FuSa 相似，比如都是基於嚴重程度、接觸機率和可控性，只是在 SOTIF 中並沒有汽車安全完整性等級 ASIL。

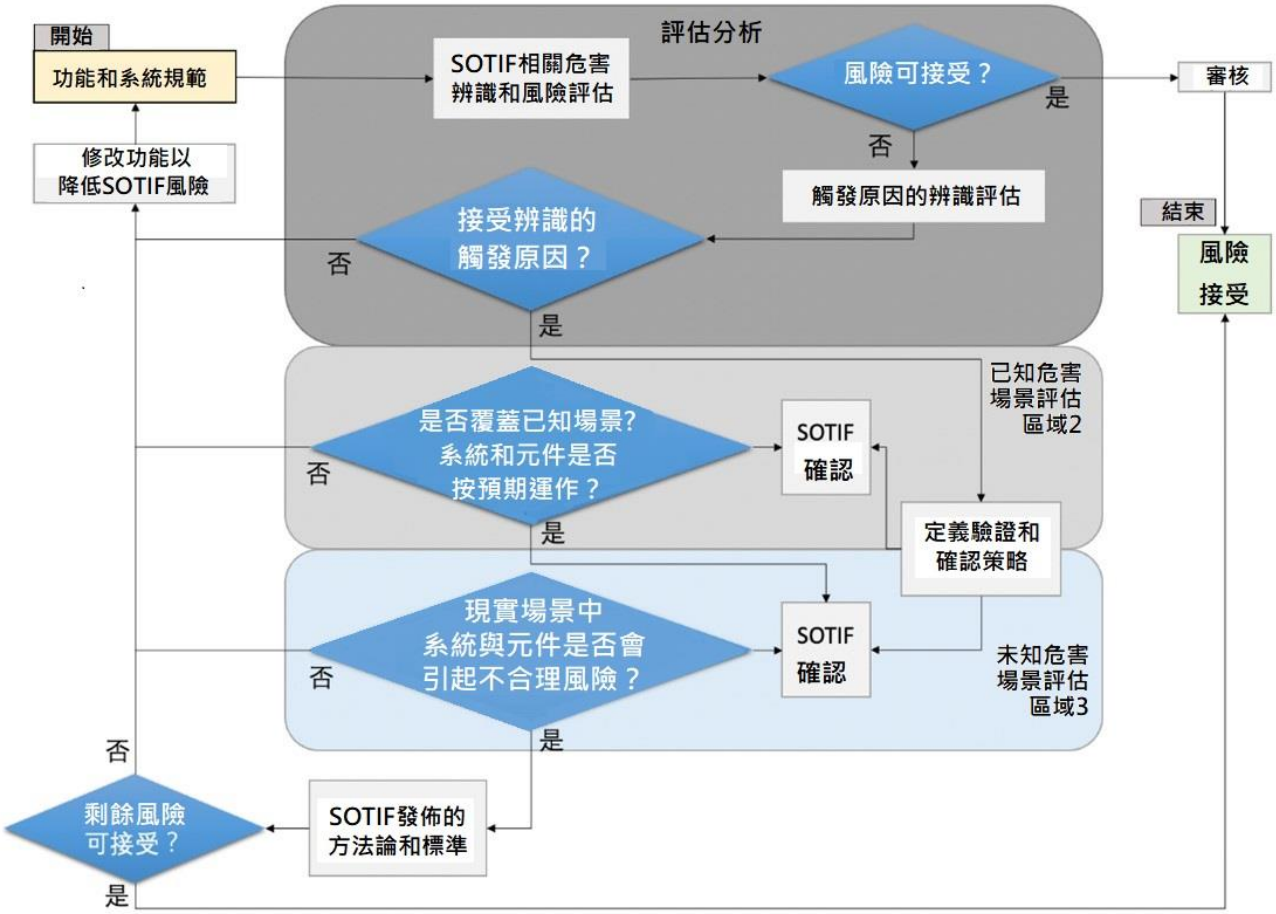

圖 8：SOTIF 流程。

汽車安全三連的後繼者

還是那句話，Safety、Security 和 SOTIF 這安全三連各自不同，但依然相輔相成，有著絲絲縷縷的聯繫。在車電領域新四化的大背景下，如果將汽車安全相關的標準和方法論完美地吸收利用，不但會提高車輛設計開發的效率，減少未

來成本，更能大幅地提高汽車產業的用戶體驗。畢竟「道路千萬條，安全第一條」嘛。不過隨著汽車產業發展反覆運算不停，新技術日新月異，本文所提到的這安全三連依然並不完善。

一切只是一個開始，更多的汽車安全相關，特別是自動駕駛車相關的標準已經開始浮現。未來的汽車安全之路將出現更多新的角色，比如 ISO TR4804 (道路汽車有關功能安全、資訊安全的自動駕駛方向的設計和確認)、ISO24089 (軟體升級工程)、ISO PAS 8800 (車載 AI)和 ISO PAS 5112 (資訊安全審計)等等。

未來的汽車安全標準和方法論必將越來越完善，汽車安全相關的設計和開發也將越來越規範。但是，總有個但是，如果標準未被納入法規，再好的標準也是個可選項，也正如我開頭所說，利益驅動下，汽車產業的新興車廠並不一定願意恪守「結硬寨、打呆仗」的形式，而更願意講究打破「枷鎖」唯快不破。究竟誰對誰錯哪？我不敢妄執可否，或許只有時間才會給我們答案……