

Artikel

Dark data: Wie sich die Nutzung übersehener Informationen als entscheidend für die F&E in der Fertigung erweisen kann

Version 1.2

Dr. Dirk Ortloff

17. Juli 2020



© DO-IT-Service GmbH
<https://www.do-it-service.de/>

Dieser Artikel zeigt die Bedeutung und den Lösungsansatz für die Bewältigung der Herausforderung durch Dark Data in Forschung und Entwicklung. Er zeigt die Auswirkungen, die hohen Kosten und die Projektverzögerungen, die durch ungenutzte Daten bei der Entwicklung von Fertigungs-technologien verursacht werden.

1 Einführung

In Zeiten wachsender internationaler Konkurrenz und schrumpfender Marktnischen wird die Innovation durch neue oder verbesserte Produkte zum Schlüssel zum Erfolg oder sogar zum Überleben. Bei der Entwicklung eines neuen Produkts oder einer neuen Iteration eines bestehenden Produkts ist es von entscheidender Bedeutung, Informationen aus einer Vielzahl von Quellen zu nutzen, um den Prozess mitzugestalten und den Erfolg zu sichern.

Der Zugang zu historischen Informationen, Wissen und Weisheit ist jedoch auch von entscheidender Bedeutung, um sicherzustellen, dass jedes neue Projekt von allen zuvor abgeschlossenen lernen kann. Erfahrungen, die aus bestehenden Herangehensweisen, früheren Forschungs- und Entwicklungsarbeiten, wissenschaftlichen Abhandlungen und alten Laborbüchern gewonnen wurden, liefern den Hauptbeitrag zur erfolgreichen Realisierung neuer Herstellungsverfahren und Produkte. Lektionen können gelernt werden, Erkenntnisse können auf neue Projekte angewandt werden, und sogar Misserfolge oder Fehler können entscheidende Hinweise dafür sein, wie ein Thema in einem zukünftigen Projekt nicht angegangen werden sollte.

2 Aufbau einer informationsgetriebenen F&E-Organisation

Für Firmen des produzierenden Gewerbes ist der Zugang zu historischen Informationen, Wissen und Weisheit der entscheidende Faktor für die Gewährleistung der Prozess- und Produktqualität während der gesamten Forschungs- und Entwicklungstätigkeit (F&E). In verschiedenen High-Tech-Branchen, z.B. Halbleiter, MEMS und Nanomaterialien, liefern die Erfahrungen und Informationen aus früheren Projekten die wesentlichen Erkenntnisse bei der Planung und Durchführung eines neuen Entwicklungsprojektes.

Mit dem Aufkommen von Industrie 4.0 wurde auch das Konzept des digitalen Zwillings eingeführt: eine virtuelle Version einer Maschine, eines Prozesses oder eines Produkts, der sich aus den Daten zusammensetzt, die das physikalische Pendant erzeugt. Zusammen mit einem oder einer Satz von Modell(en) dafür, wie sich das reale Objekt verhält, funktioniert, interagiert und sich entwickelt, kann ein digitaler Zwilling eine enorme Hilfe sein, um fundierte Entscheidungen zu treffen. Zum Beispiel kann dies bei Maschinen einen unschätzbar wertvollen Kontext dafür liefern, wie Maschinen am effizientesten arbeiten, und dabei helfen, die zur Erreichung eines bestimmten Ziels erforderlichen Entscheidungen zu treffen.

Allzu oft werden diese verschiedenen Arten von Daten jedoch nicht aufgezeichnet, geteilt oder verwendet - was den Begriff "Dank Data" hervorgerufen hat. Bedeutende Information werden entweder ignoriert, nicht am richtigen Ort oder im richtigen Format gespeichert oder einfach gar nicht erst erfasst. Abgesehen davon, dass es eine Verschwendug von wertvoller Zeit und Ressourcen ist, kann das Nicht-Betrachten oder Lernen aus diesen Daten schwerwiegende Auswirkungen auf die Effizienz eines Unternehmens und insbesondere auf seine Produktions- und F&E-Aktivitäten haben.

3 Die Entstehung von "dark data"

Die Zuverlässigkeit und Zugänglichkeit dieser Datenquellen ist jedoch nicht immer gewährleistet. Kollegen können das Unternehmen verlassen und keine vollständigen oder nur schwer nutzbare Aufzeichnungen hinterlassen, was bedeutet, dass nicht immer klar ist, welche Experimente durchgeführt wurden und was die Ergebnisse und die Schlussfolgerungen daraus waren. Laborbücher sind eine großartige Quelle für historische Daten, sind aber oft nur für diejenigen nützlich, die sie geschrieben haben. Was die Computerdateien betrifft, so können diese auf mehrere Dateiserver, Desktops und Laptops verteilt sein. Jeder Ingenieur hat seine eigene Art, Informationen zu speichern und aufzubewahren, so dass verschiedene Softwarepakete verwendet werden und es oft keine Konsistenz in der Art der Datenspeicherung gibt. Dies führt zu einer Anhäufung von Dark Data und damit zu einem Mangel an Wissen.

Studien zeigen [Barkai 2012], dass bei F&E-Aktivitäten etwa 40% der Experimente wiederholt werden. Experten in der Entwicklung von Halbleiterprozessen schätzen, dass 10-15 % der fehlgeschlagenen und doppelt durchgeführten Experimente vermieden werden könnten, wenn frühere Ergebnisse leichter zugänglich wären. Wenn Ingenieure von einem Projekt zu einem anderen wechseln, besteht zudem die Gefahr, dass Experimente gefährdet werden, wenn ein neuer Ingenieur mit einer Flut unstrukturierter und ungewohnter Daten konfrontiert wird.

Bei der Dokumentation von Experimentdaten in F&E-Projekten sollte der Schwerpunkt immer auf einem kollaborativen Wissensmanagement liegen, das einen Multi-Perspektiven-Zugang und klare aber flexible Suchkriterien ermöglicht. Beim Aufbau dieser Strukturen kann es jedoch zu Problemen kommen, da Daten teils in strukturierter Form (Tabellen, Zahlen, Einheiten) oder teils in unstrukturierter Form (Bilder, E-Mails, Dokumente) vorliegen können. Wie Studien ergeben haben, liegen bis zu 80% der digitalisierten Daten einer Firma unstrukturiert vor.

Selbst die strukturierten Daten können sich ständig oder täglich ändern, und wenn alte oder veraltete Informationen nicht gelöscht, archiviert oder korrekt kategorisiert werden, können Inkonsistenzen und Verwirrungen entstehen. Auch eine Suche in den Daten kann problematisch sein, da sich die Suchkriterien und die Dokumentation von Projekt zu Projekt und sogar innerhalb eines Projekts ändern können, je nach Aufgabe. Hinzu kommt, dass die Volltextsuche oft nicht ausreicht, um die benötigten Daten genau genug zu finden. Eine solche Suche kann auch nicht den Kontext der Daten berücksichtigen, z.B. wie die Ergebnisse erzielt wurden, wo sonst das gleiche Material oder die gleichen Bedingungen verwendet wurden oder wie ein bestimmtes Bauteil hergestellt wurde.

Diese Anhäufung von Dark Data führt zu dem unerwünschten Ergebnis, dass bestimmte Informationen nur innerhalb des Lebenszyklus eines Bauteils oder Projekts verwendet werden. Dies schränkt das Lernen für zukünftige Projekte ein, da nur diejenigen, die am Projekt beteiligt waren oder die Daten selbst erfasst haben, entweder auf sie zugreifen oder ihren Kontext verstehen können.

4 Lösung des Problems der Dark Data

Trotz dieser Herausforderungen gibt es Strategien und Prozesse, die von Organisationen eingeführt werden können, um auch aus Dark Data Kapital schlagen zu können und bei der Arbeit an neuen F&E-Projekten erhebliche Effizienzgewinne zu erzielen. Die IT und Technologie, zu der wir jetzt Zugang haben, ermöglicht ausgeklügelte und intelligente Systeme zur Aufzeichnung, zum Zugriff und zur

Sortierung von Daten, so dass diese Probleme leicht überwunden werden können.

Auf der einen Seite ist die Tatsache, dass wir mehr Daten als je zuvor speichern können, ein großer Vorteil. Da jedoch die Menge, die wir speichern können, jedes Jahr um etwa 50% zunimmt, besteht die Gefahr, dass wir in diesen Daten ertrinken, wenn sie nicht korrekt kategorisiert und in einer Weise aufgezeichnet werden, die sie für zukünftige Projekte nützlich macht.

Die Lösung besteht darin, einen intelligenteren Ansatz für die Dokumentation und Verwaltung von Daten zu finden und einen Rahmen für gemeinschaftliches Lernen zu schaffen, der sich an den F&E-Prioritäten orientiert und schnellere, billigere und effizientere Wege ermöglicht, neue Produkte auf den Markt zu bringen.

Damit die Daten für F&E gut funktionieren, muss ein solches System in der Lage sein, sowohl strukturierte als auch unstrukturierte Daten zu speichern, die Beziehungen zwischen allen Daten zu verwalten und jedem Eintrag und jeder Verknüpfung eine spezifische Beschreibung und explizite Bedeutung zu geben. In vielerlei Hinsicht ähnelt es damit einer "Single Source of Truth"-Lösung, da es Ihnen ermöglicht, alle Ihre Daten an einem Ort zu verwalten, auch wenn sie physikalisch an unterschiedlichen Stellen gespeichert sind.

Außerdem muss eine solche Datenplattform einen Audit-Trail bereitstellen, der detailliert aufführt, welche Änderungen wann und warum vorgenommen wurden, egal ob es sich um strukturierte oder unstrukturierte Daten handelt. Sie muss auch für multidisziplinäre Arbeitsumgebungen geeignet sein, z.B. muss sie den Elektroingenieur mit elektrischen Testdaten und den Maschinenbauingenieur mit Belastungstestdaten versorgen.

Eine der wichtigsten Überlegungen bei der Suche nach einer Lösung für F&E-Daten sind die Ausgangsformate. Normalerweise werden F&E-Daten auf folgende Weise gesammelt und gespeichert:

- **Semi-strukturierte Daten in Tabellenkalkulationen, Text- und CSV-Dateien (üblicherweise MS Excel):** Die Daten können leicht mit einer Vielzahl von Softwarepaketen importiert und exportiert werden, und viele Fertigungsmaschinen exportieren ihre Daten im Excel- oder CSV-Format.
- **Dateibasierte Ergebnisdaten aus verschiedenen Metrologiewerkzeugen:** Dies gilt für digitale Versuchsdatenbilder, Analyseergebnisse und Diagramme. Der Kontext kann hier leicht verloren gehen, da die Beziehungen zwischen den Datenpunkten komplex und mit den Suchfunktionen vieler Systeme nicht kompatibel sind.
- **Bestehende Labor-Datenbanken:** Sammeln von Daten aus allen Experimenten zum Zwecke der Rückverfolgbarkeit; sehr wertvoll im Hinblick auf die Optimierung von Herstellungsvorgängen und die Bereitstellung von neuen Erkenntnissen.

5 PDES: eine moderne Lösung

Obwohl alle diese Formate auf ihre eigene Art und Weise funktionieren, ist ihr Wert für den F&E-Prozess begrenzt, es sei denn, sie werden zusammengeführt und miteinander in Beziehung gesetzt, damit die Daten richtig kontextualisiert und analysiert werden können. Durch den Einsatz von PDES-Software (Process Development Execution System) [Wikipedia 2020] können diese vorhandenen Datensilos integriert und zusammengeführt werden (an einem Ort oder virtuell miteinander verbunden), um einen besser nutzbaren und aufschlussreicherem Satz historischer Informationen zu schaffen, der für zukünftige Projekte verwendet werden kann.

Die camLine XperiDesk-Software [camLine GmbH 2020] ist ein Beispiel für die PDES-Implementierung, die F&E-Aktivitäten durch das Sammeln von Daten, deren Verfolgung während ihres gesamten Lebenszykluses und die Organisation in nutzbare Informationsblöcke rationalisiert. Daten aus bestehenden Silos können durch Kopieren oder Verknüpfen in die Software integriert werden und liefern historische Einblicke, die die Leistung des aktuellen Projekts positive beeinflussen und verbessern können.

Werkzeuge innerhalb der Software ermöglichen es den Benutzern, Daten aus verschiedenen Quellen zu laden und zu verknüpfen, sie in ihrem gesamten Kontext zu verwalten und sie als verwertbare Informationen abzurufen, womit der Datenmanagementprozess einen ganzheitlichen Ansatz bietet.

Da sich die Technologien ständig weiterentwickeln und die Fertigungsmöglichkeiten immer weiter zunehmen, ist es von entscheidender Bedeutung, dass Daten aus historischen Experimenten nicht nur gesammelt, sondern so verwendet werden, dass zukünftige Innovationen ohne Zeit- und Ressourcenverschwendungen gefördert werden. PDES ist ein riesiger Fortschritt bei der Überwindung der Probleme, die durch das Phänomen der Dark Data verursacht werden. Die Bereitstellung praktisch nutzbarer, zuverlässiger Informationen aus verschiedenen Quellen ermöglichen es den Ingenieuren, ihr Arbeiten aufwandsärmer zu dokumentieren, ihr Wissen einfacher zu teilen, und Ihre F&E-Tätigkeiten effizienter zu gestalten.

Literatur

Barkai, Joe (2012): „Accelerating Science-Led Innovation for Competitive Advantage“. In: url: <http://accelrys.com/resource-center/white-papers/accelerating-science-led-innovation-request.php> (siehe S. 3).

camLine GmbH (26. Juni 2020): *XperiDesk – Streamlining Technology Development*. In: url: <https://xperidesk.com> (siehe S. 5).

Wikipedia (26. Juni 2020): *Process Development Execution System*. In: url: https://en.wikipedia.org/wiki/Process_development_execution_system (siehe S. 4).