

# Fusión y Bench - Marking: ¿Qué hemos aprendido en los últimos diez años?

Dina Raimondi  
Gilles Santini

TMPR  
IPSOS-Novaction

## 1. Introducción

Hace diez años, ya existía la teoría sobre la fusión y el bench-marking, pero muy pocos la tomaban en cuenta.

Frederic Wendt ya había impartido su plática sobre fundamentos en el Simposio Mundial de Lectores en Montreal (1983), estableciendo sus bases, y yo ya había estudiado algunos tipos potenciales de uso de la fusión en mi conferencia en SAARF en 1986.

Lucien Boucharenc ya había acumulado experiencias sobre el software de fusión IBL/AGMA y Dina ya estaba experimentando con el nuevo algoritmo de fusión IMS/FRF.

Entonces llegó el momento para que la comunidad de investigación de medios se planteara el reto de la fusión.

La duda que prevalecía era: “**¿Es la fusión una técnica decente para la investigación verdadera?**”

Aunque la respuesta no quedó muy clara durante un tiempo, pronto se diluyó la emoción y la gente empezó a experimentar y a usar la fusión y otros métodos relacionados en la práctica.

Recientemente, una nueva ola de interés ha arribado a la tierra de la investigación con preguntas más demandantes, tales como: “**¿Cuándo se debe usar y cuándo no?**” y “**¿Qué tan confiable es?**”.

Antes de intentar bosquejar el camino por recorrer, nos pareció importante compendiar lo que hemos aprendido durante los últimos diez años.

## 2. Definiciones

Ahora contamos con un conjunto de definiciones precisas:

**Adscripción:** Llamamos adscripción a cualquier operación en un archivo de estudio que le asigna a una

persona una respuesta que no fue grabada directamente de dicha persona, y que no puede deducirse de manera lógica o aritmética a partir de sus respuestas.

**Fusión:** Denominamos fusión a cualquier operación que traza a nivel individual, 1 a 1 ó 1 a n, dos muestras independientes. Los vínculos creados por el algoritmo de fusión se utilizan para adscribir a nivel individual, las respuestas que se transfieren como bloque de la muestra donante a la muestra receptora.

La fusión puede darse en una dirección o en dos direcciones.

Los métodos que usan vínculos n a 1 son de naturaleza distinta, y no deben denominarse fusión, sino bench-marking.

**Inyección:** Llamamos inyección a cualquier operación que complementa un archivo de estudio incompleto al adscribir a las personas que no respondieron bloques completos de respuestas tomadas de los que sí contestaron, gracias a los vínculos establecidos por un algoritmo de fusión.

**Bench-marking:** Conocemos como bench-marking a cualquier operación que adscribe a los individuos en un archivo de estudio información adicional de manera tal que la distribución global resulte similar a la que se observa en otra fuente, que sirve como referencia.

El bench-marking trata con información que se produce a partir de las respuestas (i.e. probabilidad o cantidades), y no con las respuestas en sí.

Utiliza técnicas como estadística, modelado, simulación Monte Carlo y correspondencia de algoritmos que producen vínculos n a 1.

Estos cuatro métodos tienen en común el hecho de basarse en **variables puente**, que se utilizan para establecer el vínculo entre la fuente donante (o primaria) y la receptora (o secundaria).



Por naturaleza, no son paramétricos y operan a nivel individual: como tales, no deben confundirse con modelado estadístico.

También debe advertirse que todos ellos nos hacen usar equipos de cómputo y software complejo, aunque la tecnología moderna de las PCs van a extender su disponibilidad.

A partir de aquí, nos concentraremos en la fusión y el bench-marking.

### 3. Prácticas Actuales

**Las expectativas y necesidades del usuario se basan en cuatro criterios principales:**

- ◆ Obtener, en todas las bases de datos, información válida sobre el público, que pueda utilizarse de la misma forma y que el mercado la considere tan efectiva como los datos recopilados de manera directa.
- ◆ Obtener datos que sean coherentes con otras fuentes en el mismo medio.
- ◆ Que sea operativo para software de planeación de medios.

- ◆ Que permita al usuario recuperar un resultado mixto válido.

Como algunos de ustedes saben, estamos usando estas técnicas en Francia, en todas su distintas formas.

Así como Gilles lo ha dicho, éstos métodos están bien establecidos en nuestra práctica actual, de manera que algunos usuarios han olvidado este paso intermedio de fusión o bench-marking entre los datos brutos y el archivo de planeación de medios que utilizan a diario.

Veamos algunos ejemplos de la aplicación de estas diferentes técnicas.

La encuesta “Cadres Actifs” de IPSOS sobre ejecutivos, que se recopila de manera formal cada año mediante un cuestionario cara a cara, con información socio-demográfica y de lectura, sólo recoge cada tercer año información detallada sobre otros asuntos, tales como computadoras y equipo de oficina, viajes, esparcimiento, etc. Estas variables tenían que fusionarse del archivo de un año al del año siguiente.

**FACTA**  
RESEARCH

*http://www.facta.net.mx*  
*info@facta.net.mx*

[5]575 16 52  
[5]575 19 22  
[5]575 00 27

[5]559 94 71  
[5]559 16 78

Contamos con dos salas para sesiones de grupo

Roberto Gayol 53  
Col. del Valle  
C.P. 03100  
México D.F.

**19 años investigando el mercado mexicano**

**ESTUDIOS CUALITATIVOS Y CUANTITATIVOS**

Esta fusión se suprimió en 1994 y se reemplazó con una inyección anual entre personas que sí habían respondido y personas que no lo habían hecho, en un cuestionario personal respondido de manera individual sobre los mismos aspectos, y la información socio-demográfica y la información sobre medios se siguió recopilando mediante un cuestionario cara a cara.

#### **También se está utilizando en bench-marking:**

◆ Para recuperar el nivel de audiencia “correcto” en una encuesta para el mercado de “multimedia” en este caso, el nivel “correcto” es el que resulta de una encuesta de audiencia de medios impresos, es decir, la “moneda” del mercado.

Cuando se recopila información sobre lectura en un cuestionario llenado de forma personal, los niveles de audiencia son distintos, lo que resulta natural. No obstante, algunos usuarios prefieren trabajar con los mismo niveles de audiencia y perfiles similares de lectores que en la encuesta de referencia de lectores

#### **Pueden surgir algunos problemas:**

◆ En el archivo fusionado de IPSOS no fue posible obtener resultados enteramente consistentes al cruzar datos de los medios (recopilados cada año), con datos producidos (por fusión cada tercer año).

Solamente se pudo sacar ventaja de las partes del cuestionario que estaban bien correlacionadas con las variables de vinculación; las demás podían producir resultados inestables.

Esto nos lleva a la introducción de restricciones funcionales más complejas.

◆ En el bench-marking, es difícil controlar la duplicación entre revistas al ajustar los niveles. Este tipo de información no deberá considerarse operativa, ya que el bench-marking se efectúa generalmente a nivel global y en una estructura de audiencia de revista por revista.

#### **La situación en otros países europeos:**

**Reino Unido.** Se realizó un experimento de fusión sobre datos NRS (de enero-junio 90 a julio-diciembre 90).

El bench-marking se aplica generalmente para ajustar los niveles de lectura de revistas recopilados en el TGI sobre resultados NRS.

**Alemania.** La fusión Agma es, en este país, la base para la planeación de “multimedia”:

- Con base en los mismos entrevistados, se hace disponible la probabilidad de exposición a televisión, radio y medios impresos.
- Se realizan dos fusiones, entre los datos de consumo Vuma y los datos de medición del panel GsK y entre el “MA elektronik tranche” sobre Radio y TV y la encuesta de medios impresos.

**Italia.** Según sabemos, la fusión y el bench-marking no se utilizan actualmente, pero se usaron anteriormente sobre datos de revistas.

**España.** AIMC (EGM) recopila datos de “multimedia”.

**Estados Unidos.** El bench-marking de niveles de lectura en el estudio de Medios y Mercados sobre los niveles obtenidos en la Encuesta de Lectores Americanos.

**Francia.** Otras aplicaciones:

- Fusión Aude (n a 1) entre el panel de audiencia Mediametrie TV y el panel de consumo SECODIP.
- Probabilidades de audiencia de radio.

## **4. Aspectos Técnicos**

Los aspectos técnicos detrás de la fusión y el bench-marking pueden resumirse bajo unas cuántas preguntas que se plantean con frecuencia:

### **■ ¿Qué es lo más importante para emprender con éxito una fusión?**

Las dos cosas más importantes a controlar antes de emprender una fusión (o, en menor grado, un proceso de bench-marking) son, en primer lugar, que las muestras donante y receptora sean comparables y que las variables puente sean genuinamente idénticas.

Por muestras comparables queremos decir que sean muestras de la misma población seleccionadas de manera similar: fusionar ejecutivos entrevistados en casa con ejecutivos entrevistados en su sitio de trabajo no es legítimo, ya que existen factores ocultos que hacen diferentes a ambas poblaciones.

Del mismo modo, aunque se ven parecidas, las preguntas sobre la frecuencia de lectura hechas cara a

cara y las preguntas sobre frecuencia de lectura en cuestionarios llenados de manera personal o por teléfono, por su naturaleza, dan niveles diferentes de lectura: es incorrecto utilizar esta información como variables puente sin un caso específico.

No debe intentarse una fusión o un proceso de benchmarking a menos que se haya verificado que las variables puente que se pretende usar den resultados casi idénticos en las muestras donante y receptora, una vez que éstas hayan sido sopesadas de nuevo para alcanzar el equilibrio en los aspectos demográficos clave.

### ■ ¿Es necesario tener muestras donantes y receptoras del mismo tamaño?

No, es lo mejor, pero no es necesario ya que la mayoría de los algoritmos de fusión y benchmarking pueden manejar tamaños desiguales de muestras.

Si la muestra donante es más grande que la muestra receptora, se presenta una situación favorable para el benchmarking, ya que las estimaciones de la información transferida tendrán varianzas menores. Sin embargo, es una situación desfavorable para la fusión, ya que podemos perder mucha de la varianza en el proceso de trazado (mapping) (se dejará fuera a muchos individuos).

Por el contrario, si la muestra donante es menor que la receptora, el benchmarking será más deficiente y la fusión funcionará bien, aunque no debe olvidarse que la muestra de menor tamaño establece el nivel de precisión.

### ■ ¿Deben equilibrarse ambas muestras?

Sí, las muestras desequilibradas afectan seriamente los algoritmos de fusión, ya que operan a nivel individual, casi siempre sin sopesarse. Aunque es menos práctico, las muestras equilibradas no son obligatorias para los algoritmos de benchmarking, ya que estos pueden manejar, en general, muestras sopesadas.

Dicho esto, las muestras estratificadas no deberán considerarse como desequilibradas, ya que en ese caso, será posible tomar cada estrato como una celda de fusión independiente.

### ■ ¿Cuál de las dos muestras deberá usarse como la muestra receptora?

La decisión deberá guiarse por cuatro consideraciones:

- La muestra receptora deberá ser la muestra de referencia del liderazgo del mercado.
- La muestra donante deberá ser, de preferencia, la de menor tamaño.
- Es más fácil transferir datos objetivos, que actitudes.
- Conforme menor número de variables transferidas, mejor.

En términos más generales, es políticamente inseguro transferir datos que se consideren como la “moneda” estándar: aun cuando sea estadísticamente posible evitar el uso de un estudio de audiencia de revistas como la muestra donante o publicar los datos de audiencia después de la fusión.

### ■ ¿Qué tipo de variable puente debe usarse?

Pueden tomarse en consideración seis tipos de variables:

- Demográficas.
- Hábitos en relación con los medios.
- Centros de interés
- Consumo.
- Geo-demográficas.
- Estilos de vida y actitudes.

Como regla empírica: como las variables demográficas son obligatorias, elija al menos una entre hábitos en relación con los medios y centros de interés, escoja al menos una entre consumo y variables geo-demográficas, y por lo general, olvide los estilos de vida.

En términos de codificación, es preferible un menor número de categorías: las dicotomías son mejores, luego serán preferibles las categorías ordenadas, siempre que sea posible, y las cantidades deberán evitarse siempre.

### ■ ¿Debemos seleccionar previamente y sopesar las variables puente?

No se ha llegado a un acuerdo en este punto: algunos especialistas consideran que el algoritmo de fusión debe ser lo suficientemente inteligente para escalar de manera automática la importancia relativa de las variables puente de entrada y para balancear las correlaciones entre ellos; otros creen que seleccionar y sopesar las variables puente es una práctica y un arte y que dicha afinación resulta valiosa.

Nuestra opinión es que, si una variable es, a priori, muy importante para el trabajo en cuestión, deberá usarse para crear una celda en donde se vaya a efectuar una fusión o bench-marking, o de manera alternativa, si el tamaño de una muestra la hace poco práctica, ésta deberá ejercer una fuerte influencia sobre el algoritmo a través de lo que comúnmente se conoce como una restricción funcional.

Si se correlaciona una variable con los datos que deseamos transferir, utilícela como una variable puente, de manera independiente a su correlación con otras variables puente.

Si una variable puede ser de interés en relación con el uso futuro del resultado del trabajo, úsela.

Si hay una variable cuyo interés para el trabajo en cuestión sea incierto, verifique su correlación con otras variables puente: si no existe mucha correlación, guárdela o descártela.

Una vez realizada esta selección un tanto prolongada, verifique cuidadosamente cada una de las posibles variables puente, dé o no salida a la misma distribución en la muestra donante y en la receptora (posiblemente dentro de cada celda de fusión, dependiendo del tamaño de la muestra). En caso que una discrepancia que no pueda corregirse mediante la recodificación, descarte la variable.

Entonces:

- Use todas las variables restantes: existe información valiosa, y probablemente oculta, en ellas.
- No sopesen las variables puente: el algoritmo sabe cómo hacerlo mucho mejor que usted.

Si algunos resultados necesitan mejorarse después de la fusión o bench-marking:

- Afine la codificación de la variable.
- Introduzca restricciones funcionales adicionales.

### ■ ¿Cómo funciona?

La base de la fusión y el bench-marking es la creación de vínculos.

Los vínculos se establecen con base en la similitud de donantes y receptores, según la medición realizada por las variables puente.

La similitud puede ser una noción tan simple como

pertenecer a la misma categoría o dar respuesta a una pregunta. Asimismo, y tal es el caso de la fusión, puede expresarse mediante una distancia en un espacio multidimensional, donde cada uno de los individuos de la muestra donante y de la muestra receptora se colocan de acuerdo con los atributos de sus variables puente.

Una vez establecidos los vínculos, la fusión transfiere directamente las respuestas de la muestra donante a la muestra receptora.

El bench-marking es un poco más complejo ya que la transferencia se hace bajo restricciones, i.e. algunos resultados globales, tales como audiencia promedio, audiencia acumulada, duplicación o el alcance de dos aspectos deben mantenerse sin variación entre la muestra donante y la muestra receptora.

### ■ ¿Cómo se establecen los vínculos?

Este es el centro del proceso. Los algoritmos simples establecen vínculos de manera tal que la distancia global entre donantes y receptores se minimice. Aunque esto parece adecuado a primera vista, no garantiza la creación de una buena solución, ya que el óptimo global puede ser la resultante de un gran número de vínculos de calidad promedio (i.e. la distancia entre donantes y receptores no es ni larga ni corta). Además, la misma muestra donante puede usarse una y otra vez, reduciendo el tamaño efectivo de la muestra, y puede romper algunas restricciones específicas.

Los buenos algoritmos apuntan hacia cuatro objetivos, al tiempo que buscan vínculos que:

- Respeten la topología local del espacio multidimensional en que se colocan los donantes y receptores.
- Produzcan una distancia global reducida entre donantes y receptores vinculados.
- Eviten el uso múltiple de la misma muestra donante.
- Logren que se cumplan las restricciones funcionales entre donantes y receptores.

Existe un trabajo pesado y crítico dentro del algoritmo de enlace, y no todos los algoritmos de fusión y bench-marking son iguales.

Aconsejamos nunca basarse en una caja negra: los métodos deberán estar publicados y ser transparentes para el usuario.

### ■ ¿Cómo verificar si funciona correctamente?

La verificación de la validez del proceso de fusión o bench-marking debe efectuarse a cuatro niveles:

El especialista a cargo del trabajo reportará el comportamiento interno global del algoritmo. Esto se hace mediante una comparación con trabajos similares y datos experimentales aprovechados a lo largo del tiempo. El comportamiento anormal de un algoritmo es una advertencia temprana de defectos potenciales, y en la mayoría de los casos, un fuerte indicador de que algo va mal.

Las distribuciones marginales y la medida “chi-square” de asociación se computarán y comprobarán de manera extensiva para asegurar la estabilidad de los datos antes y después de la fusión.

La variabilidad de los resultados posiblemente se estimará usando métodos “boot-strapping” de muestreo.

Se conducirá una campaña extensiva de pruebas de usuario con el fin de verificar que los datos producidos lleven a resultados aceptables en la práctica.

Adicionalmente a estos cuatro niveles de verificación, también puede simplemente ponerse en práctica un procedimiento “leave k-out” para generar una medida de la clasificación incorrecta. Sin embargo, debemos notar que existe un debate actual sobre qué tan adecuado es ese procedimiento, ya que el resultado de una fusión nunca debe utilizarse a nivel de entrevistado, sino sobre una sub-población o público meta de publicidad.

### ■ ¿Cuáles son los límites de la fusión y del bench-marking?

Una porción de frijoles y una porción de puerco no producen dos porciones de puerco con frijoles....

Esto mismo resulta cierto para la fusión y el bench-marking; no hay un incremento en la cantidad de información, sino tan sólo una mejor organización de la misma. y, ¡a mejor cocinero, mejor platillo!

## 5. Pros y Contras

### Ventajas de estos métodos:

◆ Garantizar una sola “moneda” para el mercado; es más fácil para todos los operadores referirse a niveles de audiencia únicos, cualesquiera que sea la fuente.

◆ Reducir la extensión del cuestionario, lo que mejora la calidad de la entrevista y reduce los costos del trabajo de campo.

◆ Permitir un mejor muestreo al dividir las muestras en encuestas independientes.

◆ Evitar el modelado del comportamiento individual.

Consideramos que es mejor tener un buen resultado de una fusión que un cuestionario demasiado largo, que lleve a resultados imprecisos.

### Desventajas:

◆ La fusión sigue siendo una técnica misteriosa y preocupante.

◆ Todas las técnicas mencionadas son sofisticadas, y consumen mucho tiempo en lo que a procesamiento se refiere, y son difíciles de comprender.

◆ Al igual que para todas las técnicas relativamente nuevas, los procesos de validación deben definirse.

◆ El margen estadístico de error de los datos proporcionados por una fusión se desconoce hasta el momento.

◆ Pueden suscitarse cuestiones legales en cuanto a la propiedad de la información.

## 6. El Camino por Recorrer

### Planee y prepárese para las fusiones:

Como sustituto para un uso correctivo, emplee una técnica establecida y planeada.

### Anticípese:

- Misma definición de muestras.
- Misma forma de conducir entrevistas.

Las encuestas cara a cara, las que se responden de manera personal y las encuestas telefónicas sobre los mismo medios, no generan los mismo niveles de respuesta, ni los mismos resultados de audiencia.

**Recopile las mismas variables de vinculación, incluyendo movilidad y momentos de presencia en casa, accesibilidad para la entrevista, hábitos de lectura, etc.**

**Tenga un marco de referencia global, como una encuesta de presupuesto de tiempo para controlar la coherencia entre los diferentes conjun-**